The Structure of Bypass of Forespore C, an Intercompartmental Signaling Factor during Sporulation in Bacillus \textsuperscript{+}\textsuperscript{S}

The Journal of Biological Chemistry Vol. 280, No. 43, pp. 36214–36220, October 28, 2005

Received for publication, June 24, 2005 Published, JBC Papers in Press, July 27, 2005, DOI 10.1074/jbc.M506910200

Hayley M. Patterson, James A. Brannigan, Simon M. Cutting, Keith S. Wilson, Anthony J. Wilkinson, Eiso AB, Tammo Diercks, Rob N. de Jong, Vincent Truffaut, Gert E. Folkers, and Robert Kaptein

From the \textsuperscript{+}Structural Biology Laboratory, Department of Chemistry, University of York, York Y010 5YW, United Kingdom, \textsuperscript{§}School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, United Kingdom, and \textsuperscript{†}Bijvoet Center, NMR Department, Utrecht University, 3584 CH Utrecht, The Netherlands

Sporulation in Bacillus subtilis begins with an asymmetric cell division giving rise to smaller forespore and larger mother cell compartments. Different programs of gene expression are subsequently directed by compartment-specific RNA polymerase \(\sigma\)-factors. In the final stages, spore coat proteins are synthesized in the mother cell under the control of RNA polymerase containing \(\sigma^K\). \(\sigma^K\) is synthesized as an inactive zymogen, pro-\(\sigma^K\), which is activated by proteolytic cleavage. Processing of pro-\(\sigma^K\) is performed by SpoIVFB, a metalloprotease that resides in a complex with SpoIIFA and bypass of forespore (Bof)A in the outer forespore membrane. Ensuring coordination of events taking place in the two compartments, pro-\(\sigma^K\) processing in the mother cell is delayed until appropriate signals are received from the forespore. Cell-cell signaling is mediated by SpoIVB and BoFC, which are expressed in the forespore and secreted to the intercompartmental space where they regulate pro-\(\sigma^K\) processing by mechanisms that are not yet fully understood. Here we present the three-dimensional structure of BoFC determined by solution state NMR. BoFC is a monomer made up of two domains. The N-terminal domain, containing a four-stranded \(\beta\)-sheet on one face of which an \(\alpha\)-helix is packed, closely resembles the third immunoglobulin-binding domain of protein G from Streptococcus. The C-terminal domain contains a three-stranded \(\beta\)-sheet and three \(\alpha\)-helices in a novel domain topology. The sequence connecting the domains contains a conserved DISP motif to which mutations that affect BoFC activity map. Possible roles for BoFC in the \(\sigma^K\) checkpoint are discussed in the light of sequence and structure comparisons.

In response to starvation, Bacillus subtilis and its relatives have the remarkable capacity to abandon growth and embark on a developmental pathway that leads to the production of dormant spores that are resistant to a variety of physical stresses. Sporulation begins with an asymmetric septation, which gives rise to two cells of unequal size but with identical chromosomes. The smaller cell is called the forespore, as it is destined to mature into the resistant spore, whereas the larger compartment is referred to as the mother cell, because it subsequently engulfs the forespore and nurtures the latter during its development. In the final stages, the mother cell lyses, and the mature spore is released into the environment where it can remain dormant indefinitely, germinating when favorable conditions for growth are restored.

A hallmark of sporulation is the utilization of a series of spatially and temporally regulated RNA polymerase \(\sigma\)-factors to effect differential gene expression from the identical chromosomes present in the forespore and the mother cell. \(\sigma^E\), \(\sigma^F\), \(\sigma^G\), and \(\sigma^K\) become activated sequentially as part of what has been termed a crisscross regulatory cascade (2). Activation of each component depends on the expression of genes under the regulation of the preceding \(\sigma\)-factor in the cascade and on signals being relayed between the two compartments. \(\sigma^G\) and \(\sigma^E\) become activated in the forespore, whereas \(\sigma^F\) and \(\sigma^K\) become active in the mother cell. The activation of these \(\sigma\)-factors is coordinated and linked to morphogenetic events to ensure that development proceeds appropriately. As a result, the sporulation process is punctuated by a series of checkpoints at which the activation of subsequent events is delayed until an appropriate set of cues has been received (3).

The last of the \(\sigma\)-factors to become activated is \(\sigma^E\). RNA polymerase containing \(\sigma^E\) (\(\sigma^E\)-factors) transcribes mother cell genes encoding spore coat proteins, as well as later-acting factors involved in the release of the spore from the mother cell and in spore germination. The \(\sigma^E\)-checkpoint ensures that \(\sigma^E\) becomes active in the mother cell only at the appropriate stage of forespore development, which is \(\sim 3\) h after the onset of sporulation (4, 5). Premature activation of \(\sigma^E\), by as little as \(30\) min, leads to aberrant spore formation (6). Transcription of sigK, which encodes \(\sigma^K\), takes place from a \(\sigma^E\)-dependent promoter and is confined to the mother cell (7). \(\sigma^K\) is translated as an inactive precursor with a 20-residue N-terminal pro-sequence that targets pro-\(\sigma^K\) to the outer forespore membrane and prevents it from binding to core RNA polymerase (5, 8). Removal of the pro-sequence to yield mature \(\sigma^K\) is catalyzed by SpoIVFB, an integral membrane metalloprotease found in the outer forespore membrane (9, 10). SpoIVFB-mediated processing of \(\sigma^K\) is under complex regulation involving mother cell and forespore-derived factors (5, 11, 12). It has been proposed that the mother cell proteins bypass of forespore (Bof)A and SpoIIFA form a ternary complex with SpoIVFB in the outer forespore membrane (see Fig. 1A).

The abbreviations used are: Bof, bypass of forespore; HSQC, heteronuclear single quantum correlation; IgG, immunoglobulin G; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy.

---

\textsuperscript{+}The work described here was funded by the European Commission as Structural Proteomics in Europe, Contract QLG2-CT-2002-00988, under the program “Quality of Life and Management of Living Resources.” The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be indicated as “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

\textsuperscript{§}The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1 and S2 and Table S1.

\textsuperscript{1} Funded by a Biotechnology and Biological Sciences Research Council studentship.

\textsuperscript{2} Funded by the Wellcome Trust.

\textsuperscript{3} To whom correspondence should be addressed. Tel.: 44-1904-328261; Fax: 44-1904-328266; E-mail: ajw@ysbl.york.ac.uk.

\textsuperscript{4} Present address: Max Planck Institute for Developmental Biology, Dept. of Protein Evolution, Spemannstr. 35, 72076 Tübingen, Germany.
ates cleavage of SpoIVFA. This relieves inhibition of SpoIVFB, which then cleaves pro-
secreted across the inner forespore membrane (H9268 to inhibit pro-
cessing, giving rise to a B of phenotype (4). This suggests that B of C is liable
at low levels under the control of an appropriate signal is received from the forespore (9).

The forespore signal is provided by SpoIVB (15), a protein expressed at
low levels under the control of σ5 and at augmented levels under σ35 control (16). SpoIVB contains a signal peptide, a PDZ domain, and a
serine protease domain (17, 18). It has been proposed that SpoIVB is
secreted into the intercompartmental space where it interacts with the
BofA-SpoIVF-SpoIVFB complex. This, in turn, leads to SpoIVB-mediated
cleavage of SpoIVF (19, 20) and relief of inhibition of pro-σk
processing by SpoIVF (21) (Fig. 1).

bofC was discovered as a gene whose deletion in a spoIIG (which encodes σp) null mutant background restores signaling of pro-σp processing,
giving rise to a B phenotype (4). This suggests that BofC is able to
inhibit pro-σk processing by SpoIVB but only when the latter is
present at the low concentrations produced by E. coli. bofC deletion in a wild-type background does not affect pro-σk processing (4).

BofC has no motifs to suggest a putative function nor does its
sequence resemble any other proteins besides BofC orthologues. To
gain further insights into forespore control of pro-σk processing in B. subtilis, we have embarked on structural studies of SpoIVB and BofC.
Here we present the solution structure of BofC determined by NMR. A full list of chemical shifts and NMR assignments have been deposited in the BioMagResBank (accession number 6731).

MATERIALS AND METHODS

Purification of BofC—The plasmid pET28BBofC was provided by Dr. T. C. Dong, Royal Holloway, University of London. pET28BBofC contains
the coding sequence for the full-length 170-residue BofC protein. For overproduction of BofC, overnight cultures of Escherichia coli BL21(DE3) pET28BBofC were used to inoculate fresh LB medium containing 30 μg/ml kanamycin. The cultures were grown at 37°C to an
OD600 of 0.6–0.7. Expression of recombinant protein was induced by
the addition of isopropyl 1-thio-β-D-galactopyranoside to a final concentration of 1 mM, and the cultures were grown at 30°C for a further
3.5 h.

A 0.1 volume of Tris-HCl at 100 mM (pH 8.5) was added to the shaking flasks, and 10 min later, the cells were harvested by centrifugation.
The cell pellet was resuspended in one-fifth of the original culture volume of a 40% sucrose, 30 mM Tris-HCl (pH 7.5), and 2 mM EDTA solution. After gentle swirling for 20 min at room temperature, the cells
were again harvested by centrifugation and resuspended in one-eighth the original volume of ice-cold water. Cellular material was removed in
a further centrifugation step, and the supernatant representing the periplasmic fraction was reduced in volume by ultra filtration (Vivascience), exchanged into buffer A (20 mM Tris-HCl, pH 8, 10% glycerol,
5 mM EDTA, 10 mM NaCl) and loaded onto a Q-Sepharose column. The
column was developed with a 10–40 mM NaCl gradient in buffer A. BofC eluted as a sharp peak at ~100 mM NaCl. The protein was subsequently purified to homogeneity by gel filtration on a Superdex 75 column in buffer A. The yield of BofC was 2–3 mg/liter of cell culture.

Labeling of Protein by 13C and 15N—For BofC structure determination
by solution NMR, an overnight culture was grown at 37°C in a
minimal medium of the following composition: 27 mM Na+, K+ phosphate (pH 6), 2 mM NaCl, 2 mM MgSO4, 0.1 mM CaCl2, 1 mg/ml
15NH4Cl, 0.4% d-glucose (for 13C 15N labeling, 0.2% d-glucose was used), 30 μg/ml kanamycin, and 1 μg/ml of each of the vitamins ribo-
flavin, niacinamide, pyridoxine monohydrochloride, and thiamine. Expression was induced at an OD600 of 0.6–0.7 by the addition of iso-
propyl 1-thio-β-D-galactopyranoside to 1 mM, and the cells were cultured
at 30°C overnight. BofC was purified as described under “Purifi-
cation of BofC.”

Protein Analysis—Electrospray ionization mass spectrometry was performed using an API QSTAR liquid chromatography/tandem mass spectrometry system. Protein masses were measured using the ExPasy ProtParam tool (us.expasy.org/tools/protparam.html).

Dynamic light scattering was performed on a ProteinSolutions DynaPro machine. Samples were centrifuged before injection into the
cells, and the data were analyzed using the Dynamics software package, version 5.

Sedimentation equilibrium experiments were conducted at 20°C on a Beckman Optima XL/A analytical Ultracentrifuge, using a Beckman
cell with a 12-mm path length in an AN-50Ti rotor. Absorbance scans
(at 280 nm) were taken at ~3 hourly intervals until sedimentation equi-
librium was achieved. The data were analyzed using the Beckman
Origin software.

NMR Spectroscopy—All spectra were recorded at 298 K on Bruker
DRX700 and DRX900 spectrometers operating at 700 and 900 MHz
proton frequencies, respectively. 15N- and 13C 15N-labeled samples
were prepared in buffers containing 20 mM NaPO4 (pH 6), 1 mM EDTA, and complete protease inhibitors (Roche Applied Science). For backbone and H3/C3 assignments, sequential (i-1) three-dimensional HNCOC, CBCA(CO)NH, and HBHA(CO)NH in combination with bifurcate
(i-1) three-dimensional HN(CA)CO, HB(CA)NH, and H(CA)HA spectra were recorded. All spectra were processed in XWinNMR (Bruker, Biospin, Rheinestetten, Germany), whereas assign-
ment was performed using PASTA (22) and AutoAssign (23) with peak
lists generated within SPARKY (www.cgl.ucsf.edu/home/sparky). Side-
chain assignment from three-dimensional C(CCO)NH total correlated
spectroscopy, H(C)CH total correlated spectroscopy (employing 15 ms of flip-flop spectroscopy-16 (25) C-C mixing each), and (H)CH correlated
spectroscopy was assisted by in-house software (to be published).
A two-dimensional H,H total correlated spectrum, employing 50 ms of YX16 mixing (26) with H-15N suppression in the direct dimen-
sion was additionally recorded for the assignment of aromatic moieties.

Distance data were derived from four three-dimensional (HQC)-
NOESY-HSQC spectra, all recorded with a 100-ms NOE evolution
time: H,N-NOESY, H,C-NOESY, (H)C,CH-NOESY and (H)C,N,
NOESY. For a verification of possible protein aggregation, a pseudo-
two-dimensional diffusion-ordered spectroscopy experiment was
measured employing double-pulsed field gradient echoes with variable
Structure of BofC from B. subtilis

gradient strength and a diffusion delay of 150 ms. Results were referenced against the known value for the water diffusion rate at 298 K ($2.3 \times 10^{-9} \text{ m}^2/\text{s}$) sampled with the same pulse sequence but with the final WATERGATE suppression module omitted.

**NOE Analysis and Structure Calculations**—Automatic NOE assignments and structure calculations were performed using the CANDID module of the program CYANA (27, 28). The quality of the structures was improved in an iterative procedure where CANDID runs were followed by manual analysis. Comparing the NOESY spectra to the preliminary structure subsequently allowed the assignment of missing resonances and improved the quality of the peak lists. Hydrogen bond restraints were defined when they were consistent with the secondary shift data, expected NOE contacts, and the calculated structure. Manual NOE peak assignments were generally not fixed in the CANDID runs but were used to create accurate spectrum-specific chemical shift lists to check the consistency of subsequent CANDID runs and to verify the manual assignments. The final CANDID run was performed using CYANA version 2.0 with Ramachandran and side-chain rotamer dihedral angle restraints in all but the last cycle. In the final cycle, fixed stereospecific assignments of prochiral groups were used when available.

The final set of NOE-based restraints determined by CANDID, in combination with restraints for 38 H-bonds and dihedral restraints for 90 residues from TALOS, were used in a water refinement run using crystallography NMR software (29) according to the standard RECOORD protocol (30). Structures were validated using WHATIF (31) and PROCHECK (32, 33).

**RESULTS AND DISCUSSION**

**Production and Characterization of Recombinant BofC**—BofC has an N-terminal signal sequence typical of proteins secreted from Bacillus by the Sec-type secretion system (34, 35). Fusion of this signal peptide N-terminal signal sequence typical of proteins secreted from Bacillus by 36216 JOURNAL OF BIOLOGICAL CHEMISTRY 34, 35). Fusion of this signal peptide N-terminal signal sequence typical of proteins secreted from Bacillus by 36216 JOURNAL OF BIOLOGICAL CHEMISTRY 36216 JOURNAL OF BIOLOGICAL CHEMISTRY this secondary structure data, a putative fifth helix (Fig. 2A), revealing two separate domains. In addition to the four helices clearly predicted by the secondary structure data, a putative fifth helix (αB) was less clearly indicated. This helix was later confirmed during quantitative structure determination, and the fold could thus be viewed as comprising two domains: one mainly β (N-terminal domain) and the other α + β (C-terminal domain) linked by a long loop partially formed by helix αB (Fig. 2).

The N-terminal domain is composed of a four-stranded β-sheet covered by an α-helix. The β-sheet has a B2-B1-B4-B3 topology as illustrated in Fig. 2A. Strands B1 and B2 and strands B3 and B4 are connected by β-turns, whereas strands B2 and B3 are joined by an α-helix that runs across one face of the β-sheet. The N-terminal domain is very well defined by a large number of NOE contacts. A sequence of 11 residues links the N- and C-terminal domains. The C terminus of this linker region includes helix αB. This helix is very short, consisting of five residues, and is separated from the core of the C-terminal domain by further unstructured residues, effectively continuing the linker. The C-terminal domain comprises three α-helices and three short anti-parallel β-strands. The β-sheet has the strand order β5-β6-β7. The three helices pack in a somewhat loose arrangement around this sheet. The
longest of the helices, αE, runs almost parallel to the sheet, and helix αD runs roughly perpendicular to the sheet with its C terminus located near to strand β5 of the sheet.

Tertiary Structure and Domain Interactions—NMR data give insight into both structure and dynamics. Often the presence of NOE contacts implies rigidity in the structure, whereas their absence indicates flexibility. From the experimental NMR data, the N-terminal domain of BofC is the better defined substructure, with an almost complete assignment of chemical shifts and a large number of NOE contacts defining the structure, with an average of 7.05 NOE contacts/residue. For the C-terminal domain, the resonance assignment is somewhat less complete, contributing to a less dense network of NOE connectivities, with an average of 4.35 NOE contacts/residue. Nonetheless, its structure is well defined, and the final structure ensemble displays very little dispersion (backbone root mean square deviation of 1.5 Å) (supplemental Table S1). Only the extensive loops connecting the elements of the secondary structure (three β-strands and three α-helices) show more substantial scattering, indicative of flexibility. Two adjacent loops near the C terminus (residues 86–100 and 119–124), located between β-strands 6 and 7 and between helices αD and αE, closely approach the N-terminal domain (Fig. 2). The loop sequence between β6 and β7 is poorly conserved, suggesting that it is not functionally important. By contrast, the sequence at the beginning of the loop between helices αD and αE is conserved; this is remarkable, as the sequences of the helices themselves are poorly conserved except for some residues at the C terminus of helix αD. These regions might form a functionally important conserved interaction surface. We cannot exclude the importance of this region for structural integrity, but no NOE contacts were found between this surface-exposed region and the N-terminal domain.

The linker region between the domains is poorly defined by the NMR data. Only a few long range NOE contacts could be unambiguously identified between the domains, i.e. between residues 19–24 (loop between β1 and β2) in the N-terminal domain and residues 97, 98, 125, 128, and 129 of the C-terminal domain. These contacts are, however, located quite close to the linker, whereas contacts farther away and thus with a larger structural "lever" are absent. Least squares superposition of the N-terminal domain structures (10–60) revealed a backbone root mean square deviation for the C-terminal domain (74–138) of 4.59 Å. Thus, the relative orientation of the domains appears quite flexible, within the steric restrictions imposed by the linker region. It could therefore be considered that both domains function in a way that does not require a specific relative orientation. To test this hypothesis, we expressed the individual N-terminal domain (1–62) and compared its $^{15}$N HSQC spectrum to that of full-length BofC under identical conditions. The substantial overall congruence in both $^{15}$N HSQC spectra proved the important point that the isolated N-terminal domain adopts the same fold as in the full-length protein. Apart from the expected shifts at the C terminus of the N-terminal domain (Lys60 and Gln61) that run into the linker region, the only significant spectral changes were in the turn between strands β1 and β2, with the largest shifts occurring for Leu20 and Asp21. This exactly correlates with the region identified as proximal to the interdomain interface from the NOE data (see above), whereas the absence of any other shift changes corroborates the conclusion that no further parts of the N-terminal domain (apart from the linker) contact the C-terminal domain. Analogous experiments with the
C-terminal domain were frustrated by the finding that this domain is insoluble when expressed in isolation. We cannot entirely rule out the possibility that minor local conformational changes occur in the C-terminal domain when present in isolation.

In conclusion, our experimental NMR data indicate that the subdomain interface is not very extensive and is largely restricted to the linker region, including loop β1-β2 in the N-terminal domain and the loop between β6 and β7 and the N terminus of αε in the C-terminal domain. Nonetheless, the observed interdomain NOE contacts near the linker limit the amplitude of possible domain reorientations and motional degrees of freedom.

The N-terminal Domain Has a Protein G-like Fold—A DALI search (41) of the protein structure data base shows that the N-terminal domain (residues 1–63) of BofC is topologically identical to the third immunoglobulin G-binding domain of protein G from Streptococcus, (Protein Data Bank (PDB) code 2igd; Z-score = 4.3, for superposition of 50 (of 61 total) equivalent C-α atoms with a root mean square deviation of 2.7 Å) (supplemental Fig. S2A). Protein G belongs to a large and diverse group of cell surface-associated proteins that bind to immunoglobulins (42). The IgG-binding domain is a highly conserved sequence of ~60 residues and this sequence similarity is reflected in a common structure. These IgG-binding domains are usually present in multiple copies within an individual protein G molecule, but they retain high affinity for IgG even when expressed individually (43).

Protein G is secreted by Streptococcus as part of a defense mechanism against phagocytosis by the host organism. By binding to the constant domains of IgG, protein G blocks the interaction of the immunoglobulin with complement proteins, preventing phagocytic cells bearing C3 receptors from taking up the complex for intracellular processing and degradation (44). This provides the bacteria with a mechanism for evading the opsonizing action of complement and increasing virulence by enhancing the capacity of the bacteria to survive in the host organism.

By analogy with protein G, we anticipate that the N-terminal domain of BofC may be a mediator of protein-protein interactions. However, the identity of the surfaces involved cannot straightforwardly be inferred. One reason for this is that structural studies have shown different surfaces on the IgG-binding domain of protein G forming interactions with IgG according to whether Fab or Fc fragments are being studied (supplemental Fig. S2). To test whether BofC binds to IgG, we carried out IgG-Sepharose affinity column chromatography. BofC was not retained on this column in contrast to protein G, which was tested in parallel. It is of further interest that the IgG-binding domains of protein G also bind to ακ-macroglobulin, a proteinase inhibitor in human plasma. This observation has led to the suggestion that protein G may be involved in proteolytic events at the cell surface (45), thus presenting another interesting parallel with the proposed role of BofC in pro-ακ processing.

The C-terminal Domain Has a Unique Topology—The C-terminal domain of BofC (residues 61–140) seems to have a new and unique fold. 

---

FIGURE 3. A, sequence alignment of BofC homologues. The alignment was generated by MultAlign (prodes.toulouse.inra.fr/multalin/multalin.html). Invariant residues have a red background, and conserved residues are boxed. The secondary structure elements of BofC are displayed above the alignment. The numbering refers to the mature form of BofC from B. subtilis. Bsu, B. subtilis; Bli, B. licheniformis; Bst, Bacillus stearothermophilus; Ban, B. anthracis; Bha, B. halodurans; Bcl, B. clausii. B, representative lowest energy NMR structure solution for BofC with the side chains of conserved residues represented as sticks; blue in cluster 1: Glu16, Gly22, Glu27, Glu31, Tyr43, Trp46, and Lys60 in the N-terminal domain; green in cluster 2: Asp21, Asp64, Ser66, Pro67, and Lys70 in the linker region; and red in cluster 3: Gly73, Gly88, Leu108, Gly119, and Ile120 in the C-terminal domain. This figure was drawn using PyMol (www.pymol.org).
No significant structural homologues were found from a DALI search of the Protein Data Bank. The highest Z-scores obtained were ≤2. The closest structural match was murine-enabled/VASP homology 1 (EVH1) domain (46), which plays a role in the spatial control of actin assembly (PDB code 1leb; Z-score = 2.0, for superposition of 43 of 111 total) equivalent C-α atoms and a root mean square deviation of 2.8 Å. The similarity occurs in the region of the three-stranded anti-parallel β-sheet (supplemental Fig. S2B). In both BofC and the EVH1 domain, an α-helix runs parallel to the face of this β-sheet. However, the intersecting elements between β-sheet and α-helix are different in the two structures. The functionally conserved residues in the EVH1 domain, required for binding its polyproline peptide ligand, lie in a region that has no counterpart in the BofC domain.

Invariant Residues Map to Three Clusters on the Structure—A BLAST search with the sequence of BoFC from B. subtilis as the search string identifies orthologues in all of the other Bacillus species whose genome sequences have been completed. BoFC was not found in the genome of Clostridium difficile in which orthologues of the other known pro-αK processing components are also missing (24). An alignment of the amino acid sequence of BoFC from B. subtilis with orthologues from Bacillus anthracis (representative of the B. anthracis, Bacillus cereus, Bacillus thuringiensis group), Bacillus licheniformis, Bacillus halodurans, and Bacillus clausii is shown in Fig. 3A. Conservation of the sequence in this type of alignment usually points to residues that play structurally or functionally important roles. Mapping of invariant residues onto the three-dimensional structure of BoFC (Fig. 3, B and C), reveals three clear clusters: cluster 1 in the N-terminal domain, cluster 2 in the segment that connects the two domains, and cluster 3 in the C-terminal domain.

Cluster 1 consists of residues exposed on the surface of the β-sheet crossed by the α-helix. Of the seven invariant residues in this cluster, it is striking that three have acidic side chains, Glu16, Glu27, and Glu31. In contrast, the invariant residues in cluster 3, within domain 2, tend to have apolar side chains and to cluster in the core of the domain, perhaps indicating a structural role.

Perhaps the most striking of the three clusters is cluster 2. Segments of polypeptide that link domains in multidomain proteins are often associated with variability in sequence. This is because their role is usually limited to tethering functional domains to one another in a single molecule. In fact, hypervariable regions within otherwise well conserved orthologous sequences are often used to identify putative domain boundaries in proteins. The strong conservation of sequence and the obvious DISP motif in the segment that constitutes the linker in the structure of BoFC suggest a more active role. It may, for example, act as a linear epitope in binding to another protein. Alternatively, it may be a recognition sequence for a specific protease. Asp21 located in domain 1 and Lys70 in the C-terminal domain are seen in the structure to belong to cluster 2, around the linker region. Lys70 may interact with Asp64 (Fig. 3D). The interdomain contacts made by these conserved residues with the linker domain restrict the relative orientation of the two subdomains with respect to each other, possibly leading to the formation of a cleft between the two subdomains formed by the C-terminal end of the loop between β6 and β7 (IQSFF), and the loop between β1 and β2 (YLDGD) and the linker region (DISP). A genetic screen identified Ser66 as an important residue for BoFC function (see next paragraph), arguing that the integrity of the linker region, and thereby the relative positioning of the two subdomains, is crucial for its function.

Mutagenesis Considerations—bofC was identified as a gene whose mutation in a spoIII (which encodes αK) null mutant background restores signaling of pro-αK processing. The first characterized bofC allele, bofC1, contains a missense mutation that results in the substitution of Ser66 of mature BoFC by phenylalanine (4). Interestingly, this serine is part of the conserved DISP motif and situated in the interdomain linker. Subsequently, a bofC insertion mutant was constructed, bofC::neo by insertion of a neomycin resistance cassette at a unique restriction enzyme cleavage site in bofC, overlapping the Ser66 codon of the open reading frame (4). Thus a protein fragment would be expressed and presumably secreted across the inner forespore membrane to produce a polypeptide consisting of residues 1–66 of the mature BoFC. The structure presented here shows that this truncated form encompasses all of the N-terminal domain and part of the interdomain linker. It is therefore likely that BoFC-(1–66) is a stable folded entity.

The bofC1 and bofC::neo alleles are distinguishable from subsequently constructed bofC null mutant in which the whole gene is deleted, in that the former are partially active as negative regulators of intercompartmental signaling of pro-αK processing. Thus, in spoIII (strains containing gerE-lacZ fusions, αK-directed expression of β-galactosidase occurs 3 h earlier and reaches higher levels in the bofC null background (bofC::neo) than in the truncated or mutated bofC alleles (bofC::neo and bofC1). This result implies that both domains of BoFC contribute to the inhibition of signaling of pro-αK processing. In the spoIII background, intact BoFC blocks signaling completely, whereas BoFC-(1–66) only delays signaling. One interpretation of the equivalent effects of the bofC1 and bofC::neo alleles is that the S66F mutation nullifies the contribution of the C-terminal domain.

The Role of BoFC in the αK Checkpoint—The role of BoFC in the αK checkpoint remains unclear. The present study reveals that BoFC is made up of two separate domains, one of which resembles an IgG-binding domain. It is tempting to use this structural similarity to infer a similarity in function and suggest that BoFC participates in protein-protein interactions following its passage through a cell membrane. Its most probable interaction partners would be other components of the pro-αK processing system with SpoIVB being the primary candidate, because it, alone among this group, is transcribed by EdJ. The levels of SpoIVB and BoFC in sporing cells suggest that the fate of each molecule is dependent on their mutual interactions (36). Thus (i) there are increased levels of BoFC in mutants unable to synthesize SpoIVB, (ii) there are decreased levels of active SpoIVB in bofC null mutants, and (iii) overproduction of BoFC inhibits SpoIVB autoproteolysis and delays pro-αK processing (36).

SpoIVB undergoes complex post-translational processing involving secretion across the inner forespore membrane, autoproteolysis in trans to release the zymogen into the intercompartmental space, and autoproteolysis in cis to produce the mature pro-αK signaling species. It has been proposed that mature SpoIVB binds the C terminus of BoFA as a prelude to cleaving SpoIVFA. This, in turn, releases SpoIVFB from its inhibition by SpoIVFA and BoFA and allows it to cleave and activate pro-αK (21). BoFC could inhibit signaling by forming a complex with SpoIVB in which interactions with other proteins are blocked or in which it acts as a competitive inhibitor of proteolysis. The implied importance of exposed residues in the BoFC linker peptide would be consistent with this mode of action.

Biochemical experiments have so far failed to establish interactions of BoFC with either the isolated PDZ domain of SpoIVB (21) or with intact SpoIVB, although performed with a mutant in which the active site serine residue is mutated to alanine (data not shown). However, the complex environment of the intercompartmental space and the flanking membranes involved in the αK checkpoint make the conditions under which these two proteins would interact extremely difficult to mimic in vitro. It is also possible that the inhibitory effect of BoFC on SpoIVB-mediated signaling depends on the presence of other components of the pro-αK processing complex. Just as PDZ domain-mediated
Structure of BofC from B. subtilis

binding of SpoI VB to Bo fA leads to the SpoI VB cleavage of SpoI VFA (20, 21), it is possible that SpoI VB–Bo fA interactions depend on the prior engagement of one or both components with SpoI VFA, SpoI VFB, or Bo fA. The observation that Bo fA is a two-domain protein with few constraints on the relative orientation of its domains is consistent with a possible function that involves interactions with a pair of protein partners.

Acknowledgments—We thank Arthur Moir (University of Sheffield) for N-terminal sequencing, Alexei Murzin (Medical Research Council Laboratory of Molecular Biology, Cambridge) for expert advice on topology, Andrew Leech (University of York) for analytical ultracentrifugation and mass spectrometry analysis and Tran Cat Dong (Royal Holloway University of London) and Jeremy Derrick (University of Manchester) for expression clones.

REFERENCES

1. Piggot, P. J., and Hilbert, D. W. (2004) Curr. Opin. Microbiol. 7, 579–586
2. Losick, R., and Stragier, P. (1992) Nature 355, 601–604
3. Stragier, P., and Losick, R. (1996) Annu. Rev. Genet. 30, 297–341
4. Gomez, M., and Cutting, S. M. (1997) Microbiology (Read.) 143, 157–170
5. Lu, S., Halberg, R., and Kroos, L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 9722–9726
6. Cutting, S., Oke, V., Driks, A., Losick, R., Lu, S., and Kroos, L. (1990) Cell 62, 239–250
7. Hilbert, D. W., and Piggot, P. J. (2004) Microbiol. Mol. Biol. Rev. 68, 234–262
8. Zhang, B., Hofmeister, A., and Kroos, L. (1998) J. Bacteriol. 180, 2434–2441
9. Resnekov, O., and Losick, R. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 3162–3167
10. Yu, Y.-T., and Kroos, L. (2000) J. Bacteriol. 182, 3305–3309
11. Cutting, S., Roels, S., and Losick, R. (1991) J. Mol. Biol. 221, 1237–1256
12. Green, D. H., and Cutting, S. M. (2000) J. Bacteriol. 182, 278–285
13. Rudner, D. Z., and Losick, R. (2002) Genes Dev. 16, 1007–1018
14. Zhou, R., and Kroos, L. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 6385–6390
15. Oke, V., Shchetkin, M., and Cutting, S. M. (1997) Mol. Microbiol. 23, 223–230
16. Gomez, M., Cutting, S., and Stragier, P. (1995) J. Bacteriol. 177, 4825–4827
17. Wakeley, P., Dorazi, R., Hoa, N. T., Bowyer, J. R., and Cutting, S. M. (2000) Mol. Microbiol. 36, 1–14
18. Hoa, N. T., Brannigan, J. A., and Cutting, S. M. (2002) J. Bacteriol. 184, 191–199
19. Kroos, L., Yu, Y.-T. N., Mills, D., and Ferguson-Miller, S. (2002) J. Bacteriol. 184, 5389–5401
20. Dong, T. C., and Cutting, S. M. (2003) Mol. Microbiol. 49, 1425–1434
21. Dong, T. C., and Cutting, S. (2004) J. Biol. Chem. 279, 43468–43478
22. Leutner, M., Gschwind, R. M., Lermann, J., Schwarz, C., Gemmecker, G., and Kessler, H. (1998) J. Biomol. NMR 11, 31–43
23. Zimmerman, D. E., Kulikowski, C. A., Huang, Y., Feng, W., Tashiro, M., Shimotakahara, S., Chien, C., Powers, R., and Montelione, G. T. (1997) J. Mol. Biol. 269, 592–610
24. Stragier, P. (2002) in Bacillus subtilis and Its Closest Relatives: From Genes to Cells (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds), pp. 519–525, American Society for Microbiology, Washington, D. C.
25. Kadkhodaei, M., Rivas, O., Tan, M., Mohebbi, A., and Shaka, A. J. (1991) J. Magn. Reson. 91, 437–443
26. Furer, I., Kramer, F., Marino, J. P., Glaser, S. J., and Lay, B. (2004) J. Magn. Reson. 166, 39–46
27. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) J. Mol. Biol. 273, 283–298
28. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) J. Mol. Biol. 319, 209–227
29. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–921
30. Nederveen, A. J., Doreleijers, J. F., Vranken, W., Miller, Z., Sproot, C. A. E. M., Nabuurs, S. B., Güntert, P., Livny, M., Markley, P. L., Nilges, M., Ulrich, E. L., Kaptein, R., and Bonvin, A. M. J. J. (2004) Proteins Struct. Funct. Genet. 59, 662–677
31. Vriend, G. (1990) J. Mol. Graph. 8, 52–66
32. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
33. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) Proteins 12, 345–364
34. Nagarajan, Y. (1993) in Bacillus subtilis and Other Gram-positive Bacteria (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds), pp. 713–726, American Society for Microbiology, Washington, D. C.
35. Tjalsma, H., Antelmann, H., Jongbloed, J. D. H., Braun, P. G., Darmon, E., Dorenbos, R., Dubois, J.-Y. F., Westers, H., Zanen, G., Quax, W. I., Kuipers, O. P., Bron, S., Hecker, M., and van Dijl, J. M. (2004) Microbiol. Mol. Biol. Rev. 68, 207–233
36. Wakeley, P., Hoa, N. T., and Cutting, S. M. (2000) Mol. Microbiol. 36, 1415–1424
37. Wishart, D. S., and Sykes, B. D. (1994) J. Biomol. NMR 4, 171–180
38. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) Biochemistry 31, 1647–1651
39. Diercks, T., Coles, M., and Kessler, H. (1999) J. Biomol. NMR 15, 177–180
40. Cornilescu, G., Delaglio, F., and Bax, A. (1999) J. Biomol. NMR 13, 289–302
41. Holm, L., and Sander, C. (1993) J. Mol. Biol. 233, 123–138
42. Boyle, M. D. P. (ed) (1990) Bacterial Immunoglobulin-binding Proteins, Academic Press, New York
43. Derrick, J. P., and Wigley, D. B. (1994) J. Mol. Biol. 243, 906–918
44. Muñoz, E., Vidarte, L., Pastor, C., Casado, M., and Vivanco, F. (1998) Eur. J. Immunol. 28, 2591–2597
45. Björk, L., and Åkerstro¨m, B. (1990) in Bacillus subtilis and Other Gram-positive Bacteria (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds), pp. 519–525, American Society for Microbiology, Washington, D. C.
46. Prehoda, K. E., Lee, D. J., and Lim, W. A. (1999) Cell 97, 471–480