

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

Y. S. Chung, F. Breidt<sup>†</sup> 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 (Facijs 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; Breittling 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. <sup>†</sup>Present address: 322 Schaub Hall, Box 7624, NC State University Raleigh, Raleigh, 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; Breittling 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 (Breittling 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

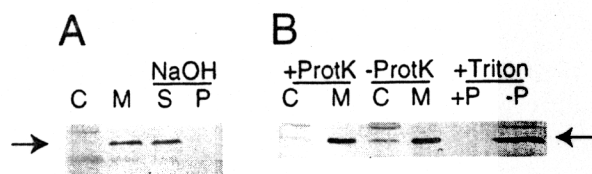


Fig. 1. Localization of ComGA.

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 ( $2.5 \text{ mg ml}^{-1}$ ). 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.

