Molecular Microbiology (1998) **29**(3), 905–913

Cell surface localization and processing of the ComG proteins, required for DNA binding during transformation of *Bacillus subtilis*

Y. S. Chung, F. Breidt† and D. Dubnau*
Public Health Research Institute, 455 First Avenue, New York, NY 10016, USA.

Summary

The *comG* operon of *Bacillus subtilis* encodes seven proteins essential for the binding of transforming DNA to the competent cell surface. We have explored the processing of the ComG proteins and the cellular localization of six of them. All of the proteins were found to be membrane associated. The four proteins with N-terminal sequence motifs typical of type 4 prepilins (ComGC, GD, GE and GG) are processed by a pathway that requires the product of *comC*, also an essential competence gene. The unprocessed forms of ComGC and GD behave like integral membrane proteins. Pre-ComGG differs from pre-ComGC and pre-ComGD, in that it is accessible to proteolysis only from the cytoplasmic face of the membrane and at least a portion of it behaves like a peripheral membrane protein. The mature forms of these proteins are translocated to the outer face of the membrane and are liberated when peptidoglycan is hydrolysed by lysozyme or mutanolysin. ComGG exists in part as a disulphide-cross-linked homodimer *in vivo*. ComGC was found to possess an intramolecular disulphide bond. The previously identified homodimer form of this protein is not stabilized by disulphide bond formation. ComGF behaves as an integral membrane protein, while ComGA, a putative ATPase, is located on the inner face of the membrane as a peripheral membrane protein. Possible roles of the ComG proteins in DNA binding to the competent cell surface are discussed in the light of these and other results.

Introduction

Transformation of naturally competent bacteria proceeds via the binding, processing and internalization of exogenous DNA. To accomplish this, a set of proteins are required that were first identified in *Bacillus subtilis* (reviewed in Dubnau, 1993; 1997). Similar proteins are essential for transformation in a variety of bacterial species, both Gram positive and Gram negative (Facius and Meyer, 1993; Clifton *et al.*, 1994; Fleischmann *et al.*, 1995; Freitag *et al.*, 1995; Tønjum *et al.*, 1995; Fussenegger *et al.*, 1997; Lunsford and Roble, 1997; Pestova and Morrison, 1997; Campbell *et al.*, 1998).

Among the competence proteins with orthologues in the other competence systems are those encoded by the *B. subtilis comG* operon (Albano *et al.*, 1989). This operon specifies seven gene products that are individually required for the binding of DNA to the competent cell (Hahn *et al.*, 1987; Breitling and Dubnau, 1990; Chung and Dubnau, 1998). The ComG proteins belong to several families, each with members required for the assembly of pili and for the secretion of proteins across the outer membranes of certain Gram-negative bacteria, in addition to those needed for competence.

The first of these families is exemplified by ComGA, PilB and PulE, required, respectively, for transformation in *B. subtilis*, pilus assembly in *Pseudomonas aeruginosa* (Nunn *et al.*, 1990) and the secretion of pullulanase in *Klebsiella oxytoca* (Possot *et al.*, 1992). The members of this family possess consensus nucleotide binding motifs and have been postulated to be energy-transducing proteins, coupling the hydrolysis of ATP to the movement of macromolecules across the cell membrane. The second family contains ComGB, PilC (Nunn *et al.*, 1990) and PulF (Possot *et al.*, 1992), the loss of which cause deficiencies in transformation, pilus assembly and pullulanase secretion respectively. ComGB possesses three predicted membrane-spanning segments. The third family consists of a set of small proteins, each absolutely required in its respective system and each with a single predicted membrane-spanning segment near its N-terminus [ComGC, GD, GE, GG for competence, PilA, E, V, W, X and FimU for pilus assembly (Russell and Darzins, 1994; Alm and Mattick, 1995; 1996; Alm *et al.*, 1996) and PulG, H, I, J for pullulanase secretion (Reyss and Pugsley, 1990)]. This protein family includes the major structural subunit of the type 4 pilus (pilin, encoded by *pilA* in *P. aeruginosa*) and, in several cases, its members have been shown to be synthesized in precursor form with a consensus cleavage site for processing by a dedicated peptidase/transmethylase.

Received 7 April, 1998; revised 28 May, 1998; accepted 1 June, 1998. †Present address: 322 Schaub Hall, Box 7624, NC State University Rahleigh, Rahleigh, NC 27695-7624, USA. *For correspondence. E-mail dubnau@phri.nyu.edu; Tel. (212) 578 0842; Fax (212) 578 0804.

The similarities among the members of this third family are essentially confined to a group of hydrophobic N-terminal amino acid residues, which include the cleavage site. Other than those members of this family that encode the structural components of the pilus, the precise roles of the other pilin-like proteins are unknown. The remaining *comG* product, ComGF, is a small protein with a single predicted membrane-spanning segment near its N-terminus. It does not appear to possess a processing site, has no known orthologue and its role is unknown, although it is required for DNA binding to the competent cell surface (Chung and Dubnau, 1998).

In addition to the *comG* proteins, a fourth gene family, exemplified by *comC* (Mohan *et al.*, 1989), *pilD* (Strom *et al.*, 1993) and *pulO* (Dupuy *et al.*, 1992), encodes the peptidases required for processing of pre-pilin-like proteins. Indeed, we have shown that ComGC is processed by a pathway that requires the product of *comC* (Chung and Dubnau, 1994). As loss-of-function mutations in *comC* eliminate DNA binding, it appears that processing of at least one of the pilin-like proteins is required for transformation and, specifically, for the first step in this process.

It has been proposed that the pilin-like proteins are required for the assembly of a structure embedded in the cell surface. It is postulated that this structure participates in the transport of DNA, in the case of transformation, or of protein, in the case of the secretion systems (Albano *et al.*, 1989; Breitling and Dubnau, 1990; Hobbs and Mattick, 1993; Pugsley, 1993; Mattick and Alm, 1995). Arguing from the example of pilus biogenesis, one or more of the pilin-like proteins in each system might act as a component of the proposed structure, while the other proteins might be required for assembly but would not themselves be present in the structure. Although these ideas have been discussed widely, little direct evidence for or against them has been presented, and it is clearly important to explore the subcellular localization of these proteins and their interactions. It is likely that information concerning the role of a given protein in any one system will provide insights helpful in understanding the related systems.

Unprocessed ComGC (in a *comC* background) is localized as an integral membrane protein, with its C-terminus outside the membrane (Breitling and Dubnau, 1990; Chung and Dubnau, 1994). Upon cleavage, a portion of the total ComGC pool is translocated to the outside of the membrane (Chung and Dubnau, 1994). Cross-linking experiments have shown that ComGC, in both the processed and the unprocessed state, is present as a dimer (probably a homodimer), while no evidence for the association of ComGC with any other proteins has been obtained (Chung and Dubnau, 1994).

In the present study, we have investigated the localization of ComGA, GC, GD, GE, GF and GG, as well as the processing of the pilin-like proteins.

Results

Localization of ComGA

Immunoblots of fractionated competent cell extracts were used to determine the subcellular localization of ComGA. Protoplasts were collected by centrifugation, osmotically lysed and the membrane and cytoplasmic fractions were recovered after further centrifugation. Figure 1A shows that ComGA co-sedimented with the membranes, with only a trace amount visible in the cytoplasmic fraction. In other fractionation experiments, the protoplast supernatant fraction was also tested, and no ComGA signal was detected (not shown). Treatment of the membranes with NaOH solubilized all of the detectable ComGA (Fig. 1A), demonstrating that this protein behaves as a peripheral membrane protein (Russel and Model, 1982). Indeed, hydropathy analysis of ComGA had not predicted the presence of a membrane-spanning segment (Albano *et al.*, 1989), and at least one other member of the ComGA family has also been shown to be a peripheral membrane protein (Sandkvist *et al.*, 1995).

In order to determine whether ComGA is exposed on the inner or outer surface of the membrane, protoplasts were incubated in the presence of proteinase K, with and without prior treatment with 1% Triton X-100 (Fig. 1B). Treatment of intact protoplasts with proteinase K had no effect on the strength of the immunoblot signal, whereas prior disruption of the protoplasts with detergent rendered the ComGA accessible to proteolysis. These results demonstrated that the protein is located on the inner face of the membrane, consistent with the presence of a potential nucleotide binding site on ComGA. Further support for this conclusion was derived from the following cross-linking experiment. Treatment of protoplasts with disuccinimidyl tartarate resulted in the cross-linking of ComGA to yield higher molecular weight forms, whereas use of the

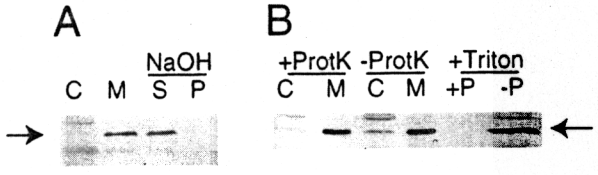

Fig. 1. Localization of ComGA.
A. An immunoblot using anti-ComGA antiserum. Subcellular fractions were obtained from competent BD630. M, total membrane fraction; C, cytoplasmic fraction. Membranes were treated with NaOH where indicated and centrifuged. The ComGA signal present in the resulting supernatant (S) and pellet (P) is shown.
B. Immunoblot using anti-ComGA antiserum with fractions derived from protoplasts of BD630, previously incubated in the presence and absence of proteinase K (2.5 mg ml^{-1}). M, total membrane fraction; C, cytoplasmic fraction. The two lanes to the extreme right show the effect of incubation with (+P) and without (−P) proteinase K treatment when the protoplasts were previously lysed with Triton X-100.

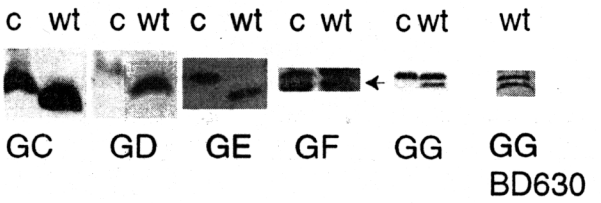

Fig. 2. Processing of ComG proteins requires ComC. Immunoblots using antisera for the indicated proteins were run using membrane preparations from wild-type (wt) and *comC* (c) mutant strains. The backgrounds of the strains (except for the one used in the panel marked BD630) were BD2528 for ComGC, ComGF and ComGG, and BD2519 for ComGD and ComGE (see Table 1). The arrow indicates the ComGF signal, and the upper band corresponds to a cross-reacting protein.

impermeable hydrophilic compound disulphodisuccinimidyl tartarate did not (not shown). These forms are large and polydisperse. The accessibility of ComGA only to the hydrophobic cross-linker is consistent with the proposed location of this protein.

ComGC, GD, GE and GG, but not ComGF, are processed by a ComC-dependent pathway

These five small proteins each possess a single predicted N-terminal membrane-spanning segment, but only ComGC, GD, GE and GG exhibit potential cleavage sites and similarity to the larger family of pre-pilin-like proteins (Albano *et al.*, 1989). We have shown previously that ComGC is processed by a pathway that is dependent on ComC (Chung and Dubnau, 1994). This observation has now been extended to ComGD, GE and GG (Fig. 2). As expected, no evidence for ComGF processing was obtained; the upper band shown for ComGF in Fig. 2 is a cross-reacting protein. In the case of ComGG, the processing was often incomplete; in many experiments, a variable fraction of the total ComGG signal in extracts of wild-type bacteria migrated at the same position in SDS–PAGE as the signal obtained from *comC* extracts. Although most of the results with ComGC, GD, GE, GF and GG shown in Fig. 2 were obtained with overproducing strains, it can be seen that the partial processing of ComGG was also detected in a wild-type background (BD630) and was therefore not caused by saturation of the processing machinery. ComGG lacks a nearly invariant glutamyl residue, present in the type 4 pilin-like proteins, in the fifth position downstream from the cleavage site. This departure from the consensus may be responsible for the partial processing observed. However, when the glutamyl residue in PilA of *Pseudomonas aeruginosa* was altered by mutagenesis, processing was unaffected, although the N-terminal methylation usually observed in mature type 4 pilin protein was absent (Pasloske and Paranchych, 1988; Strom and Lory, 1991). When the conserved glutamyl residue of PulG was changed to a valine or alanine residue, processing and methylation were unaffected (T. Pugsley, personal communication; Pugsley, 1996). We do not know whether ComGG or the other pilin-like competence proteins are N-methylated.

Localization of ComGC, GD, GE, GF and GG

As all the data presented in Fig. 2 were obtained with fractionated membranes, it is apparent that ComGD, GE, GF and GG, like ComGC, are at least partially membrane localized. ComGC has been shown to be localized in the cytoplasmic membrane when unprocessed, while a portion of the total ComGC pool is translocated to the outer face of the membrane after processing to a form that is NaOH soluble and therefore no longer present as an integral membrane protein (Chung and Dubnau, 1994). The localization of the other pilin-like proteins and of ComGF was examined by immunoblotting of fractionated competent cell extracts, and that of ComGC was re-examined.

Figure 3 shows that ComGC, in addition to its previously observed presence in association with the membrane, was also detected in the protoplast supernatant. In repeated experiments, a variable fraction of the ComGC, averaging roughly 25% of the total, was detected in the protoplast supernatant. In a *comC* mutant background, the signal detected in the protoplast supernatant fraction was greatly diminished. Together with our previous results, these observations reveal that the ComC-dependent translocation of ComGC leads to localization of the processed protein on the outer face of the membrane and as a form that may be associated with cell wall material and is solubilized upon hydrolysis of peptidoglycan. Roughly half of the processed ComGC molecules remain localized as integral membrane proteins with their C-termini outside the membrane (Breitling and Dubnau, 1990; Chung and Dubnau, 1994).

Figure 4 reveals equivalent results for ComGD. About 25% of the total signal was detected in the protoplast supernatant fraction and the remainder in association with the membrane (Fig. 4A). Again, the fraction detected in the protoplast supernatant in repeated experiments was

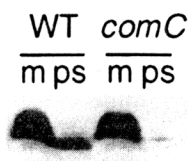

Fig. 3. Localization of ComGC. Immunoblot of ComGC in membrane (m) and protoplast supernatant (ps) fractions from wild-type and *comC* strains in the BD2528, multicopy *comS* background. For this experiment, a short SDS–polyacrylamide gel was used, which poorly resolved pre-ComGC from its processed form.

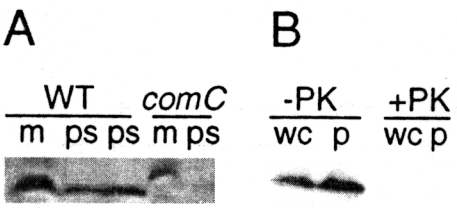

Fig. 4. Localization of ComGD.
A. Immunoblot of ComGD in membrane (m) and protoplast supernatant (ps) fractions from wild-type and *comC* strains in the BD2528 background.
B. Survival of ComGD in total extracts when whole cells (wc) or protoplasts (p) of strain BD2528 were treated with proteinase K (PK).

variable. No ComGD was detected in the cytoplasmic fraction (not shown). In a *comC* mutant background, pre-ComGD was detected in association with membrane, and no signal was detected in the protoplast supernatant. Treatment of whole cells or of protoplasts with proteinase K demonstrated that, as in the case of ComGC (Breitling and Dubnau, 1990), all of the ComGD signal was susceptible to proteolysis (Fig. 4B). We conclude that both the moiety associated with the membrane and that present in the protoplast supernatant were accessible to exogenous protease. ComGE was also shown to be associated with the membrane (see Fig. 2) but was not studied further.

The properties of ComGF were distinct from those of ComGC and ComGD. Figure 5A shows that ComGF was associated with the membrane and that no detectable signal was present in the protoplast supernatant fraction. In other experiments, no signal was observed in the cytoplasmic fraction (not shown). The ComGF signal was insoluble in NaOH (Fig. 5B), indicating that it was present as an integral membrane protein. We conclude that ComGF, which is not processed, is an integral membrane protein. Hydropathy analysis of ComGF predicted the presence of a single N-terminal membrane-spanning segment (Albano *et al.*, 1989). The orientation of ComGF in the membrane was not determined.

We next addressed the location of ComGG. Figure 6A illustrates the positions in a gel of the precursor and mature

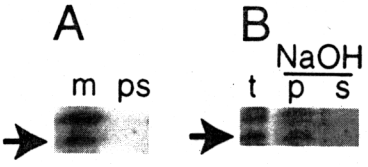

Fig. 5. Localization of ComGF.
A. Immunoblot of ComGF in membrane (m) and protoplast supernatant (ps) fractions from BD2528.
B. Immunoblot of ComGF in the total membrane fraction (t) and in the NaOH-soluble (s) and -insoluble (p) fractions derived from BD2528 membranes. The arrows mark the ComGF signal. The band above that of ComGF represents a cross-reacting protein.

forms of this protein (compare the wild-type sample in lane 7 with the *comC* sample in lane 8). Figure 6A also shows that the membrane fraction contained both processed and unprocessed molecules (Fig. 6A, lanes 1 and 7), whether the membranes were isolated with or without prior protoplast formation. A portion of the total ComGG signal (Fig. 6A, lane 2) was located in the protoplast supernatant fraction, and this consisted solely of processed protein. The percentage of total mature ComGG protein released into the supernatant fraction varied somewhat in different experiments, as did the proportion of total mature and precursor ComGG. In fact, in some experiments, all of the mature ComGG was found in the protoplast supernatant fraction (not shown). These results demonstrate that mature ComGG can be released upon protoplasting, as noted above for ComGC and ComGD (Figs 3 and 5). Proteinase K treatment of protoplasts (Fig. 6A, lane 3) or of intact cells (Fig. 6A, lane 4) destroyed the processed form and did not affect the pre-ComGG band. Triton X-100 treatment of membranes isolated from protoplasts (Fig. 6A, lanes 5 and 6) rendered all of the ComGG proteinase K sensitive, suggesting that pre-ComGG may not be intrinsically proteinase K resistant. In this respect, pre-ComGG behaves differently from pre-ComGC (Breitling and Dubnau, 1990), which is proteinase K accessible in intact cells and

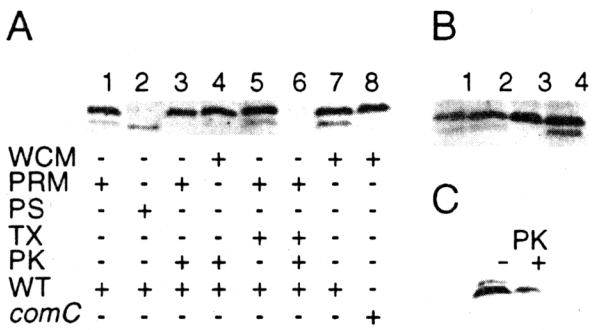

Fig. 6. Localization and properties of ComGG.
A. Immunoblot of ComGG from wild-type (WT) and *comC* strains. WCM, membranes isolated from whole cells after treatment with mutanolysin. PRM, membranes isolated from protoplasts. PK indicates that the whole cells or protoplasts were treated with proteinase K before the lysis step. PS, protoplast supernatant fraction; TX, Triton X-100 treatment of membranes isolated from protoplasts. One half of the Triton X-100 treated sample was incubated with proteinase K (lane 6).
B. NaOH solubility of ComGG. Lanes 1, 2 and 4 show the ComGG signal from the wild-type strain and lane 3 from *comC* membranes. Lanes 3 and 4 contain total membranes and are included to establish the positions in the gel of the processed and unprocessed forms of ComGG and to show the extent of processing in the wild-type preparation. Lanes 1 and 2 display the NaOH-insoluble and NaOH-soluble ComGG signals from the membrane preparation shown in lane 4.
C. Immunoblot of ComGG from everted membrane vesicles incubated with and without proteinase K.
The strain background used for the experiments shown in A and B was that of BD2528.

in protoplasts. Pre-ComGG differs in another respect from the other pre-ComG proteins, in that about half of it is NaOH soluble (Fig. 6B). The ComGG data presented so far can be accommodated by the following model. We propose that about half of the pre-ComGG is arranged as a peripheral membrane protein, and about half is present as integral membrane protein. All of the pre-ComGG is inaccessible to proteinase K in intact cells and must, therefore, be exposed on the cytosolic face of the membrane. Upon processing, ComGG is translocated to a position outside the membrane where it is proteinase K susceptible. Some of this material can be liberated by treatment with lysozyme or mutanolysin and some remains NaOH insoluble. All of the processed ComGG is proteinase K accessible and therefore exposed outside the membrane. To test this model further, membrane vesicles from competent cells were prepared by French pressure cell lysis and treated with proteinase K (Fig. 6C). The French pressure cell yields vesicles at least some of which are oriented inside out (Kaback, 1971; Kerppola *et al.*, 1991). After French pressure cell lysis, both the processed and unprocessed forms of ComGG were susceptible to degradation. As the pre-ComGG protein was now proteinase K accessible, this result was in accord with our model. In the sample used for Fig. 6C, a larger proportion of the total ComGG had been processed than in the experiment shown in Fig. 6A. It is possible that processing continues after cell lysis.

ComGG dimers are stabilized by disulphide bond formation

As ComGC, GE, GF and GG contain two, two, three and one cysteine residues respectively, it was of interest to determine whether disulphide bond formation was involved in the dimerization of each of these proteins. Based on cross-linking experiments, we have reported previously that ComGC forms homodimers *in vivo* (Chung and Dubnau, 1994). Competent cells were treated with *N*-ethylmaleimide (NEM) before lysis in order to block the *in vitro* formation of disulphide bonds. Membranes were prepared in the continued presence of NEM and incubated in sample buffer with and without β-mercaptoethanol (β-ME). Figure 7A shows the result obtained in immunoblots using anti-ComGC antiserum. A dimer band was detected only in the absence of both NEM and β-ME. This suggests *in vitro* intermolecular disulphide bond formation. It also suggests that the ComGC dimer, previously detected by the treatment of intact cells with cross-linking reagents, was not stabilized *in vivo* by disulphide bond formation. It is interesting that the migration of the monomer form of ComGC is slightly retarded in the absence of β-ME, irrespective of the treatment with NEM. This suggests the presence of an intramolecular disulphide bond, formed *in vivo*, which alters the migration of ComGC during electrophoresis. ComGF antiserum used with the same samples revealed no evidence of the formation of a disulphide-mediated dimer (not shown).

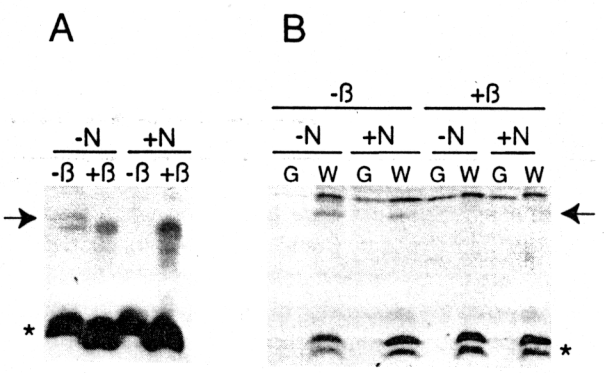

Fig. 7. Dimerization of ComGC (A) and ComGG (B). Intact cells were incubated with and without 10 mM *N*-ethyl maleimide (N) and washed. Membrane fractions were then prepared and subjected to electrophoresis in the presence and absence of 1.44 M β-mercaptoethanol (β).
A. Strain BD2528.
B. Results are shown for BD2528, the wild-type strain (W) and for a strain carrying a Tn*917* insertion in *comGG* in the BD2528 background (BD2709).
The arrows and the asterisks indicate the positions of dimer and monomer forms respectively. The band immediately below the dimer form in (A) represents a cross-reacting protein. In (B), the asterisk is placed between the processed and unprocessed ComGG bands.

A different result was obtained with ComGG (Fig. 7B). In the absence of β-ME, an apparent dimer form was detected in both the NEM-treated and untreated samples. The apparent dimer signal was derived from ComGG, as both it and the monomer form were absent in a *comG210* mutant strain, which carries a Tn*917* transposon insertion in *comGG* (Albano *et al.*, 1989).

We conclude from these studies that the single cysteine residue in ComGG participates in the *in vivo* formation of an intermolecular disulphide bond, although only a minor fraction of the total ComGG monomer is involved. We also conclude that the previously reported ComGC dimer, detectable when cross-linking reagents were used, is not stabilized by a disulphide bond and that there is no evidence for disulphide bond formation in ComGF, either *in vitro* or *in vivo*. Finally, it appears that an intramolecular disulphide bond may form *in vivo* in the case of ComGC, which contains a pair of cysteines.

Discussion

We have established that ComGA, GC, GD, GE, GF and GG are membrane associated. The hydrophobicity of ComGB suggests that this protein is also located in the membrane (Albano *et al.*, 1989). Membrane localization is consistent with the roles of the ComG proteins in the

binding of DNA to competent cells. We have also shown that the four ComG proteins with N-terminal sequences resembling the processing sites of type 4 pre-pilins are cleaved by a mechanism that requires ComC. As the latter is similar in amino acid sequence to the protease/transmethylase that processes the *P. aeruginosa* pre-pilin (Strom *et al.*, 1993), it is likely that the cleavage reactions are directly catalysed by ComC. As loss-of-function *comC* mutants are transformation deficient and fail to bind DNA, it is likely that the processing of at least one and probably all of these essential competence proteins is necessary for DNA binding.

Processing of pre-ComGC is followed by the translocation of the mature protein from its position as an integral membrane protein to a new location outside the cell membrane, although roughly half of the processed molecules remain as NaOH-insoluble membrane-spanning proteins (Chung and Dubnau, 1994). In the present study, we have observed that a previously undetected fraction of the mature protein is released when the cell wall is hydrolysed by treatment with lysozyme or mutanolysin in an osmoprotective medium. This is also true for ComGD and ComGG, and processing also appears to be essential for the translocation of these proteins. This behaviour of the mature forms of ComGC, ComGD and ComGG suggests that they may be associated with cell wall material, loosely attached to the outer surface of the membrane or present in the space between the cell wall and the membrane. NaOH-solubility experiments demonstrated that some of the mature ComGC and ComGG remains as proteinase K-accessible, integral membrane protein molecules. This was not tested for ComGD and ComGE.

Although mature ComGG is localized similarly to ComGC and ComGD, it appears to follow a markedly different processing pathway. Pre-ComGC is detected as a bitopic integral membrane protein with its N-terminus in and C-terminus out (Breitling and Dubnau, 1990). In contrast, pre-ComGG is found on the cytosolic aspect of the membrane, to which it is peripherally attached. In each case, a fraction of the mature form is present as bitopic integral membrane protein. However, the apparent disparities in the processing pathways may actually reflect kinetic differences. Both precursor forms may associate peripherally with the membrane at first, with this being a long-lived intermediate in the case of pre-ComGG and short lived in the case of pre-ComGC. The hydrophobic N-terminal sequences of both precursor proteins may then be buried in the membrane as a loop to yield an NaOH-insoluble form inaccessible to proteinase K added from the outside. This form would also be long lived in the case of pre-ComGG and undetectable in the case of pre-ComGC. Cleavage by ComC could then occur, as the cleavage site of the membrane-embedded loop form would be exposed on the cytoplasmic face of the membrane, where it is accessible to the active site of ComC. Cleavage would be followed closely by translocation of the C-terminal domains, while a fraction of the mature molecules of both proteins is released from the membrane. This formal scheme accounts for our observations and is consistent with the persistence of pre-ComGG (Fig. 2), as the cleavage of this protein is postulated to occur slowly.

To understand the possible functions of the competence proteins, it is instructive to compare protein secretion mediated by the Pul system with transformation in *B. subtilis*. In the *Klebsiella oxytoca* system, transport of pullulanase across the inner membrane is accomplished by the Sec proteins (Pugsley, 1993). For transformation, the proteins ComEC, ComFA and ComEA, which are required for the transport of DNA across the cell membrane, can therefore be regarded as the functional equivalents of the Sec machinery. These competence proteins do not resemble the Sec proteins in amino acid sequence, other than in the nucleotide binding sequence motifs present in ComFA and SecA, both of which are localized primarily on the cytoplasmic face of the inner membrane. The Pul system includes PulD, which has been hypothesized to form an outer membrane export channel (Russel, 1994; Linderoth *et al.*, 1997; Pugsley *et al.*, 1997; Russel *et al.*, 1997), and it is reasonable that a search of the recently completed *B. subtilis* genome sequence (Kunst *et al.*, 1997) reveals no apparent PulD orthologue (not shown). *B. subtilis* is a Gram-positive organism, which lacks an outer membrane. If the remaining Pul products are needed for events occurring in the periplasm during pullulanase secretion, what are the equivalent steps mediated by their orthologues, the ComG proteins, as Gram-positive bacteria have no outer membrane and do not possess a periplasm in the strict sense? The events certain to occur outside the cell membrane during transformation include DNA binding, transport through the cell wall and cleavage of the transforming DNA to yield internal termini for uptake. In addition, the degradation of one of the incoming strands (Dubnau, 1993) may occur outside the membrane. The ComG proteins may be required to carry out any or all of these postulated roles, in addition to their documented roles in DNA binding. The location of some of these proteins outside the membrane, possibly in association with wall material, is certainly consistent with these possibilities. The ComG proteins may be needed directly to carry out these steps, or they may be needed for the correct positioning or folding of other competence proteins.

Notable among these other proteins is ComEA, an integral membrane protein situated with its C-terminus outside the membrane. ComEA is required for DNA binding to the competent cell and also for the transport of DNA across the membrane (Inamine and Dubnau, 1995). Recently, we have found that ComEA is a DNA binding protein, with higher affinity for double-stranded than for single-stranded

DNA (Provvedi and Dubnau, 1998). Mutants deficient in any of the ComG proteins also fail to bind DNA (Chung and Dubnau, 1998), although ComEA is still present in such mutants (Provvedi and Dubnau, 1998). *In vitro* experiments designed to detect DNA binding to the ComG proteins have yielded uniformly negative results (Provvedi and Dubnau, 1998). If the ComG proteins do not themselves interact with DNA, perhaps they permit DNA to gain access to ComEA. The ComG proteins may assemble to form a channel that traverses the wall, or they may remodel cell wall material, in either case permitting exogenous DNA to interact with ComEA. Alternatively, they may cause or permit ComEA to fold into an active configuration. As noted above, the ComG proteins may be required for steps subsequent to binding and, indeed, ComEA and the ComG proteins may together comprise a machine that binds DNA and then presents it to the transporter/pore-forming complex, made up of ComEC and ComFA (Dubnau, 1997).

It is possible that a complex of ComGC with other competence proteins exists. Previous attempts to obtain evidence for such interactions among the pilin-like proteins using cross-linking reagents have failed in both the competence and pullulanase systems, although homodimers have been demonstrated in the cases of ComGC (Chung and Dubnau, 1994), ComGG (Fig. 7) and PulG (Pugsley and Possot, 1993). In the present report, we have shown that the ComGG dimers, but not those of ComGC, involve disulphide bond formation (Fig. 7), and that ComGC, like many of the pilin-like proteins in other systems, probably contains an intramolecular disulphide bond. However, no evidence has been obtained for disulphide bond-mediated heteromultimer formation. A possible non-trivial explanation for the failure to demonstrate heteromultimers is that a multiprotein complex assembles transiently and, as a consequence, only a fraction of the total ComG or Pul proteins are in complexes at a given time. In fact, some evidence for heterogeneity of at least ComGC, ComGD and ComGG is provided by our observations. For instance, only a minor fraction of the ComGG molecules are present in disulphide bond-mediated dimers. In addition, the mature forms of these proteins are found in several forms: as integral membrane proteins, peripherally associated with the outer face of the membrane and released when cell wall is hydrolysed. Some or all of these forms may be related to specific stages during DNA binding and uptake. It is also possible that certain of the ComG and Pul proteins do not form heteromultimers at all, but act individually, perhaps remodelling cell wall material, or as cell surface chaperone-like proteins, adjusting the conformation of other competence or secretion proteins.

Experimental procedures

Strains

All the strains used were derivatives of *B. subtilis* 168 and are described in Table 1. Strains BD2519, BD2528, BD2708 and BD2709 express competence proteins in all or nearly all the cells in the culture and were used to amplify Western blot signals, as indicated in the figure legends.

Antibody preparation and immunoblotting

Antibodies against ComGA, ComGC, ComGD, ComGE, ComGF and ComGG were raised in rabbits using the synthetic peptides listed in Table 2. The peptides for ComGA and ComGC were coupled to maleimide-activated keyhole limpet haemocyanin (Breitling and Dubnau, 1990; Chung and Dubnau, 1994), and those for ComGD, ComGE, ComGF and ComGG were coupled to maleimide-activated bovine serum albumin (BSA). In each case, a C-residue was incorporated at the N-terminus of the peptide to permit coupling. Immunoblotting was carried out as described previously (Breitling and Dubnau, 1990; Chung and Dubnau, 1994). Signals were detected using either alkaline phosphatase-conjugated secondary antibodies or by ECL (Amersham). For each antibody, the identity of the band in Western blots assigned to the cognate ComG protein was confirmed by using extracts derived from appropriate null mutants.

Cell fractionation

Cells from 100 ml of culture were harvested by centrifugation at the time of maximal competence and resuspended in 5 ml of protoplasting buffer [50 mM Tris-HCl, pH 8.0, 15 mM $MgCl_2$, 50 mM NaCl, 30% sucrose (w/v)]. Lysozyme (50 mg ml^{-1}) or mutanolysin (100 U ml^{-1}) were added together with a protease inhibitor cocktail [aprotinin (330 μg ml^{-1}), leupeptin (165 μg ml^{-1}) and pepstatin (165 μg ml^{-1})]. Mutanolysin was used in most experiments, as lysozyme migrated in gels at about the same position as the smaller ComG proteins and

Table 1. Strains.

Strain	Genotype	Source
BD630	*his leu-8 metB5*	Albano *et al.* (1987)
BD2519	*his leu-8 metB amyE::(xylR Pxyl-comK(cat)) comG-lacZ(kan)*[a]	Hahn *et al.* (1996)
BD2528	*his leu-8 metB5 [pUB110(comS)]*	Hahn *et al.* (1996)
BD2708	*his leu-8 metB5 comCΩTn917 [pUB110(comS)]*	This work
BD2709	*his leu-8 metB5 comGGΩTn917 [pUB110(comS)]*	This work

a. The *comG–lacZ* construct is inserted at the *comG* locus by single reciprocal recombination, and the strain is therefore Com^+.

Table 2. Synthetic peptides for antibody production.

Protein	Peptide[a]
ComGA	**C**DHALLKKRDMKKEE
ComGC	**C**FELDHEGQTPSLAD
ComGD	**C**YLGSGRVNVERK
ComGE	**C**LSILQTEWLHAS
ComGF	**C**TAFPVYSYLGGG
ComGG	**C**DQKQKKLLRWTE

a. The bold face C residues were added to permit coupling of the peptides to carrier proteins.

often resulted in distortion of the ComG bands. The suspension was incubated for 30–60 min at 37°C, and the cells were examined microscopically at intervals, until protoplast formation was judged to be nearly complete. The protoplasts were sedimented, and the supernatant (protoplast supernatant fraction) was retained and 1 mM phenylmethylsulphonyl fluoride (PMSF) was added. The protoplasts were lysed by resuspension in 1 ml of 50 mM Tris-HCl, pH 8.0, 5 mM $MgSO_4$, 50 mM NaCl, 10 $\mu g\,ml^{-1}$ DNase, 10 $\mu g\,ml^{-1}$ RNase, with the protease inhibitors included. The membrane fraction was sedimented at 50 000 r.p.m. for 20 min at 4°C in a TLA 100.2 rotor in a TL-100 centrifuge (Beckman) and resuspended in SDS sample buffer for electrophoresis. For some experiments, membranes were prepared using the French pressure cell as described previously (Chung and Dubnau, 1994).

Proteinase K and NaOH treatment

Treatment of membrane fractions or of cells with proteinase K or with 0.1 N NaOH (Russel and Model, 1982) was performed as described previously (Breitling and Dubnau, 1990).

Treatment with N*-ethyl maleimide*

To detect *in vivo* disulphide bonds, *N*-ethyl-maleimide, at a final concentration of 10 mM, was added to competent cultures immediately before harvesting by centrifugation. Membranes were isolated as described above, using the French pressure cell, and suspended in SDS sample buffer in the presence and absence of 1.44 M β-mercaptoethanol.

Acknowledgements

We thank R. Provvedi and other members of our laboratory for useful discussions. This work was supported by NIH grant GM43756.

References

Albano, M., Hahn, J., and Dubnau, D. (1987) Expression of competence genes in *Bacillus subtilis*. *J Bacteriol* **169:** 3110–3117.

Albano, M., Breitling, R., and Dubnau, D. (1989) Nucleotide sequence and genetic organization of the *Bacillus subtilis comG* operon. *J Bacteriol* **171:** 5387–5404.

Alm, R.A., and Mattick, J.S. (1995) Identification of a gene, *pilV*, required for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*, whose product possesses a pre-pilin-like sequence. *Mol Microbiol* **16:** 485–496.

Alm, R.A., and Mattick, J.S. (1996) Identification of two genes with prepilin-like leader sequences involved in type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. *J Bacteriol* **178:** 3809–3817.

Alm, R.A., Hallinan, J.P., Watson, A.A., and Mattick, J.S. (1996) Fimbrial biogenesis genes of *Pseudomonas aeruginosa*: *pilW* and *pilX* increase the similarity of type 4 fimbriae to the GSP protein secretion systems and *pilY1* encodes a gonococcal PilC homolog. *Mol Microbiol* **22:** 161–173.

Breitling, R., and Dubnau, D. (1990) A pilin-like membrane protein is essential for DNA binding by competent *Bacillus subtilis*. *J Bacteriol* **172:** 1499–1508.

Campbell, E.A., Choi, S., and Masure, H.R. (1998) A competence regulon in *Streptococcus pneumoniae* revealed by genomic analysis. *Mol Microbiol* **27:** 929–939.

Chung, Y.S., and Dubnau, D. (1994) ComC is required for the processing and translocation of ComGC, a pilin-like competence protein of *Bacillus subtilis*. *Mol Microbiol* **15:** 543–551.

Chung, Y.S., and Dubnau, D. (1998) All seven *comG* open reading frames are required for DNA binding during the transformation of competent *Bacillus subtilis*. *J Bacteriol* **180:** 41–45.

Clifton, S.W., McCarthy, D., and Roe, B.A. (1994) Sequence of the *rec-2* locus of *Haemophilus influenzae*: homologies to *comE*-ORF3 of *Bacillus subtilis* and *msbA* of *Escherichia coli*. *Gene* **146:** 95–100.

Dubnau, D. (1993) Genetic exchange and homologous recombination. In Bacillus subtilis *and Other Gram-positive Bacteria: Biochemistry, Physiology and Molecular Genetics*. Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: American Society for Microbiology Press, pp. 555–584.

Dubnau, D. (1997) Binding and transport of transforming DNA by *Bacillus subtilis*: the role of type-IV pilin-like proteins – a review. *Gene* **192:** 191–198.

Dupuy, B., Taha, M.-K., Possot, O., Marchal, C., and Pugsley, A.P. (1992) PulO, a component of the pullulanase secretion pathway of *Klebsiella oxytoca*, correctly and efficiently processes gonococcal type IV prepilin in *Escherichia coli*. *Mol Microbiol* **6:** 1887–1894.

Facius, D., and Meyer, T.F. (1993) A novel determinant (*comA*) essential for natural transformation competence in *Neisseria gonorrhoeae*. *Mol Microbiol* **10:** 699–712.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269:** 496–512.

Freitag, N.E., Seifert, H.S., and Koomey, M. (1995) Characterization of the *pilF–pilD* pilus-assembly locus of *Neisseria gonorrhoeae*. *Mol Microbiol* **16:** 575–586.

Fussenegger, M., Rudel, T., Barten, R., Ryell, R., and Meyer, T.F. (1997) Transformation competence and type-4 pilus biogenesis in *Neisseria gonorrhoeae* – a review. *Gene* **192:** 125–134.

Hahn, J., Albano, M., and Dubnau, D. (1987) Isolation and

characterization of competence mutants in *Bacillus subtilis*. *J Bacteriol* **169:** 3104–3109.

Hahn, J., Luttinger, A., and Dubnau, D. (1996) Regulatory inputs for the synthesis of ComK, the competence transcription factor of *Bacillus subtilis*. *Mol Microbiol* **21:** 763–775.

Hobbs, M., and Mattick, J.S. (1993) Common components in the assembly of type-4 fimbriae, DNA transfer systems, filamentous phage and protein secretion apparatus; a general system for the formation of surface-associated protein complexes. *Mol Microbiol* **10:** 233–243.

Inamine, G.S., and Dubnau, D. (1995) ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *J Bacteriol* **177:** 3405–3051.

Kaback, H.R. (1971) Bacterial membranes. *Methods Enzymol* **13:** 99–120.

Kerppola, R.E., Shyamala, V.K., Klebba, P., and Ames, G.F.-L. (1991) The membrane bound proteins of periplasmic permeases form a complex. *J Biol Chem* **266:** 9857–9865.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.V.A., and Bertero, M.G. (1997) The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390:** 249–256.

Linderoth, N.A., Simon, M.N., and Russel, M. (1997) The filamentous phage pIV multimer visualized by scanning transmission electron microscopy. *Science* **278:** 1653–1638.

Lunsford, R.D., and Roble, A.G. (1997) comYA, a gene similar to comGA of *Bacillus subtilis*, is essential for competence-factor-dependent DNA transformation in *Streptococcus gordonii*. *J Bacteriol* **179:** 3122–3126.

Mattick, J.S., and Alm, R.A. (1995) Common architecture of type-4 fimbriae and complexes involved in macromolecular traffic. *Trends Microbiol* **3:** 411–413.

Mohan, S., Aghion, J., Guillen, N., and Dubnau, D. (1989) Molecular cloning and characterization of *comC*, a late competence gene of *Bacillus subtilis*. *J Bacteriol* **171:** 6043–6051.

Nunn, D., Bergman, S., and Lory, S. (1990) Products of three accessory genes, *pilB*, *pilC*, and *pilD*, are required for biogenesis of *Pseudomonas aeruginosa* pili. *J Bacteriol* **172:** 2911–2919.

Pasloske, B.L., and Paranchych, W. (1988) The expression of mutant pilins in *Pseudomonas aeruginosa*: fifth position glutamate affects pilin methylation. *Mol Microbiol* **2:** 489–495.

Pestova, E.V., and Morrison, D.A. (1997) *Streptococcus pneumoniae* chromosomal loci specifically induced during competence for genetic transformation. In *Abstracts of the 97th General Meeting of the American Society for Microbiology*. p. 293. American Society for Microbiology: Washington, DC.

Possot, O., d'Enfert, C., Reyss, I., and Pugsley, A.P. (1992) Pullulanase secretion in *Escherichia coli* K-12 requires a cytoplasmic protein and a putative polytopic cytoplasmic membrane protein. *Mol Microbiol* **6:** 95–105.

Provvedi, R., and Dubnau, D. (1998) ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Mol Microbiol* (submitted).

Pugsley, A.P. (1993) The complete general secretory pathway in Gram-negative bacteria. *Microbiol Rev* **57:** 50–108.

Pugsley, A.P. (1996) Multimers of the precursor of a type IV pilin-like component of the general secretory pathway are unrelated to pili. *Mol Microbiol* **20:** 1235–1245.

Pugsley, A.P., and Possot, O. (1993) The general secretory pathway of *Klebsiella oxytoca*: no evidence for relocalization or assembly of pilin-like PulG protein into a multiprotein complex. *Mol Microbiol* **10:** 665–674.

Pugsley, A.P., Francetic, O., Possot, O.M., Sauvonnet, N., and Hardie, K.R. (1997) Recent progress and future directions in studies of the main terminal branch of the general secretory pathway in Gram-negative bacteria – a review. *Gene* **192:** 13–19.

Reyss, I., and Pugsley, A.P. (1990) Five additional genes in the *pulC–O* operon of the Gram-negative bacterium *Klebsiella oxytoca* UNF5023 which are required for pullulanase secretion. *Mol Gen Genet* **222:** 176–184.

Russel, M. (1994) Phage assembly: a paradigm for bacterial virulence factor export? *Science* **265:** 612–614.

Russel, M., and Model, P. (1982) Filamentous phage precoat is an integral membrane protein: analysis by a new method of membrane preparation. *Cell* **28:** 177–184.

Russell, M.A., and Darzins, A. (1994) The *pilE* gene product of *Pseudomonas aeruginosa*, required for pilus biogenesis, shares amino acid sequence identity with the N-terminus of type-4 prepilin proteins. *Mol Microbiol* **13:** 973–985.

Russel, M., Linderoth, N.A., and Sali, A. (1997) Filamentous phage assembly: variation on a protein export theme. *Gene* **192:** 23–32.

Sandkvist, M., Bagdasarian, M., Howard, S.P., and DiRita, V.J. (1995) Interaction between the autokinase EpsE and EpsL in the cytoplasmic membrane is required for extracellular secretion in *Vibrio cholerae*. *EMBO J* **14:** 1664–1673.

Strom, M.S., and Lory, S. (1991) Amino acid substitutions in pilin of *Pseudomonas aeruginosa*. Effect on leader peptide cleavage, amino terminal methylation, and pilus assembly. *J Biol Chem* **266:** 1656–1664.

Strom, M.S., Nunn, D.N., and Lory, S. (1993) A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proc Natl Acad Sci USA* **90:** 2404–2408.

Tønjum, T., Freitag, N.E., Namork, E., and Koomey, M. (1995) Identification and characterization of *pilG*, a highly conserved pilus-assembly gene in pathogenic *Neisseria*. *Mol Microbiol* **16:** 451–464.