THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 31, pp. 21720–21727, August 4, 2006

A Macromolecular Complex Formed by a Pilin-like Protein in Competent *Bacillus subtilis*[*][S]

Received for publication, April 28, 2006, and in revised form, June 1, 2006 Published, JBC Papers in Press, June 1, 2006, DOI 10.1074/jbc.M604071200

Inês Chen, Roberta Provvedi[1], and David Dubnau[2]
From the Public Health Research Institute, Newark, New Jersey 07103

In competent *Bacillus subtilis*, the ComG proteins are required to allow exogenous DNA to access to membrane-bound receptor ComEA during transformation. Here we describe a multimeric complex containing the pilin-like protein ComGC. Due to similarities to the type 4 pilus and the type 2 secretion system pseudopilus, we have tentatively named it the "competence pseudopilus." The ComGC multimer is released from cells upon digestion of the cell wall with lysozyme and has a heterogeneous size, estimated to range between 40 and 100 monomers, covalently linked by disulfide bonds. We determined that the prepilin peptidase ComC, the thiol-disulfide oxidoreductase pair BdbDC, and all seven ComG proteins are necessary to form the pseudopilus. Furthermore, these proteins are also sufficient to form a functional complex, *i.e.* able to facilitate binding of exogenous DNA to ComEA. The initial steps of pseudopilus biogenesis include the processing of ComGC in the cytoplasmic membrane and consist of two independent events, proteolytic cleavage by ComC and formation of an intramolecular disulfide bond by BdbDC. The other ComG proteins are required to assemble the mature ComGC monomers in the membrane into a multimeric complex proposed to span the cell envelope. We discuss the possible role of the competence pseudopilus in DNA binding and uptake during transformation.

Genetically competent bacteria are able to take up exogenous DNA and undergo transformation. In the Gram-positive *Bacillus subtilis*, the development of competence is a finely regulated process, requiring the expression of the transcriptional regulator ComK, which leads to the expression of the late competence proteins, involved in the binding, uptake, and processing of transforming DNA (1, 2).

Several late competence proteins from *B. subtilis* have been identified and characterized (3). ComEA is a membrane-bound DNA-binding protein (4), required both for DNA binding and transport (5). In addition, all seven genes contained in the *comG* operon are necessary for DNA binding (6), and it is thought that the ComG proteins allow exogenous DNA to contact its receptor ComEA (4). The ComG proteins share similarities with those involved in the formation of type 4 pili and in type 2 secretion from Gram-negative organisms. ComGA belongs to the family of traffic NTPases; ComGB is a conserved protein with several membrane-spanning domains; ComGF is predicted to be an integral membrane protein; ComGC, -GD, -GE, and -GG are similar to type 4 pilins. Pilin-like proteins are required for type 2 secretion, but because they were not known to form a pilus, they have been called "pseudopilins" (7), a term we have adopted here for ComGC, -GD, -GE, and -GG. The competence pseudopilins are made as precursors and undergo proteolytic cleavage by the prepilin peptidase ComC, also required for translocation of ComGC across the membrane (8). Finally, the protein disulfide oxidoreductase pair BdbDC is needed for the stability of the pilin-like protein ComGC in the membrane (9), probably by catalyzing the formation of the intramolecular disulfide bond in ComGC (10).

Type 4 pilins assemble into pili, long polymeric fibers that function in bacterial adhesion and motility (11). When overexpressed, secretion pseudopilins can also form pilus-like structures, presumably elongated versions of the complexes functioning in secretion, *i.e.* secretion pseudopili (12, 13). A pilus-like structure formed by the competence pseudopilins has been proposed (14), but no experimental evidence for such structure is available so far. Here we report that the major pseudopilin in *B. subtilis*, ComGC, forms a high molecular weight multimeric complex, covalently linked by disulfide bonds. By analogy with the pseudopilus of type 2 secretion systems, we have tentatively named it the "competence pseudopilus."

EXPERIMENTAL PROCEDURES

Strain Construction, Growth, and Transformation Assay—B. subtilis strains used in this work were derived from strain BD630 (*hisA1 leu-8 metB5*) and are listed in Table 1. Most strains contain the gene *comS* in a multicopy plasmid, which results in a higher proportion of competent cells in the population (15). Strains were routinely grown through the competence regimen (16). For strains with inducible promoters, 1 mM IPTG[3] or 1% xylose were added to the medium. To measure transformation frequency, cells grown through the competence regimen were transformed with DNA from a *leu*+ strain (BD170) and selected for leucine prototrophy.

To express the competence genes *comC*, *comG*, and *bdbDC* from IPTG-inducible promoters, fragments containing the 5′

* This work was supported by National Institutes of Health Grant GM43756. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[S] The on-line version of this article (available at http://www.jbc.org) contains additional text and references.
[1] Present address: Dipartimento di Biologia, Università degli Studi di Padova, Padova 35131, Italy.
[2] To whom correspondence should be addressed: Public Health Research Institute, 225 Warren St., Newark, NJ 071103. Tel.: 973-854-3400; Fax: 973-854-3401; E-mail: dubnau@phri.org.

[3] The abbreviations used are: IPTG, isopropyl 1-thio-β-D-galactopyranoside; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

TABLE 1
Strains used in this work

Strain number	Relevant genotype	Reference
BD2528	pUB110-*comS*, or multicopy *comS* (mcS)	15
BD2780	mcS *comG12*-Em	39
BD2708	mcS *comC*-Em	10
BD2718	mcS *comE*-Em	4
BD2996	mcS *comF*-Em	40
BD2999	mcS *bdbC*-Em	9
BD3002	mcS *bdbDC*-Em	9
BD3355	mcS *bdbDC*-Em *amyE::pxyl-bdbC*	9
BD3844	mcS *comC*-Cm	This work
BD3845	mcS *bdbDC*-Em *comC*-Cm	This work
BD2786	mcS *comG*ΔA	6
BD2787	mcS *comG*ΔB	6
BD2789	mcS *comG*ΔD	6
BD2790	mcS *comG*ΔE	6
BD2791	mcS *comG*ΔF	6
BD2709	mcS *comGG*	41
BD4140	mcS *comGA*-ATP$^-$	25
BD3944	mcS *amyE::comGC*-wt *comG12*	This work
BD3942	mcS *amyE::comGC*-C41S *comG12*	This work
BD3943	mcS *amyE::comGC*-C81S *comG12*	This work
BD3945	mcS *amyE::comGC*-C41/81S *comG12*	This work
BD3999	mcS *amyE::comGC*-C41/81S *comG comC*	This work
BD2121	*comK*	42
BD3502	*comK* pSPAC-*comC* pSuperSPAC-*comG* pSPAC-*bdbDC*	This work
BD4148	*comK amyE::pxyl-comEA*	This work
BD4149	*comK* pSPAC-*comC* pSuperSPAC-*comG*	This work
BD4164	*comK amyE::pxylA-comEA* pSPAC-*comC* pSuperSPAC-*comG*	This work

region of each gene or operon were amplified by PCR and cloned into plasmids containing pSPAC (17) for *comC* and *bdbDC*, or pSuperSPAC, a kind gift from David Rudner, for *comG*. These constructs were used for Campbell-like integration at the native loci of the aforementioned genes. Expression of inducible *comEA* was achieved from an ectopic locus, *amyE*, under the control of the xylose-inducible promoter p*xylA* (15). These constructs were introduced into *B. subtilis* strain BD630 and combined to obtain the strains described above, with *comK* moved in at the last step.

Mutagenesis of *comGC* to substitute the cysteine residues by serine was performed by overlapping PCR (18). The mutated genes were ectopically expressed from the *amyE* locus, using the *comG* promoter as previously described (6). To assess transformability, the constructs were moved by transduction into a strain containing an in-frame deletion within *comGC*, BD2788 (19). For immunoblot analysis of the mutant ComGC forms, a complete knock-out of the *comG* operon was moved in.

All constructs generated by PCR were confirmed by DNA sequencing. Details on the constructions are provided in the Supplemental Material.

Preparation of Subcellular Fractions and Immunoblotting—Cells grown through the competence regimen were harvested and resuspended in protoplast buffer (25% sucrose, 50 mM Tris-HCl, pH 8, 50 mM NaCl, 5 mM $MgCl_2$) with 200 μg/ml lysozyme. Usually, a 10-ml culture was resuspended in 500 μl of protoplast buffer. After 30–45 min at 37 °C, protoplasts were centrifuged at 10,000 × *g* for 10 min, and the protoplast supernatant was carefully removed.

For membrane protein preparations, protoplasts were osmotically lysed by resuspension in 1 ml of 50 mM Tris-HCl, pH 8, 50 mM NaCl, 5 mM $MgCl_2$, with DNase I, RNase, and protease inhibitors. After 10 min at room temperature, membrane vesicles were pelleted for 20 min at 16,000 × *g*, and membrane proteins were extracted for 1 h at room temperature with 1% Triton X-100 in 100 μl of the same buffer. Insoluble material was removed by centrifugation at 16,000 × *g* for 20 min at room temperature, and the supernatant containing solubilized membrane proteins was transferred to a new tube. Protein concentration was measured using the BCA assay (Pierce).

For SDS-PAGE and immunoblot analyses of protoplast supernatant preparations, 10 or 20 μl were loaded on 12.5% acrylamide gels in Tris-Tricine buffer (20). For immunoblots of membrane proteins, 16% acrylamide gels were used, and different amounts of protein were loaded, as indicated in the figure legends. To reduce samples, 20 mM β-mercaptoethanol or 5 mM dithiothreitol were used. After SDS-PAGE, proteins were transferred to nitrocellulose filters, probed with anti-ComGC rabbit antiserum, and detected by chemiluminescence.

Analytical Gel-Filtration Chromatography—Sephacryl S-500HR (GE Health Sciences) was equilibrated with 50 mM Tris-HCl, pH 8.0, 0.3 M NaCl and packed into a 50- × 1.5-cm column, with 1 ml/min flow at room temperature. These same conditions were used for calibration and analytical runs. The total volume (V_t), including column and tubing, was 90.5 ml. The void volume (V_o) was 30 ml, determined by injection of fixed bacteria. For calibration of the column, the following proteins were used as standards (with their Stokes radii and V_e, mean elution volumes from two measurements): aldolase (4.8 nm and 73 ml), catalase (5.2 nm and 69.5 ml), ferritin (6.1 nm and 66.5 ml), thyroglobulin (8.5 nm and 62.5 ml), IgM (12.5 nm and 58.5 ml), and myosin (20 nm and 31.5 ml). For each protein, K_{av} was calculated from the equation $K_{av} = (V_e - V_o)/(V_t - V_o)$. The Stokes radius of the ComGC complex was determined by interpolation using a calibration curve constructed by plotting the Stokes radii of the standard proteins *versus* $(-\log K_{av})^{1/2}$. For analytical assays, 1 ml of protoplast supernatant was loaded into the column, and 6-ml fractions were collected, precipitated with trichloroacetic acid, and analyzed for the presence of ComGC by SDS-PAGE and immunoblotting. For chro-

matography under reducing conditions, the column was equilibrated and run with 2 mM dithiothreitol added to the buffer.

Sucrose Gradient Ultracentrifugation—Protoplast supernatant preparations were diluted to 12.5% sucrose in 50 mM Tris-HCl, pH 8, 0.3 M NaCl, and loaded on top of a 12-ml isokinetic 15–45% sucrose gradient in the same buffer. Samples were run for 7 h at 37,000 rpm in a SW41 rotor at room temperature; 1-ml fractions were collected from the top, precipitated with trichloroacetic acid, and analyzed for the presence of ComGC by SDS-PAGE and immunoblotting. The sedimentation coefficient of the ComGC complex was determined by interpolation using a calibration curve, constructed with proteins standards run in the same way; protein concentration in the fractions were measured by Bradford assay (Bio-Rad). The protein standards (and their sedimentation coefficient values) were aldolase (7.3 S), catalase (11.3 S), and thyroglobulin (19 S).

Estimation of Molecular Weight—The values for Stokes radius (R_s, in nanometers) and sedimentation coefficient (s, in Svedberg units, 10^{-13} S) obtained experimentally were used to estimate the molecular weight of the ComGC complex in the protoplast supernatant, as previously described (21), with the equation, $M = 6\pi\eta N s R_s/(1 - \upsilon\rho)$, where M is molecular mass, η is the viscosity of the medium ($\eta = 0.01$ g/(cm·s)), υ is the partial specific volume of the protein ($\upsilon = 0.73$ cm^3/g), ρ is the density of the medium ($\rho = 1$ g/ml), and N is Avogadro's number, leading to, $M = 4.2 \times s \times R_s$ (in kilodaltons).

DNA Binding Assay—A 2.5-kb DNA fragment was radiolabeled by PCR with [α-^{32}P]dCTP, with specific activity of $\sim 5 \times 10^6$ cpm/μg of DNA. Cells ($\sim 1.0 \times 10^9$ colony forming units) were incubated with 1 μg/ml DNA for 15 min at 37 °C. The cells were centrifuged, washed three times with 1× Spizizen salts (22), and cell-associated DNA was measured in a scintillation counter.

RESULTS

The Pilin-like Protein ComGC Forms a Polymeric Complex—It was previously shown that the pilin-like protein ComGC can be found in membrane fractions as well as associated with the cell wall (10). The latter form is released from the cells by lysozyme treatment, resulting in formation of protoplasts. When the protoplasts are removed by centrifugation, ComGC is found in the protoplast supernatant fraction. By immunoblot analysis of ComGC in this fraction, several bands forming a "ladder" pattern could be observed under non-reducing conditions (Fig. 1, *left panel*). The predicted molecular mass of mature ComGC is 10.3 kDa, thus each band corresponds to monomer and higher molecular weight forms of ComGC (dimer, trimer, tetramer, and so on). A significant amount of ComGC signal was at the top part of the gel, not resolved into individual bands. All higher forms collapsed into a single band (corresponding to ComGC monomer) when the samples were treated with reducing agents. These results indicate that ComGC forms high molecular weight multimers, stabilized by the presence of disulfide bonds between the monomers under the denaturing conditions of SDS-PAGE. The amino acid sequence of mature ComGC contains two cysteine residues, and these could participate in formation of such intermolecular

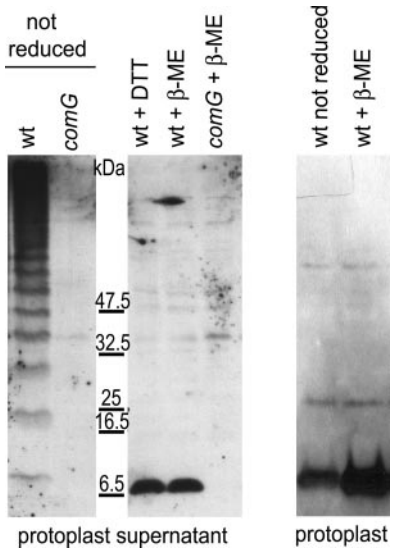

FIGURE 1. **The pilin-like protein ComGC forms a multimeric complex associated with the cell envelope.** *Left panel*, anti-ComGC immunoblots with protoplast supernatant fractions (equivalent to 200 μl of culture) prepared from wild-type strain BD2528 (*wt*) and *comG* mutant BD2780. Samples, with or without treatment with reducing agents β-mercaptoethanol (*β-ME*) or dithiothreitol (*DTT*), were separated by SDS-PAGE. *Right panel*, anti-ComGC immunoblots from protoplast fractions (equivalent to 100 μl of cell culture) from wild-type strain BD2528, with or without reduction with β-mercaptoethanol.

bonds. The multimeric forms of ComGC can also be observed in protoplast supernatant fractions from cells treated with *N*-ethylmaleimide (not shown), an alkylating agent that attacks free cysteine residues. This demonstrates that the intermolecular disulfide bonds occur *in vivo* and are not artifactual. In Fig. 1, the monomer bands in the lanes with and without added reductant show slightly different mobilities. This could be due to the presence of an intramolecular disulfide bond in the protoplast supernatant monomer.

When the protoplast fraction was analyzed by immunoblotting, only the band corresponding to ComGC monomer was observed (Fig. 1, *right panel*). However, there is a clear increase in the intensity of the signal upon reduction of the samples, suggesting that multimeric ComGC remains associated to the protoplasts. Bands corresponding to higher forms could not be seen in the protoplast fraction nor in whole cell extracts (not shown), perhaps due to the abundance of other proteins that could interfere with immunodetection.

The ComGC Complex Has a High Molecular Weight—The SDS-PAGE/immunoblot results shown above indicate that the ComGC multimer released from protoplasts is heterogeneous, with high molecular weight forms (Fig. 1). These characteristics were confirmed by experiments performed under native conditions. By analytical gel filtration, we determined the Stokes radius of the ComGC complex to range from 4 to 25 nm, with the bulk of the material between 5.5 and 11.7 nm (Fig. 2*A*). The elution profile of ComGC did not change when gel-filtration chromatography was performed under reducing conditions (not shown), suggesting that protein-protein interactions other than disulfide bonds contribute to the structure of the released ComGC complex. It should be noted that these non-covalent interactions would be disrupted under the denaturing condi-

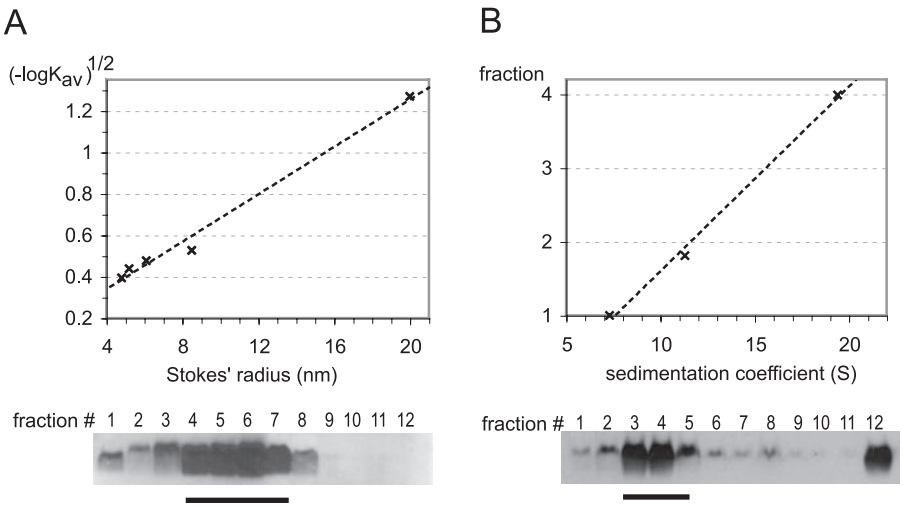

FIGURE 2. **Determination of hydrodynamic parameters of the ComGC complex.** *A*, gel filtration. The calibration curve for a Sephacryl S-500HR column was established as described under "Experimental Procedures." Protoplast supernatant preparations from wild-type strain BD2528 were loaded and fractionated, and 6-ml fractions were collected and probed with anti-ComGC antibody after SDS-PAGE. The *line* underneath the immunoblot indicates the fractions considered to contain most of the ComGC complex. In the various fractions, which were concentrated by trichloroacetic acid precipitation prior to loading, the abundance of other proteins sometimes resulted in distortions of the *lanes* in SDS-PAGE. *B*, sucrose gradient centrifugation. Samples were run through a 12-ml 15–45% sucrose gradient. The calibration curve was established as described under "Experimental Procedures." Protoplast supernatant preparations were loaded on the top of the gradient and centrifuged; 1-ml fractions were collected from the *top*, precipitated with trichloroacetic acid, and analyzed for the presence of ComGC by immunoblotting. The *line* underneath the immunoblot indicates the fractions considered to contain most of the ComGC complex. The signal at the bottom of the gradient (fraction 12) was considered to be precipitated ComGC material.

tions of SDS-PAGE. By ultracentrifugation in sucrose gradients, ComGC could be found in almost all fractions along the sucrose gradient column. Most of the ComGC signal was present in fractions corresponding to particles with sedimentation coefficient between 15 and 20.6 S (Fig. 2*B*), although some material was found at the bottom of the gradient (fraction 12), presumably precipitated. Both in gel-filtration and sucrose gradient analyses, the wide distribution of ComGC among the fractions points toward the polydisperse nature of the ComGC complex released from the cells by lysozyme treatment.

These two experimentally obtained hydrodynamic parameters, Stokes radius and sedimentation coefficient, were used to estimate the molecular weight of the ComGC complex, resulting in a range of 400–1000 kDa. This is roughly equivalent to 40–100 subunits. Such a complex is evocative of supramolecular structures formed by orthologs of ComGC, the type 4 pilus and the type 2 secretion pseudopilus. In light of these similarities, we tentatively propose the term competence pseudopilus for the ComGC complex, although the full validation of this term awaits fuller characterization of the structure of the complex. We will refer to the fraction released from the cells by lysozyme as the pseudopilus, although this may be a broken or incomplete structure, as suggested by the multimeric ComGC associated with the protoplasts (Fig. 1, *right column*).

Genetic Requirements for ComGC Complex Formation—We asked which competence factors are necessary to form the competence pseudopilus. ComGC is initially synthesized as a precursor, and is cleaved by the prepilin peptidase ComC (8). ComGC is also oxidized (10), presumably by the action of the putative protein oxidoreductase pair BdbDC (9). As expected, mutations in either *comC* or *bdbDC* abrogated formation of the ComGC complex (Fig. 3*A*). In contrast, the products of late competence operons *comE* and *comF*, necessary for DNA binding/uptake (23, 24), are dispensable for formation of the pseudopilus (Fig. 3*A*).

The *comG* operon contains the gene for ComGC and six other open reading frames, each of which is required for DNA binding and transformation (6). Non-polar mutations in any of the other *comG* genes result in the absence of multimeric ComGC from the protoplast supernatant (Fig. 3*B*, *upper panel*), and only the band corresponding to ComGC monomer could be detected, in amounts that varied from experiment to experiment. A point mutation in the Walker A motif of ComGA eliminates transformation (25) and results in the absence of complexed ComGC from the protoplast supernatant, indicating that the ATP-binding site of ComGA is required to form the pseudopilus. Upon longer exposures, a faint ladder pattern could be detected in immunoblots of *comGD*, *-GE*, and *-GF* mutants (not shown); in fact, these strains also exhibit residual transformation, albeit at least 10,000-fold less than a wild-type strain. Because the mutations in *comGD*, *-GE*, and *-GF* are small in-frame internal deletions (6), it is possible that the truncated proteins have residual activity, allowing formation of pseudopilus; alternatively, the requirement for these proteins to assemble the pseudopilus may not be absolute. The mutants in the *comG* genes seemed to accumulate ComGC in the membrane (Fig. 3*B*, *lower panel*), which is consistent with the failure to translocate ComGC to the cell wall.

Processing of ComGC at the Membrane: Cleavage and Oxidation—ComGC is cleaved by the prepilin peptidase ComC (8), whereas formation of an intramolecular disulfide bond is presumably catalyzed by BdbDC (9). Both BdbD and BdbC are predicted to reside in the membrane: BdbD is similar to *E. coli* DsbA, and BdbC has similarity to DsbB. In the absence of either BdbD or BdbC, the amount of ComGC in the membrane is reduced, presumably due to the instability of the non-oxidized form (9). We have investigated these processing steps of ComGC in more detail.

The transformation deficiency of a mutant lacking BdbC can be partially rescued by treatment with oxidizing agents: the addition of L-cystine to cells increased the transformation frequency of a *bdbC* strain up to 100-fold (Fig. 4*A*). On the other hand, L-cystine had a much less pronounced impact on transformation frequencies of strains lacking BdbD, or both BdbD and BdbC (Fig. 4*A*). Another oxidizing agent, oxidized glutathione, had similar effects (not shown). These results support the role of BdbDC as a thiol-disulfide oxidoreductase pair: BdbD

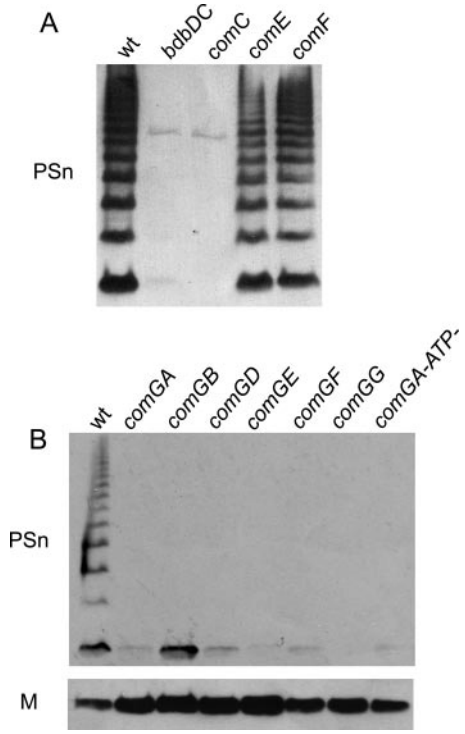

FIGURE 3. **Factors necessary for ComGC complex formation.** Protoplast supernatant (*PSn*) and membrane fractions (*M*) from mutants in different late competence genes were analyzed by immunoblotting. *A*, wild-type strain (BD2528) and mutants in *comG* (BD2780), *bdbDC* (BD3002), *comC* (BD2708), *comE* (BD2718), and *comF* (BD2996). *B*, wild-type strain (BD2528) and mutants in each of the *comG* genes: *comGA* (BD2786), *comGB* (BD2787), *comGD* (BD2789), *comGE* (BD2790), *comGF* (BD2791), and *comGG* (BD2709); mutant in the *comGA* Walker A ATP-binding site (*comGA*-ATP$^-$, BD4140). 1 μg of membrane protein was loaded in the *lower panel* (*M*).

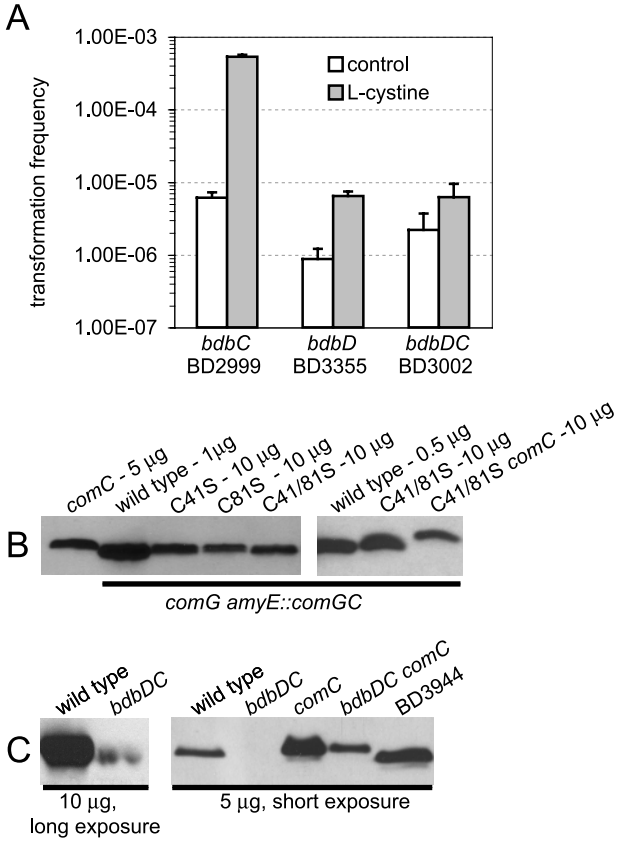

FIGURE 4. **Processing of ComGC in the membrane.** *A*, the transformation deficiency of a *bdbC* mutant can be partially relieved by oxidizing agents. Treatment of cells with 5 mM L-cystine enhanced transformation frequencies of the *bdbC* strain (BD2999); a less pronounced effect was seen with the strain lacking *bdbD* (BD3355) or the double mutant *bdbDC* (BD3002). *B*, the cysteine residues of ComGC are important for protein stability but not for cleavage by ComC. Immunoblot analysis of membrane preparations from strains expressing different forms of ComGC from an ectopic locus; to increase the ComGC signal in these strains, the *comG* operon was inactivated. In the *left panel*, 1 μg of protein from the strain expressing wild-type ComGC (BD3944), and 10 μg of protein from strains expressing ComGC-C41S (BD3942), ComGC-C81S (BD3943), and ComGC-C41/81S (BD3945) were loaded; an uncleaved ComGC sample from a *comC* mutant (BD3845, 5 μg) was loaded as a standard. In the *right panel*, 0.5 μg of protein from the strain expressing wild-type ComGC (BD3944), and 10 μg of protein from strains expressing ComGC-C41/81S (BD3945) and ComGC-C41/81S *comC* (BD3999) were loaded. *C*, oxidation of ComGC does not depend on cleavage. Immunoblot analysis of membrane preparations (10 μg of protein and longer exposure in the *left panel*, 5 μg of protein in the *right panel*) from wild-type strain (BD2528), *bdbDC* (BD3002), *comC* (BD3844), and *bdbDC comC* (BD3845). Strain BD3944 (see *B*) is included as a cleaved ComGC standard in the *last lane* on the *right*.

must be oxidized to catalyze the oxidation of ComGC. As a result of this reaction, BdbD becomes reduced and has to interact with BdbC to be re-oxidized. It appears that, in the absence of BdbC, re-oxidation of BdbD can be artificially achieved by oxidizing agents. Such substances are presumably also able to oxidize ComGC directly, albeit with lower efficiency than BdbD, resulting in the small increase in transformation frequencies seen in the *bdbD* or *bdbDC* mutants.

We substituted the cysteine residues with serine (C41S and/or C81S), generating ComGC variants incapable of forming the internal disulfide bond. The mutated ComGC forms do not support transformation, and no ComGC complex could be detected in the supernatant (not shown). Furthermore, the mutated proteins are much less abundant in membrane fractions than wild-type ComGC (Fig. 4*B*). Note that different amounts of protein were loaded in the gels to permit visualization of the mutant ComGC forms. The result with the C41/81S double mutant rules out the possibility that the presence of an unpaired thiol leads to ComGC instability. The instability of the cysteine mutants is consistent with the notion that oxidation of ComGC is necessary to achieve a stable conformation.

Because the cleavage and oxidation of ComGC both occur in the membrane, we asked whether these processing events occur in some obligatory order, or are independent. To approach this problem, we analyzed membrane fractions for the amount and cleavage states of ComGC in the wild-type and various mutant strains. We first asked whether ComC required oxidized ComGC as a substrate. ComGC C41/81S cannot be oxidized but is nevertheless cleaved by ComC, as demonstrated by the shift in its migration when *comC* is eliminated by mutation (Fig. 4*B*). We conclude that the ComC peptidase can cleave unoxidized ComGC.

We next asked the converse question: can the BdbDC oxidoreductase act on uncleaved ComGC? For this, we compared the strength of ComGC signals in immunoblots of membrane fractions prepared from wild-type, *comC*, *bdbDC*, and double mutant *comC bdbDC* strains (Fig. 4*C*). As noted previously (9), when *bdbDC* was inactivated, the amount of ComGC was drastically reduced (Fig. 4*C*, *left panel*), indicating that unoxidized ComGC is improperly folded and subject to degradation in the membrane. In contrast, inactivation of *comC* alone resulted in

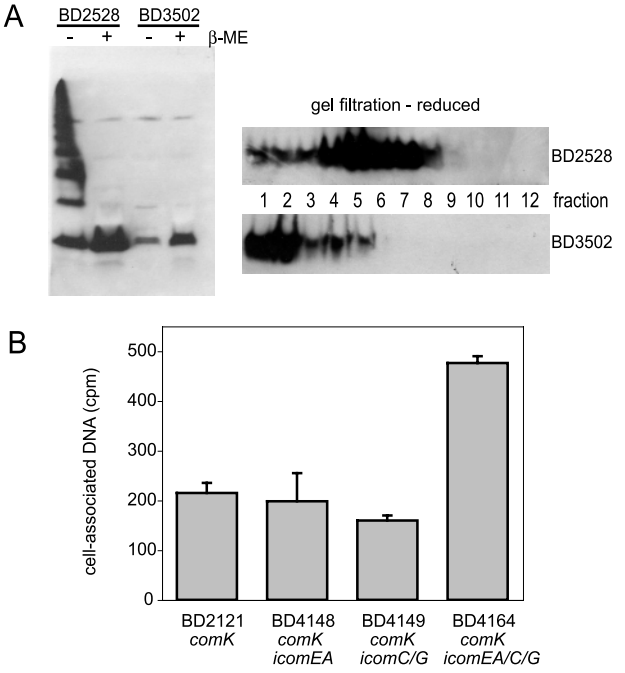

FIGURE 5. **Factors that are sufficient for ComGC complex formation.** Expression of *comC*, *bdbDC*, and the *comG* operon from IPTG-inducible promoters in a *comK* background is sufficient to form ComGC complex. In *A*: *left panel*, immunoblot analysis of protoplast supernatant preparations from wild-type strain BD2528 and IPTG-inducible strain BD3502, with or without β-mercaptoethanol reduction. *Right panel*, immunoblot analysis of fractions from gel-filtration chromatography: protoplast supernatant preparations from wild-type strain (BD2528) and IPTG-inducible strain BD3502 were separated in a Sephacryl S-500HR column. The distortion of the bands in this panel is due to intentional overloading, to enable detection of minor amounts of ComGC in some of the fractions. *B*, expression of *comEA*, *comC*, and the *comG* operon from inducible promoters (*icomEA/C/G*, strain BD4164) is sufficient to promote binding of radiolabeled DNA (compared with BD2121, $p < 0.0001$, Student's *t* test). Expression of *comEA* from a xylose-inducible promoter (*icomEA*, BD4148) or of *comC*/*comG* by themselves (*icomC*/G, BD4149) does not promote DNA binding to the cells. The *graph* presents data from one representative experiment out of three independent experiments, performed in triplicate, with results shown as mean ± S.D. All strains are *comK*.

the accumulation of the pre-ComGC, presumably due to the failure of unprocessed protein to translocate out of the membrane (8). Mutation of *bdbDC* in the *comC* background decreased the amount of pre-ComGC, indicating that oxidation is required for the stability of pre-ComGC. Thus, ComGC oxidation by BdbDC can occur in the absence of cleavage. We conclude from these experiments that BdbDC can oxidize uncleaved ComGC and that ComC can cleave unoxidized ComGC. It therefore appears that these processing events are independent of one another and can occur in either order.

Factors That Are Sufficient to Assemble the ComGC Complex—To determine whether ComC, BdbDC, and the ComG proteins are sufficient to form the ComGC pseudopilus, we constructed a strain expressing these proteins from IPTG-inducible promoters. To ensure that no other late competence gene would be expressed in this strain, the gene encoding the competence transcription factor ComK was knocked out. Upon induction with IPTG, ComGC can indeed be found in the protoplast supernatant, but in amounts lower than in a wild-type strain, expressing all genes from their native promoters (Fig. 5*A*). Reduction with β-mercaptoethanol increased the intensity of the ComGC signal in immunoblots, indicating that ComGC is present as higher molecular weight forms stabilized by disulfide bonds under the denaturing conditions of SDS-PAGE. However, the "ladder" pattern could not be observed in the inducible strain, even when the non-reduced sample was run on a separate gel, to avoid possible diffusion of β-mercaptoethanol between lanes. A possible explanation is that the complex formed by this strain might be larger than the wild type and thus would not enter the polyacrylamide gel or be transferred efficiently to the nitrocellulose membrane. To test this hypothesis, we performed analytical gel-filtration chromatography of the protoplast supernatant and found that ComGC elutes earlier and in a less disperse pattern in the preparation from the inducible strain, compared with the wild type (Fig. 5*A*), resulting in an estimated Stokes radius ranging from 8.3 to 25 nm. This indicates that the ComGC complex in the inducible strain is larger than the wild-type strain, for which the Stokes radius ranges from 5.5 to 11.7 nm. We could not determine the sedimentation coefficient of the induced ComGC complex, because after ultracentrifugation in sucrose gradients, ComGC was found in the bottom of the sucrose column, presumably due to formation of aggregates.

Thus, the complex formed by the inducible strain has physical characteristics different from those of the wild-type complex, which are reflected in its migration pattern in SDS-PAGE and elution profile in gel filtration, as well as its precipitation during ultracentrifugation. Because the putative function of the ComG proteins is to allow DNA to bind its receptor ComEA, we asked whether the inducible ComGC complex was functional, *i.e.* whether it would support DNA binding by ComEA. Strains expressing IPTG-inducible ComC and ComG proteins, together with xylose-inducible ComEA, were constructed in a *comK* background. We did not include IPTG-inducible *bdbDC* in these strains, because BdbD is expressed from its own promoter even in the absence of ComK, although at reduced levels (not shown). All the strains were assayed for their ability to bind DNA (Fig. 5*B*). Expression of ComEA alone or of ComC/ComG by themselves did not promote binding of DNA to cells, but when ComEA was expressed together with ComC/ComG, cells were able to associate with DNA. We conclude that BdbDC, ComC, and ComG proteins are not only necessary but are also sufficient to form a pseudopilus that can facilitate DNA binding to ComEA.

DISCUSSION

Pilin-like proteins and orthologs of pilus biogenesis components are required for DNA uptake in *B. subtilis*, as well as in other competent organisms, both Gram-positive and Gram-negative. In Gram-negative piliated organisms, it has been assumed that pili themselves participate in the process of DNA uptake, whereas a structure formed by the competence pilin-like proteins has been proposed for non-piliated organisms (14). Here we provide the first evidence for such a complex, which we have tentatively named the competence pseudopilus.

The *B. subtilis* competence pseudopilus is a multimeric complex of the pilin-like protein ComGC. Proteins from the type 4 pilin family have a conserved hydrophobic α-helical domain at their N-terminal region, which renders monomeric pilin insoluble (11). This hydrophobic region is thought to be embedded

in the membrane in the monomer (26), or buried inside the core of the pilus fiber, where their interactions hold the polymer together (11, 27, 28). ComGC was identified as a pilin-like protein due to its conserved hydrophobic region, which anchors it in the membrane (19). When expressed in *E. coli*, ComGC also associated with membrane fractions, and it was insoluble in the absence of detergent (not shown). Nevertheless, ComGC complexes can be extracted in a soluble form from the cell wall of competent *B. subtilis*. We do not have evidence that the competence pseudopilus is a filament. However, the properties of ComGC and the overall sequence similarities of the ComG proteins with those from type 4 pili and type 2 secretion systems are strong indications of a common basic structure.

If the competence pseudopilus has the same organization as type 4 pili, it would be anchored in the cytoplasmic membrane. In accordance with this, multimeric ComGC seems to remain associated with protoplasts (Fig. 1). The release of pseudopilus material from protoplasts might occur simply by mechanical shearing of a protruding structure. Modeling studies of the type 4 pilus and the secretion pseudopilus have indicated that each monomer contributes ~1 nm to the length of the fiber (11, 29). Applying this relationship to the estimated number of subunits of the complexed ComGC results in a length of 40–100 nm. This is consistent with a structure spanning the periplasmic space and cell wall of *B. subtilis*, which together are ~55-nm thick (30).

The other competence pseudopilins (ComGD, -GE, and -GG) are required to assemble the pseudopilus, but we do not know whether they are a structural part of the pseudopilus. In type 4 pilus and type 2 secretion systems, a major pilin or pseudopilin subunit is assembled into the pilus structure, but minor pilins are required for its biogenesis or function (12, 31, 32). The minor pseudopilins from *B. subtilis* are also processed by the prepilin peptidase ComC, and ComGD and ComGG can be found in the protoplast supernatant fraction (10).

We have attempted to purify pseudopili from protoplast supernatant, to enable further characterization. These efforts have been hampered by the heterogeneous size of the structures. Moreover, addition of histidine tags at the C terminus of ComGC, even with the inclusion of linker sequences, prevented the assembly of pseudopili and did not allow transformation (not shown), although derivative forms of the PulG pseudopilin from *Klebsiella oxytoca*, containing various tags added to its C-terminal end, could be assembled into pseudopili and functioned in secretion (33). Attempts to visualize structures in crude protoplast supernatants and partially purified samples using electron microscopy with negative staining also failed, presumably because of the low abundance of the pseudopili. This issue will be pursued when other approaches to purification prove successful.

Although no such image is available, we can estimate the degree of asymmetry of the ComGC-containing complex (34). This procedure relies on determination of the ratio S_{max}/S for a given particle, where S_{max} denotes the sedimentation coefficient of a sphere with the mass of the particle. Calculation of this ratio for the ComGC complex with a mass of 1000 kDa and $S = 20.6$, yielded a ratio of 1.8. The work of Schürmann *et al.* (34) suggests that this ratio corresponds to an asymmetric particle, with an axial ratio of ~10. We suggest, therefore, that the ComGC complex is elongated, as would be expected for a pseudopilus. The diameter of the type 4 pilus and secretion pseudopilus is 6–8 nm (11, 29). An axial ratio of 10 would correspond to a length of 60–80 nm, in agreement with the proposed length of 40–100 nm, given above for the competence pseudopilus.

The late competence genes *comC*, *bdbDC*, and the whole *comG* operon are both necessary and sufficient to form a ComGC complex. The amount of ComGC in the inducible *comC/comG/bdbDC* strain was reduced compared with the wild-type strain, but the size of the complexed ComGC was apparently larger (Fig. 5*A*). One possible explanation is the lack of a factor controlling the size of the pseudopilus in *comK* cells; alternatively, there might be an imbalance in the amounts of the induced proteins. Another potential issue is the mislocalization of the induced complex: ComGA localizes preferentially at the pole of competent cells, where DNA binding and uptake actually take place, and polar localization of ComGA appears to be reduced in the absence of ComK (35). Nevertheless, even with physical characteristics that differed from those of the wild type, the induced ComGC complex allowed exogenous DNA to access its membrane bound receptor, ComEA (Fig. 5*B*).

ComGC contains an intramolecular disulfide bond, which seems necessary for its stability in the membrane. In contrast, in the competence pseudopilus, ComGC subunits are covalently joined by intermolecular disulfide bonds. Thus, intramolecular disulfide bonds from ComGC in the membrane must be reduced to allow formation of the intermolecular ones in the pseudopilus, but no disulfide bond isomerase could be identified in the *B. subtilis* genome sequence. ComGG has one cysteine residue and was shown to form a dimer sensitive to reducing agents (10), suggesting it might participate in the bond exchange reaction. However, substitution of the cysteine residue in ComGG by serine did not have any measurable impact on transformation frequency (not shown).

The disulfide bond exchange between ComGC monomers might occur in the assembled pseudopilus. In gel-filtration chromatography, the elution pattern of ComGC from the protoplast supernatant did not change under reducing conditions, indicating that interactions other than the intermolecular disulfide bonds stabilize the structure of the pseudopilus. In type 4 pili, the hydrophobic interactions between the N termini in the core of the fiber are thought to hold the polymer together (27, 28); the same kind of interactions might occur in the competence pseudopilus. Thus, if the cysteine residues from one ComGC subunit are held near those from neighboring subunits by these non-covalent interactions, the disulfide bond exchange could occur spontaneously. In fact, the bond exchange reactions may not be complete, and some ComGC subunits might retain their intramolecular disulfide bond or form only one intermolecular bond, which could explain the monomeric and other smaller forms seen in SDS-PAGE/immunoblots.

The ComGC complex released in the protoplast supernatant has a wide range of size distribution, from 40 to 100 nm. This could be due to shearing, yielding breaks at random points of a rigid structure, although one might expect most breaks to occur

near the base. The broad size distribution is suggestive of another feature of the type 4 pilus: its ability to retract. Type 4 pili mediate so-called twitching motility, powered by the retraction (disassembly) and extension (assembly) of the pilus fibers (36). The secretion pseudopilus is thought to function as a piston, pushing substrates from the periplasm through the secretin channel in the outer membrane of Gram-negative bacteria (7, 14); thus, it might also extend and retract. Similarly, the competence pseudopilus might also have a dynamic nature, capable of extending and retracting by assembly/disassembly events, resulting in the observed heterogeneous size distribution of the pseudopilus population.

Retraction of the pseudopilus may bring DNA from the extracellular environment to its membrane-bound receptor ComEA. In addition, retraction of the pseudopilus might be required for ComEA to carry out its role in uptake of transforming DNA. ComEA has a potentially flexible region located next to the DNA-binding domain, proposed to function as a hinge that allows ComEA to bend and deliver bound DNA to the translocation channel in the cytoplasmic membrane formed by ComEC (37). Pseudopilus retraction may be required to bend ComEA, driving DNA transport. Whatever the molecular details, a model in which pseudopilus retraction plays a role in DNA uptake by *B. subtilis* is supported by results obtained from single-cell experiments, where the unique force characteristics of DNA uptake were found to be similar to those from type 4 pilus retraction (38).

Acknowledgments—We thank all the members of our laboratory for helpful discussion and J. Dubnau for reviewing the manuscript.

REFERENCES

1. van Sinderen, D., Luttinger, A., Kong, L., Dubnau, D., Venema, G., and Hamoen, L. (1995) *Mol. Microbiol.* **15,** 455–462
2. Hamoen, L. W., Venema, G., and Kuipers, O. P. (2003) *Microbiology* **149,** 9–17
3. Chen, I., and Dubnau, D. (2004) *Nat. Rev. Microbiol.* **2,** 241–249
4. Provvedi, R., and Dubnau, D. (1999) *Mol. Microbiol.* **31,** 271–280
5. Inamine, G. S., and Dubnau, D. (1995) *J. Bacteriol.* **177,** 3045–3051
6. Chung, Y. S., and Dubnau, D. (1998) *J. Bacteriol.* **180,** 41–45
7. Pugsley, A. P. (1993) *Microbiol. Rev.* **57,** 50–108
8. Chung, Y. S., and Dubnau, D. (1995) *Mol. Microbiol.* **15,** 543–551
9. Meima, R., Eschevins, C., Fillinger, S., Bolhuis, A., Hamoen, L. W., Dorenbos, R., Quax, W. J., van Dijl, J. M., Provvedi, R., Chen, I., Dubnau, D., and Bron, S. (2002) *J. Biol. Chem.* **277,** 6994–7001
10. Chung, Y. S., Breidt, F., and Dubnau, D. (1998) *Mol. Microbiol.* **29,** 905–913
11. Craig, L., Pique, M. E., and Tainer, J. A. (2004) *Nat. Rev. Microbiol.* **2,** 363–378
12. Sauvonnet, N., Vignon, G., Pugsley, A. P., and Gounon, P. (2000) *EMBO J.* **19,** 2221–2228
13. Durand, E., Bernadac, A., Ball, G., Lazdunski, A., Sturgis, J. N., and Filloux, A. (2003) *J. Bacteriol.* **185,** 2749–2758
14. Hobbs, M., and Mattick, J. S. (1993) *Mol. Microbiol.* **10,** 233–243
15. Hahn, J., Luttinger, A., and Dubnau, D. (1996) *Mol. Microbiol.* **21,** 763–775
16. Albano, M., Hahn, J., and Dubnau, D. (1987) *J. Bacteriol.* **169,** 3110–3117
17. Hahn, J., Roggiani, M., and Dubnau, D. (1995) *J. Bacteriol.* **177,** 3601–3605
18. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* (*Amst.*) **77,** 51–59
19. Breitling, R., and Dubnau, D. (1990) *J. Bacteriol.* **172,** 1499–1508
20. Schagger, H., and von Jagow, G. (1987) *Anal. Biochem.* **166,** 368–379
21. Siegel, L. M., and Monty, K. J. (1966) *Biochim. Biophys. Acta* **112,** 346–362
22. Anagnostopoulos, C., and Spizizen, J. (1961) *J. Bacteriol.* **81,** 741–746
23. Hahn, J., Inamine, G., Kozlov, Y., and Dubnau, D. (1993) *Mol. Microbiol.* **10,** 99–111
24. Londoño-Vallejo, J. A., and Dubnau, D. (1993) *Mol. Microbiol.* **9,** 119–131
25. Haijema, B. J., Hahn, J., Haynes, J., and Dubnau, D. (2001) *Mol. Microbiol.* **40,** 52–64
26. Nunn, D. (1999) *Trends Cell Biol.* **9,** 402–408
27. Forest, K. T., and Tainer, J. A. (1997) *Gene* (*Amst.*) **192,** 165–169
28. Keizer, D. W., Slupsky, C. M., Kalisiak, M., Campbell, A. P., Crump, M. P., Sastry, P. A., Hazes, B., Irvin, R. T., and Sykes, B. D. (2001) *J. Biol. Chem.* **276,** 24186–24193
29. Kohler, R., Schafer, K., Muller, S., Vignon, G., Diederichs, K., Philippsen, A., Ringler, P., Pugsley, A. P., Engel, A., and Welte, W. (2004) *Mol. Microbiol.* **54,** 647–664
30. Matias, V. R., and Beveridge, T. J. (2005) *Mol. Microbiol.* **56,** 240–251
31. Alm, R. A., and Mattick, J. S. (1997) *Gene* (*Amst.*) **192,** 89–98
32. Filloux, A. (2004) *Biochim. Biophys. Acta* **1694,** 163–179
33. Vignon, G., Kohler, R., Larquet, E., Giroux, S., Prevost, M. C., Roux, P., and Pugsley, A. P. (2003) *J. Bacteriol.* **185,** 3416–3428
34. Schurmann, G., Haspel, J., Grumet, M., and Erickson, H. P. (2001) *Mol. Biol. Cell* **12,** 1765–1773
35. Hahn, J., Maier, B., Haijema, B. J., Sheetz, M., and Dubnau, D. (2005) *Cell* **122,** 59–71
36. Merz, A. J., So, M., and Sheetz, M. P. (2000) *Nature* **407,** 98–102
37. Dubnau, D. (1999) *Annu. Rev. Microbiol.* **53,** 217–244
38. Maier, B., Chen, I., Dubnau, D., and Sheetz, M. P. (2004) *Nat. Struct. Mol. Biol.* **11,** 643–649
39. Hahn, J., Albano, M., and Dubnau, D. (1987) *J. Bacteriol.* **169,** 3104–3109
40. Draskovic, I., and Dubnau, D. (2005) *Mol. Microbiol.* **55,** 881–896
41. Albano, M., Breitling, R., and Dubnau, D. A. (1989) *J. Bacteriol.* **171,** 5386–5404
42. Hahn, J., Bylund, J., Haines, M., Higgins, M., and Dubnau, D. (1995) *Mol. Microbiol.* **18,** 755–767