Molecular Microbiology (1994) **12**(2), 333–334

MicroCorrespondence

The *C*-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both Gram-positive and Gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan

Sir,

The major outer membrane protein OprF from *Pseudomonas* species displays strong homology to several outer membrane proteins from unrelated species, including OmpA from enteric bacteria (De Mot *et al.*, 1992, *Mol Gen Genet* **231:** 489–493). However, this homology is confined to the respective *C*-terminal regions (about 100–140 residues), and there is no obvious similarity between the *N*-terminal regions. Such remarkable intergeneric sequence conservation presumably reflects a similar, but as yet unidentified, function of this domain. For both OmpA of *Escherichia coli*, and OprF of *Pseudomonas aeruginosa*, a structural role in stabilizing the outer membrane has been proposed (Gotoh *et al.*, 1989, *J Bacteriol* **171:** 983–990; Woodruff and Hancock, 1989, *J Bacteriol* **171:** 3304–3309). In addition, pore-forming activity has been demonstrated for both proteins (Nikaido *et al.*, 1991, *J Biol Chem* **266:** 770–779; Sugawara and Nikaido, 1992, *J Biol Chem* **267:** 2507–2511). The extended *C*-terminal homology is also found in lipoproteins that are tightly, but non-covalently bound to peptidoglycan. These peptidoglycan-associated lipoproteins (PALs) are important structural elements for the cell envelope (Lazzaroni and Portalier, 1992, *Mol Microbiol* **6:** 735–742). The functions of the other outer membrane proteins in this family are poorly characterized. However, it is noteworthy that for several of them strong, non-covalent association with peptidoglycan has been described (Lugtenberg and van Alphen, 1983, *Biochim Biophys Acta* **737:** 51–115; Hancock *et al.*, 1990, *Mol Microbiol* **4:** 1069–1075), although the protein domains interacting with the peptidoglycan layer remain to be identified.

Here we show that the *C*-terminal sequence homology extends beyond proteins from the outer membrane of Gram-negative bacteria, and is also found in a number proteins from Gram-positive species. The latter bacteria lack the outer membrane structure, but they do contain peptidoglycan in their cell envelopes (Shockman and Barrett, 1983, *Annu Rev Microbiol* **27:** 501–527). As outlined below, the conserved *C*-terminal region may be important for the interaction with peptidoglycan.

Using the BLASTP program (Altschul *et al.*, 1990, *J Mol Biol* **215:** 403–410), we searched the PDB, SwissProt, PIR and GenPept data bases for novel proteins homologous to the *C*-terminal part of *Pseudomonas fluorescens* OprF (106 residues). Significant extended *C*-terminal homology was detected in 28 other proteins and a multiple alignment of 14 representative sequences was generated (Fig. 1). Their hydropathy profiles were also quite similar and indicated that no amino acid stretches were present that were sufficiently long or hydrophobic enough to span the cytoplasmic membrane (data not shown).

To our surprise, the updated list contained three proteins from the Gram-positive species *Bacillus subtilis* (MotB, OrfB) and *Bacillus megaterium* (OrfB) (Fig. 1). MotB is required for the motility of *B. subtilis* (Mirel *et al.*, 1992, *J Bacteriol* **174:** 4197–4204), but the function of the OrfB proteins remains to be determined. In *E. coli*, MotB links the rotational machinery of the flagellar motor to the cell wall by interacting with both the energizing integral membrane protein MotA (through the highly hydrophobic membrane-anchored *N*-terminal part of MotB) and the peptidoglycan layer (through the hydrophilic *C*-terminal part of MotB) (Blair *et al.*, 1991, *J Bacteriol* **173:** 4049–4055). Interestingly, characterization of a number of mutant MotB proteins of *E. coli* revealed that most mutations affecting MotB function were located in the *C*-terminal domain (Blair *et al.*, ibid.). As shown in Fig. 1, the hydroxylamine-induced amino acid substitutions occurred mainly in the regions of MotB that are most conserved among the various OmpA-related proteins. Two of the three perfectly conserved arginine residues were essential for MotB function. These basic residues are potential sites for interaction with the carboxylates present on peptidoglycan. The observation that the cell envelopes of mutants lacking a major outer membrane protein such as OmpA, OprF, or PAL are far more fragile may be explained by a diminished stabilizing interaction with peptidoglycan. On the other hand, there are no indications in the literature that such mutants would be affected in motility.

In conclusion, the extent of *C*-terminal sequence conservation between MotB from both *B. subtilis* and *E. coli*, and several peptidoglycan-associated outer membrane proteins from a wide range of Gram-negative species suggests a common function, probably in the interaction with peptidoglycan.

```
             *                                  **                *  *    *
MotB[Ec]  FRTGSADVEPYMRDILRAIAPVLNGI---PNRISLSGHTDDFPYASGEKGYSNWELSADRANASRRELMVG-GL 230
MotB[Bs]  FDSGKATIRKEDVPLAKEISNLLVINP--PRNIIISGHTDNMPIKNS-EFQSNWHLSVMRAVNFMGLLIEN-PK 214
OrfB[Bs]  FDTGEAKVLKNAETLLHQIAVLLQTI---PNDIQVEGHTDSRNISTY-RYPSNWELSAARASGVIQYFTSK-EK 198
OrfB[Bm]  FETGQADILKKGTPFLDELGRLFSTI---PNDIKVEGHTDNRPIHTY-AYPSNWELSAARASGVIRYLTNH-FS 176
Omp5[Hi]  FAFGKANLKPQAQATLDSIYGEMSQVK--SAKVAVAGYTDRIGSDAF-----NVKLSQERADSVANYFVAK-GV 305
OmpA[Ec]  FNFNKATLKPEGQAALDQLYSQLSNLDPKDGSVVVLGYTDRIGSDAY-----NQGLSERRAQSVVDYLISK-GI 290
Omp3[Ng]  FGFDKDSLRAEAQDNLKVLAQRLSRTN--VQSVRVEGHTDFMGSEKY-----NQALSERRAYVVANNLVSN-GV 164
OmpX[Hs]  FDFDQDTLTSKGEEAVDNVAMQLEAFS--AKEIKIVGFTDRLGTDSY-----NLDLSQRRADRVKERLIEK-GL 207
Pal [Ec]  FDLDKYDIRSDFAQMLDAHANFLRSNP--SYKVTVEGHADERGTPEY-----NISLGERRANAVKMYLQGK-GV 138
Pal [Lp]  FAYDDSTLASKYLPSVNAQAEYLKTHP--GARVMIAGHTDERGSREY-----NVALGERRADTVAEILRMA-GV 137
Omp [Ba]  FDFDKSTLKPEGRQLLDQVAQQARAID--LETIIAVGNTDSIGTEAY-----NMKLSERRAASVKAYLVSK-GI 154
OprF[Pf]  FDFDKSVVKPNSYGDVKNLADFMAQYP--ATNVEVAGHTDSIGPDAY-----NQKLSQRRADRVKQVLVKD-GV 286
CD  [Bc]  FDYDKSIIKPQYREEVAKVAAQMREFP--NATATIEGHASRDSARSSARY--NQRLSEARANAVKSMLSNEFGI 401
YFIB[Ec]  FAKNDYKLLPESQQQIQTMAAKLASTG--LTHARMDGHTDNYGEDSY-----NEGLSLKRANVVADAWAMGGQI 122
          •                                   •               •  •   ••

                     *                                     *
MotB[Ec]  DSGKVLRVVGMAATMRLSDRGPDD----------------------AVNRRISLLVLNKQAEQAILHENAESQN 282
MotB[Bs]  LDAKVFSAKGYGEYKPVASNKTAEGR--------------------SKNRRVEVLILPRGAAETNEK        261
OrfB[Bs]  LPSKRFIAVGYADTKPVKDNKTNEHM--------------------KENRRVEIVIKKSKTTSS           242
OrfB[Bm]  LSANRFEALGYGDTKPLVPNTSNDNL--------------------QKNRRVEIIISDPEAQ             218
Omp5[Hi]  AADAI-SATGYGKANPVTGATCDQVKGRK-----------ALIACLAPDRRVEIAVNGTK               353
OmpA[Ec]  PADKI-SARGMGESNPVTGNTCDNVKQRA-----------ALIDCLAPDRRVEIEVKGIKDVVTQPQA       346
Omp3[Ng]  PASRI-SAVGLGESQAQMTQVCQAEVAKLGAKASKAKKREALIACIEPDRRVDVKIRSIVTRQVVPARNHHQH  138
OmpX[Hs]  NIDII--AIGYGKTQQI--KACNDVPAKE------------LKDCLRPNRRVEISAYGNISKKYGNGELKGGTT 264
Pal [Ec]  SADQI-SIVSYGKEKPAVLGHDEAAY--------------------SKNRRAVLVY                   173
Pal [Lp]  SRQQI-RVVSYGKERPANYGHDEASH--------------------AQNRRVEFIYEATR               189
Omp [Ba]  DPNRI-YTEGKGKLNPIASNKTAEGR--------------------ARNRRVEIEIVGSRK              194
OprF[Pf]  APSRI-TAVGYGESRPVADNATEAGR--------------------AVNRRVEASVEAQAQ              326
CD  [Bc]  APNRL-NAVGYGFDRPIAPNTTAEGK--------------------AMNRRVEAVITGSKTTTVDQTKDMIVQ  453
YFIB[Ec]  PRSNL-TTQGLGKKYPIASNKTAQGR--------------------AENRRVAVVITTP                160
                                                           ••
```

Fig. 1. Multiple alignment (CLUSTAL; PC Gene, IntelliGenetics) of the *C*-terminal regions of OmpA-related proteins from *Bordetella avium* (*Ba*), *Branhamella catarrhalis* (*Bc*), *B. subtilis* (*Bs*), *E. coli* (*Ec*), *Haemophilus influenzae* (*Hi*), *Haemophilus somnus* (*Hs*), *Legionella pneumophila* (*Lp*), *Neisseria gonorrhoeae* (*Ng*), and *P. fluorescens* (*Pf*). Omp5 (Munson *et al.*, 1993, *Infect Immun* **61:** 4017–4020), OmpA (Beck and Bremer, 1980, *Nucl Acids Res* **8:** 3011–3027), Omp3 (Gotschlich *et al.*, 1987, *J Exp Med* **165:** 471–482), OmpX (Won and Griffith, 1993, *Infect Immun* **61:** 2813–2821), Omp (Gentry-Weeks *et al.*, 1992, *J Bacteriol* **174:** 7729–7742), OprF (De Mot *et al.*, 1992, *Mol Gen Genet* **231:** 489–493), CD (Murphy *et al.*, 1993, *Mol Microbiol* **10:** 87–97) and PAL (Chen and Henning, 1987, *Eur J Biochem* **163:** 73–77; Engleberg *et al.*, 1991, *Mol Microbiol* **5:** 2021–2029) are (major) outer membrane proteins. YFIB is a putative lipoprotein (Byström *et al.*, 1983, *EMBO J* **2:** 899–905). Proteins with high homology (>70%) to one of the above-listed proteins were not included in the alignment: the OmpA homologues from 10 enterobacterial species (Lawrence *et al.*, 1991, *J Gen Microbiol* **137:** 1911–1921), the Omp5-homologous fimbrial protein from *H. influenzae* (L08448), the Omp3 homologue from *N. meningitidis* (Klugman *et al.*, 1989, *Infect Immun* **57:** 2066–2071), the PAL from *H. influenzae* (Deich *et al.*, 1988, *J Bacteriol* **170:** 489–498; Nelson *et al.*, 1988, *Infect Immun* **56:** 128–134), and the OprFs from *P. aeruginosa* (Duchêne *et al.*, 1988, *J Bacteriol* **170:** 155–162) and *Pseudomonas syringae* (Ullstrom *et al.*, 1991, *J Bacteriol* **173:** 768–775). The MotB proteins are required for flagellar motion (Stader *et al.*, 1986, *J Bacteriol* **166:** 244–252; Mirel *et al.*, 1992, *J Bacteriol* **174:** 4197–4204), but the functions of their distant OrfB homologues from *B. subtilis* (Grundy *et al.*, 1993, *Mol Microbiol* **10:** 259–271) and *B. megaterium* (GenBank accession number L26052) are not known. The complete *C*-terminal regions of the respective proteins are shown except for MotB(*Ec*) and OmpX (last 26 and 8 residues, respectively, not included). Identical or similar residues (F-W-Y; V-L-I-M; E-D; N-Q; K-R; S-T-A) occurring in a certain position of at least 8 out of the 14 protein sequences are in bold to highlight the most conserved regions. Dashes represent gaps introduced for optimal alignment. Residues changed in non-functional MotB mutants of *E. coli* (see text) are marked with asterisks. Positions that are perfectly conserved throughout the 29 sequences are labelled with dots.

René De Mot and Jos Vanderleyden
F. A. Janssens Laboratory of Genetics, Catholic University of Leuven, W. De Croylaan 42, B-3001 Heverlee, Belgium.
Tel (16) 220921; Fax (16) 200720;
E-mail: Janssens%faj%agr@CC3.KULEUVEN.AC.BE.
Received 20 December, 1993.