Topology of the *Erwinia chrysanthemi* oligogalacturonate porin KdgM

Teijo PELLINEN, Helena AHLFORS, Nicolas BLOT[1] and Guy CONDEMINE[2]

Unité Microbiologie et Génétique, UMR CNRS-INSA-UCB 5122, Batiment Lwoff, 10 rue Raphaël Dubois, 69622 Villeurbanne, France

The *Erwinia chrysanthemi* oligogalacturonate-specific monomeric porin, KdgM, does not present homology with any porins of known structure. A model of this protein, based on sequence similarity and the amphipathy profile, was constructed. The model depicts a β-barrel composed of 14 antiparallel β-strands. The accuracy of this model was tested by the chemical labelling of cysteine residues introduced by site-directed mutagenesis. The protein has seven surface-exposed loops. They are rather small with the exception of one, loop L6. Deletion of this loop allowed the entry of maltopentaose into the bacteria, a molecule too large to enter through the wild-type KdgM. Loop L6 could fold back into the lumen of the pore and play the role of the constriction loop L3 of general porins. With 14 transmembrane segments, the KdgM porin family could represent the smallest porin characterized to date.

Key words: β-barrel, membrane protein, pectin degradation.

INTRODUCTION

In Gram-negative bacteria, exchange of ions and small molecules between the external medium and the periplasm occurs through β-barrel pore-forming proteins. Among them, the porins are one-domain β-barrel proteins inserted in the outer membrane that form water-filled channels, allowing the diffusion of molecules from the extracellular medium into the periplasm [1]. General porins allow the diffusion of hydrophilic molecules, with a size < 600 Da, without any particular substrate specificity. Substrate-specific porins are specific for one class of solutes. Molecules that could not cross the outer membrane, or cross only very slowly through general porins, can use these specific channels to enter into the periplasm.

The three-dimensional structure of several porins has been determined: structures of the *Escherichia coli* general porins OmpF and PhoE [2], and of the substrate-specific porins LamB [3] and ScrY [4] are known. The *E. coli* general porins OmpF and PhoE are trimeric. In each monomer, 16 antiparallel β-strands cross the outer membrane and form a barrel. The β-strands are linked on the periplasmic side by short turns, whilst on the external side they are connected by loops of varying length. Loop L3 is longer and folds into the mouth of the channel, forming a constriction zone. This reduction of the size of the pore contributes to the substrate exclusion limit and to the selectivity of the pore. Moreover, the presence of charged amino acids in the constriction zone determines the ion selectivity of the pore: OmpF is cation-selective while PhoE is anion-selective [5]. Loop L2 folds back into the channel of a neighbouring subunit, contributing not only to the stabilization of the trimer but also to the selectivity of the pore [6]. The substrate-specific maltoporin, LamB, can transport maltodextrins specifically, while the ScrY substrate is sucrose. These two porins, whose crystal structures can be superimposed, also form homotrimers but their monomers consist of 18 antiparallel β-strands [3,4]. As observed with the general porins, loop L3 folds back inside the pore. In the maltoporin, a row of aromatic amino acids guides the maltodextrins inside the channel [7].

Recently, monomeric porins have been described. CymA is a maltodextrin-specific porin of *Klebsiella oxytoca* [8]. OmpG has been characterized in *E. coli*, and homologues exist in other Gram-negative bacteria. OmpG does not seem to have a substrate specificity and forms unusually large channels [9,10]. The oligogalacturonate-specific porin, KdgM, of *Erwinia chrysanthemi* is the first member to be described of a family of proteins found also in *Yersinia pestis*, *E. coli* and *Salmonella typhimurium*. It is thought that these porins could have oligosaccharides as substrates [11]. The structure of any of these monomeric porins is not known. It has been suggested that OmpG may possess 14 transmembrane β-segments [12] because the size of the barrel appears to be smaller than that of OmpF and OmpC, but no experimental data have confirmed this hypothesis. The size of proteins of the KdgM family is much smaller than that of other porins: KdgM is 216 amino acids long, compared with amino acid lengths of 280 for OmpG, 324 for CymA, 340 for OmpF and 421 for LamB. Thus the number of transmembrane segments of KdgM could also be less than that of the other known porins. In this study a model of KdgM topology was constructed. To test the model, cysteine residues were introduced into the cysteine-free protein and these residues were chemically labelled in intact or permeabilized cells. The predicted model, in which KdgM has 14 transmembrane segments, was confirmed by the results obtained during the present study.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

E. coli NM522 [*supE, thi-1, Δ*(*lac-proAB*) *Δ*(*hsdSM-mcrB*)*5* (r_k^-, m_k^-)], BL21(DE3)omp8 (*ΔlamB ompF*::Tn5 *ΔompA ΔompC*) [13] and MB1 [*araD139 Δ*(*argF-lac*)*U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rsbR22*], from the *E. coli* Genetic Stock Center (Yale University, New Haven, CT, U.S.A.), were used as the hosts for expressing *kdgM* mutant plasmids. KdgM activity tests were carried out with *E. chrysanthemi* strain A3573 (*lmrT*^{c} *lacZ kdgM*::*uidA-Km*) [11]. The pUC18-based plasmid pKM2 containing *kdgM* [11], was used as the source for site-directed mutagenesis. *E. chrysanthemi* and *E. coli* cells were grown in Luria–Bertani (LB) medium at 30 °C and 37 °C, respectively. KdgM activity tests were performed on M63 agar

Abbreviations used: LB, Luria–Bertani; PEO-MAB, EZ-Link™ PEO maleimide-activated biotin.

[1] Present address: Max-Planck-Institut für Terrestrische Mikrobiologie, Karl-von-Frisch-Strasse, 35043 Marburg, Germany.

[2] To whom correspondence should be addressed (e-mail guy.condemine@insa-lyon.fr).

plates supplemented with 0.4 % polygalacturonate and 0.2 % glycerol. Pectate lyase activity was detected by flooding the plates with a saturated solution of copper acetate. When required, ampicillin (100 μg/ml) was used in selective media.

Antibiotic-sensitivity assay

Sensitivity to antibiotics and drugs was examined by measuring the zone of growth inhibition around antibiotic-containing discs. Strains were grown to late exponential phase, 0.1 ml of culture was mixed with 5 ml of 0.4 % agar in LB medium and poured over LB agar plates containing ampicillin (50 μg/ml). Paper discs (6-mm-diameter) containing 10 μl of antibiotic solution were placed on top of the lawn of cells. The plates were incubated at 30 °C overnight and the zone of growth inhibition around the disc was measured.

Molecular biology techniques

Plasmid DNA preparation, DNA restriction and electrophoresis, and bacterial electroporation were all performed according to Sambrook et al. [14]. Site-directed mutagenesis was performed with the QuikChange mutagenesis Kit (Stratagene, La Jolla, CA, U.S.A.). The mutagenesis was performed with mutagenic oligonucleotides containing mismatches for the desired amino acid replacements and silent mismatches to introduce or remove restriction-enzyme cleavage sites. The oligonucleotides used for site-directed mutagenesis are shown in Table 1. To construct loop L6 mutants, *Xba*I restriction sites were introduced at nucleotides 516 or 554 of *kdgM*. After cutting the two mutated plasmids with *Xba*I and *Eco*RI, the fragments encoding the C-terminal part of *kdgM* were swapped and ligated to the other plasmid to create *kdgMΔ 154–165* and *kdgMrep153–164*.

Biotin labelling of cysteine residues

E. coli cells, containing mutagenized plasmids, were grown in ampicillin-supplemented (100 μg/ml) LB medium. When a D_{600} of 1.0 was reached, two aliquots (2.5 ml) of cells were harvested by centrifugation (12000 ***g*** for 2 min), for biotin labelling of intact cells and permeabilized cells. To prepare the permeabilized cells, one aliquot of cells was resuspended in 150 μl of 0.2 M Tris/HCl, pH 8.0, and sucrose was added to a final concentration of 20 % (w/v). After centrifugation (12000 ***g*** for 2 min) the pellet was resuspended in 200 μl of 100 mM Tris/HCl, pH 8.0, containing 20 % (w/v) sucrose. Lysozyme was added to a final concentration of 1 mg/ml. After a 10 min incubation at 20 °C, EDTA was added to a final concentration of 5 mM, and incubation was continued for another 10 min. The permeabilized cells were centrifuged (12000 ***g*** for 2 min) and the pellet was gently resuspended in 1 ml of buffer A (100 mM phosphate buffer, pH 7.0, containing 20 % sucrose and 5 mM $MgCl_2$). The bacteria from the second aliquot, to be used for intact-cell biotin labelling, were washed once with 1 ml of buffer A and resuspended after centrifugation (12000 ***g*** for 2 min) in 1 ml of the same buffer. The biotin labelling of cysteine residues was carried out by adding 100 μl of 5 mM EZ-Link™ PEO maleimide activated biotin (PEO-MAB; Pierce Chemical Company, Rockford, IL, U.S.A.) to 1 ml of bacterial suspension and gently agitating for 15 s (for intact cells) or for 60 s (for permeabilized cells). The reactions were terminated by adding 50 μl of 1 M dithiothreitol (final concentration of 50 mM) and agitating for 60 s. The bacterial samples were washed twice with 1 ml of buffer A to eliminate unbound biotin and dithiothreitol. After the final centrifugation (12000 ***g*** for 2 min) the pellets were resuspended in 100 μl of 0.8 % SDS and boiled for 5 min to lyse the cells.

Table 1 Primers used for site-directed mutagenesis

Mutated amino acids are denoted by their single-letter abbreviations and positions in the sequence, followed by + and − to indicate if the primer was identical to the coding or to the non-coding strand of DNA, respectively. The mutated nucleotides are in bold and underlined.

Primer	Sequence
T12 +	CGTCATGAAATGCA**T**GAT**TG**CGCTCAGGCCGGTC
T12−	GACCGGCCTGAGCG**CA**ATC**A**TGCATTTCATGACG
R20 +	CCGGTCATAAAGAT**T**GTCT**A**TTAATCTCTCACCG
R20−	CGGTGAGAGATTAA**T**AGAC**A**ATCTTTATGACCGG
N29 +	GTTAATCTCTCACCGTTTTGC**ATG**TGGTTTCGGATTATCTTCC
N29−	GGAAGATAATCCGAAACCA**CAT**GCAAAACGGTGAGAGATTAAC
G30 +	GTTAATCTCTCACCGTTTTGC**C**AAT**T**GTTTCGGATTATCTTCCG
G30−	CGGAAGATAATCCGAAAC**A**ATT**G**GCAAAACGGTGAGAGATTAAC
K38 +	GGATTATCTTCCGAAGTA**TGT**TGGGC**C**CAATCCAGTGCGG
K38−	CCGCACTGGATTG**G**GCCCA**ACA**TACTTCGGAAGATAATCC
K46 +	GCAATCCAGTGCGGAT**TGT**AC**A**CCAAATAAACCATTTAATG
K46−	CATTAAATGGTTTATTTGG**T**GT**ACA**ATCCGCACTGGATTGC
K50 +	CCAGTGCGGATAAAACCCC**C**AAT**TGT**CAATTTAATGAGCAGGTC
K50−	GACCTGCTCATTAAATGG**ACA**ATT**G**GGGGTTTTATCCGCACTGG
G59 +	GCAGGTCAGTAAC**T**G**T**AC**A**GAAGTTGTCGCC
G59−	GGCGACAACTTC**T**GT**A**C**A**GTTACTGACCTGC
S72 +	CGTTTATAAATTCAAT**T**GCGTCTTTTCTATCGAACC
S72−	GGTTCGATAGAAAAGACGC**A**ATTGAATTTATAAACG
S86 +	CCAGGTTTTTCTTTAGAATC**C**GG**A**T**G**TTCAAATAACAACTATCG
S86−	CGATAGTTGTTATTTGAA**C**AT**C**C**G**GATTCTAAAGAAAAAACCTGG
N100 +	CCATACCTGCGTGGACGCGC**ATGC**GTGACCGATGATTTGTCTG
N100−	CAGACAAATCATCGGTCAC**GCAT**GCGCGTCCACGCAGGTATGG
S106 +	GTGACCGATGATTTGT**G**TGT**A**GC**GC**TGCGTTATCGCCC
S106−	GGGCGATAACGCA**GC**GC**T**ACA**C**ACAAATCATCGGTCAC
K116 +	CGTTATCGCCCTTACTT**T**AAA**T**GTAATAGCGGTAATATTGG
K116−	CCAATATTACCGCTATTAC**A**TTT**A**AAGTAAGGGCGATAACG
K124 +	CGTAATAGCGGTAATAT**A**GGT**TGT**GATAACACCATGGACAAGGG
K124−	CCCTTGTCCATGGTGTTATC**ACA**ACC**T**ATATTACCGCTATTACG
D129 +	GGTAAAGATAACACCATG**TG**CAAGGGTTATACCCTGACC
D129−	GGTCAGGGTATAACCCTTG**CA**CATGGTGTTATCTTTACC
D139 +	CCCTGACCGGTAATATC**TG**TTATACCTTCCTGAAGGATTAC
D139−	GTAATCCTTCAGGAAGGTATAA**CA**GATATTACCGGTCAGGG
T141 +	CGGTAATATCGATTAT**TG**CTTCCT**T**AAGGATTACACCATCG
T141−	CGATGGTGTAATCCTT**A**AGCAAG**CA**ATAATCGATATTACCG
S159 +	GGAATACAAAAAAGGTAC**C**T**G**TGGCAAAACTATTCTGTC
S159−	GACAGAATAGTTTTGCCA**C**A**G**GTACCTTTTTTGTATTCC
Y170 +	CTGTCTGACAACGATGATT**G**TGA**T**ATCACTCATAACGTGAAATTG
Y170−	CAATTTCACGTTATGAGTGAT**A**TCA**C**AATCATCGTTGTCAGACAG
I172 +	GACAACGATGATTATGAC**TG**CACTCATAACGTGAAATTG
I172−	CAATTTCACGTTATGAGTG**CA**GTCATAATCATCGTTGTC
K184 +	GAAATTGTCTTATAAATGGGA**CTGC**AACTGGAAACCGTATGTTG
K184−	CAACATACGGTTTCCAGTT**GCAG**TCCCATTTATAAGACAATTTC
K187 +	CTTATAAATGGGATAAAAACTGG**TGT**CC**A**TATGTTGAAGTCGGCAAC
K187−	GTTGCCGACTTCAACATA**T**GG**ACA**CCAGTTTTTATCCCATTTATAAG
S198 +	CGGCAACGTTTC**C**GG**A**T**G**TGAAACCACCGATGAGC
S198−	GCTCATCGGTGGTTTCA**C**AT**C**C**G**GAAACGTTGCCG
R204 +	CTGAAACCACCGATGA**AT**GCCAGACGCGTTACCG
R204−	CGGTAACGCGTCTGGC**AT**TCATCGGTGGTTTCAG
Loop6–1	TACACCATCGGCTATGA**T**CT**A**GAATACAAAAAAGG
Loop6–2	CCTTTTTTGTATTC**T**AG**A**TCATAGCCGTGGTGTA
Loop6–3	TCTGGCAAAACTATTCT**AGA**TGACAACGATGATTA
Loop6–4	TAATCATCGTTGTCA**TCT**AGAATAGTTTTGCCAGA

Immunoprecipitation of KdgM

To eliminate other PEO-MAB-labelled proteins, KdgM was immunoprecipitated. Anti-KdgM antibodies (5 μl), diluted 1/1000 in TBS-T (20 mM Tris/HCl, pH 8.0, 140 mM NaCl and 0.1 % Tween 20) and saturated with an *E. coli* lysate, were added to the PEO-MAB-treated cells, lysed with SDS (as above)

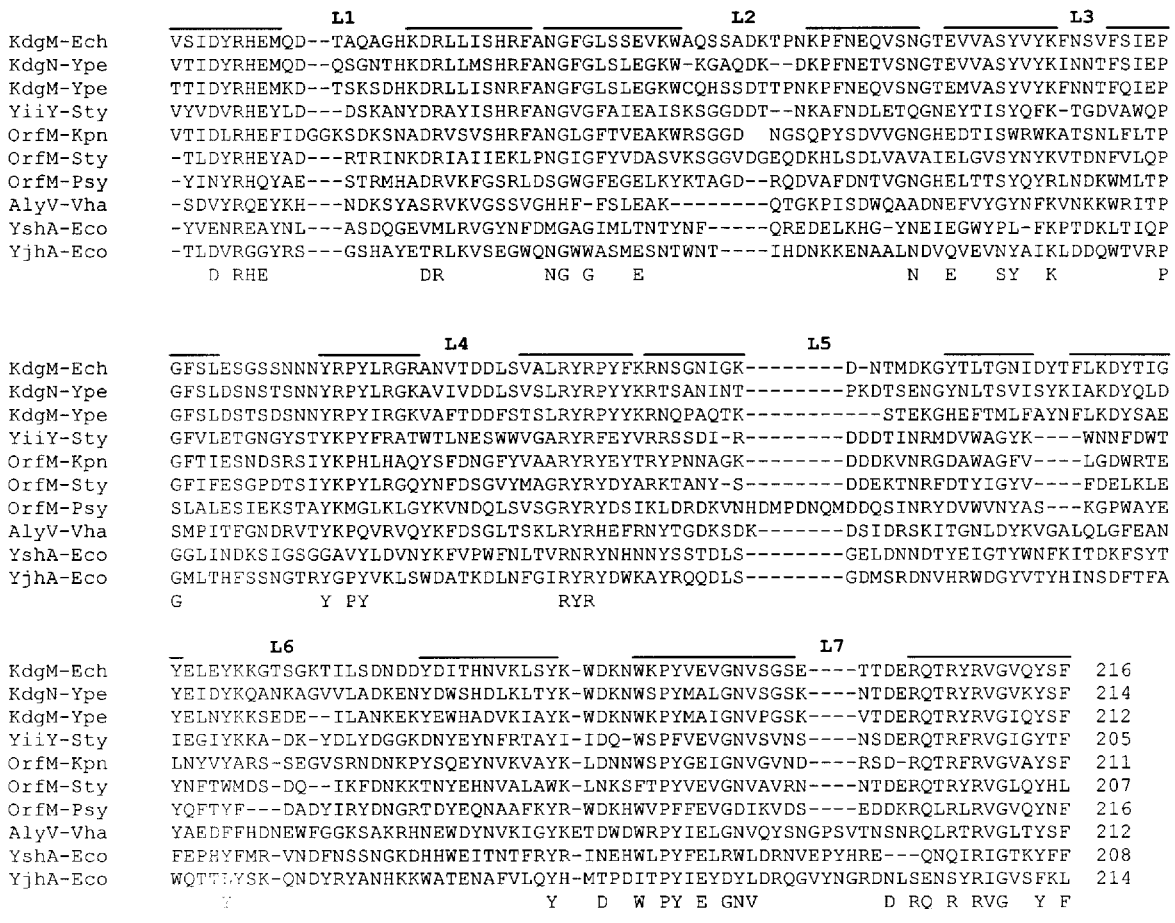

Figure 1 Alignment of the sequences of KdgM homologues

The origin and the SwissProt accession number of the proteins are: KdgM-Ech, *E. chrysanthemi* protein Q934G3; KdgN-Ype, *Yersinia pestis* protein Q8ZA25; KdgM-Ype, *Y. pestis* protein Q8ZFI5; Yiiy-Sty, *Salmonella typhimurium* protein P43022; OrfM-Kpn, *Klebsiella pneumoniae*; OrfM-Sty, *S. typhimurium*; OrfM-Psy, *Pseudomonas syringae*; AlyV-Vha, *Vibrio halioticoli* protein AlyVGIII Q9RGQ2; YshA-Eco, *E. coli* protein OmpL P76773; YjshA-Eco, *E coli* protein P39372. OrfM-Kpn, OrfM-Sty (http://genome.wustl.edu/projects/bacterial) and OrfM-Psy (http://tigrblast.tigr.org/ufmg) have been identified in unfinished genome sequencing projects. Amino acids conserved in seven or more sequences are indicated under the alignment. The position of transmembrane segments is indicated by bars over the alignment, the loops are numbered L1 to L7.

and agitated for 30 min at 20 °C. A 25 μl aliquot of Protein A–Sepharose 6MB (Amersham Biosciences, Piscataway, NJ, U.S.A.), washed twice with TBS-T, was added and the samples were incubated for 30 min at 20 °C with agitation. After two washes with 1 ml of TBS-T, 50 μl of SDS/PAGE Laemmli sample loading buffer [15] were added and the samples were boiled for 5 min.

SDS/PAGE and immunodetection

Separation of proteins by SDS/PAGE was performed according to the method of Laemmli [15]. After electrophoresis, proteins were electroblotted onto nitrocellulose, in a semidry apparatus at 2 mA/cm^2, in transfer buffer containing 40 mM glycine, 50 mM Tris, 0.4 % SDS and 10 % (v/v) methanol. The nitrocellulose was then saturated with gelatin and incubated with horseradish peroxidase-conjugated streptavidin for biotin detection or with anti-KdgM antibodies (1/10 000 dilution). The secondary antibodies (anti-IgG horseradish peroxidase-conjugated) were used at a 1/25 000 dilution. Chemiluminescence detection was performed using ECL® detection reagents (Amersham Biosciences).

RESULTS

Topology model of KdgM

To construct a working model of the topology of KdgM, several criteria were used. First, the structure of the protein was assumed to be a β-barrel as has been established for all the porins studied so far. The amino acid sequence of KdgM was aligned with those of nine homologues (Figure 1). This alignment was used to identify conserved regions, which usually correspond to the membrane-embedded domains of outer membrane proteins [16]. Putative β-strands were determined by the method of Schirmer and Cowan [17]. β-Strands with flanking aromatic residues were favoured because a ring of aromatic residues surrounds the β-barrels on both sides of the membrane in all the porins whose structure has been resolved. The first and the last residues of porins always face the periplasm. Thus, the first segment has a N_{in}-C_{out} topology, while the last segment is N_{out}-C_{in}. All this information was used to construct the model shown in Figure 2. In this model, KdgM has 14 β-strands, linked by seven loops which are, in general, short in comparison with porins with a known three-dimensional structure. Loop L3 is very short and loop L6 is the longest on the model, with 19 residues.

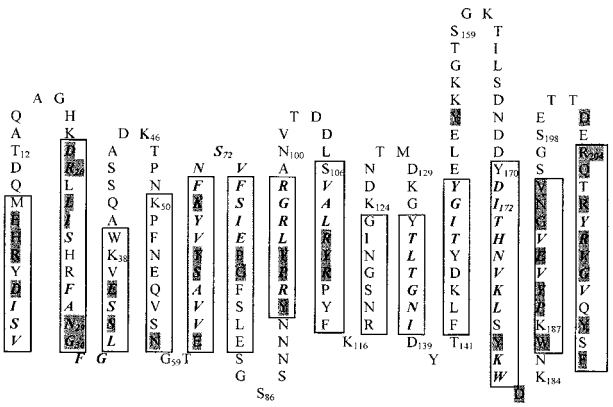

Figure 2 Putative topology model of KdgM

The transmembrane segments are represented within boxes. The 24 amino acids (single letter notation, e.g. R is arginine) substituted with cysteine are shown with sequence number. The residues present in seven or more of the members of the KdgM family aligned are shaded. The residues predicted to be in the outer membrane (17) are in bold and italic.

Construction of KdgM mutants

The KdgM protein does not contain cysteine residues. To confirm the topological model of KdgM, 24 mutants were constructed by introducing, through site-directed mutagenesis, one cysteine residue at a position that could correspond with either an outer loop or a periplasmic turn. Chemical labelling of the cysteines allowed the determination of their true cellular localization (see below). The stability of the mutant proteins was tested by immunodetection. All the proteins, except KdgM-R204C (where Arg-204 was replaced by a cysteine residue), were stable and expressed at the same level as the wild-type protein. Modification of residues in the β-strands is less permissive than in loops and turns, and often leads to instability of the porin. Arg-204 and the neighbouring residues are well conserved in all the KdgM homologues (Figure 1). This conservation and the instability of the mutant suggest that Arg-204 is in the last transmembrane segment.

The functionality of the mutant proteins was tested by their ability to restore polygalacturonate utilization to an *E. chrysanthemi kdgM* mutant. All the mutants, except KdgM-R204C, were functional. This indicates that none of the other mutated amino acids are required for the function of KdgM and that the introduction of cysteines, even at positions that are within the β-strands, does not disturb the protein architecture.

Testing of the membrane-topology model

To determine the position of the cysteines introduced by mutagenesis (outer loops, periplasmic turns or outer membrane), intact or permeabilized cells were incubated briefly with PEO-MAB, a compound that biotinylates cysteines. Since this reagent is water soluble and membrane impermeable, only cysteines present at the surface of the bacteria should be labelled in intact and permeabilized cells, whilst only cysteines exposed in the periplasm should be labelled in permeabilized cells, and the cysteines embedded in the outer membrane should not be labelled at all. After immunoprecipitation of KdgM, modification of cysteines by biotin was detected by peroxidase-conjugated streptavidin. Detection of KdgM immunoprecipitated in the intact cell samples and the permeabilized cell samples was performed in order to calibrate the results (Figure 3). The results of cysteine-labelling experiments are shown in Figure 3.

Cysteines were labelled in intact cells in KdgM-T12C, KdgM-K46C, KdgM-S72C, KdgM-N100C, KdgM-K124C, KdgM-D129C and KdgM-S198C, confirming the position of the six loops L1, L2, L3, L4, L5 and L7. Cysteine residues (denoted by the single-letter code and the position of the residue they replaced) N29C, K50C, I172C and K187C were not labelled, either in intact or in permeabilized cells. They belong to membrane β-strands. The modification of periplasmic residues in permeabilized cell samples was clearly observed for cysteine residues G59C, S86C, K116C, D139C, T141C and K184C. These residues belong to periplasmic turns.

Some other cysteine residues, such as R20C, G30C, K38C, S106C and Y170C, were weakly labelled in the permeabilized cell sample. According to the model, they should not be periplasmic but belong to β-strands. They may be facing the water-exposed central hole of the pore where they could react weakly with PEO-MAB. Thus, these residues are very probably membrane residues.

Modification of loop L6

Our model suggests that residue Ser-159 is at the top of loop L6. However, cysteine S159C was never labelled with the same intensity as other surface-exposed cysteine residues and behaved more like a periplasmic residue. Loop L6 is the only long loop of the model. It does not contain conserved residues and does not show any amphipathy. We supposed that it could fold into the barrel and play the role of the constriction loop L3 of general porins. This would reduce the accessibility of S159C to PEO-MAB and could explain the weaker labelling of S159C. Loops are more permissive to insertions and deletions than β-strands or turns [18]. To confirm that residues 151–169 form a loop, we constructed a deletion and an insertion mutant in this region and tested the stability and the functionality of the resulting proteins. We engineered protein KdgMΔ154–165 by the deletion of amino acids 154–165, and the introduction of residues E151D and E153D. Another protein, KdgMrep153–164, was engineered so that amino acids 153–164 were repeated in tandem. Although KdgMΔ154–165 was partially unstable, it restored growth of a *kdgM* mutant on polygalacturonate, indicating that it was functional. The KdgMrep153–164 insertion mutant was stable and functional. The stability and functionality of these mutants confirmed that region 151–169 is not a transmembrane region and forms a loop. The deletion of constriction loops often makes porins more permeable to antibiotics, indicating larger pore channels. We tested the sensitivity of the *E. coli* porin-less strains BL21(DE3)omp8, expressing KdgM, KdgMΔ154–165, and KdgMrep153–164 to several antibiotics and drugs (polymyxin, gentamycin, kasugamycin, phosphomycin, vancomycin, trimethoprim, erythromycin, fusidic acid, nalidixic acid, rifampicin, chloramphenicol, streptomycin, spectinomycin and SDS) but we could not detect any modification of the resistance level of these strains to these compounds. Oligomaltosides enter *E. coli* through the porin LamB. A *ΔlamB E. coli* mutant was unable to grow on maltopentaose as the sole carbon source. We tested whether KdgM or the loop L6 mutants introduced into a *ΔlamB E. coli* strain MB1 could restore growth on maltopentaose. While a liquid culture in M63 minimal medium [11] of the *lamB*$^+$ strain had a doubling time of 1.5 h with maltopentaose as the sole carbon source, no growth was observed for the strain expressing KdgM, and a slow growth was observed for the strains expressing KdgMΔ154–165 and KdgMrep153–164 (doubling time of 12 h and 22 h, respectively), indicating

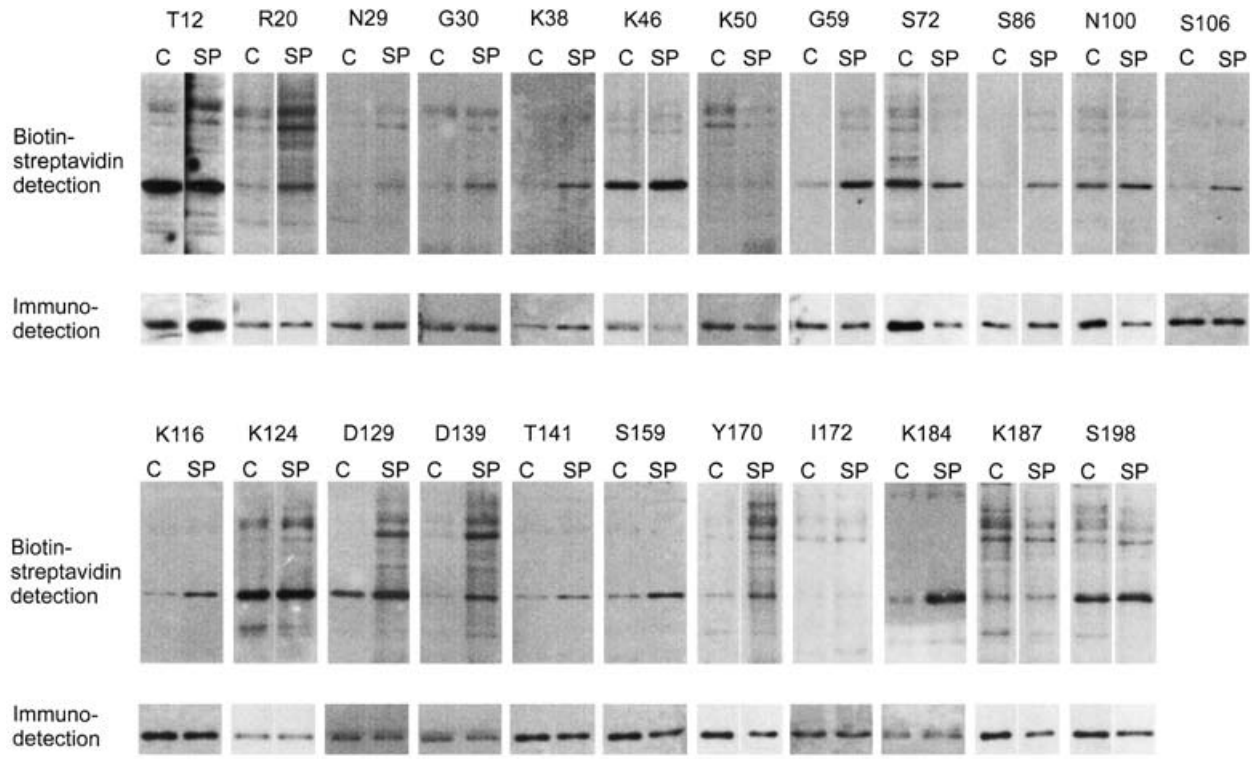

Figure 3 Labelling of cysteine residues

Intact or permeabilized cells of *E. coli* expressing mutant KdgM proteins were labelled with PEO-MAB. After immunoprecipitation of KdgM, labelling was quantified by steptavidin detection and the amount of KdgM in the samples estimated by immunodetection. C, intact cells; SP, permeabilized cells.

that the maltodextrin can enter the bacteria through the modified KdgM porins.

DISCUSSION

The strategy used for the topological analysis of KdgM was the specific labelling of cysteine residues introduced by site-directed mutagenesis. Cysteine is a relatively hydrophobic, non-bulky residue, and its insertion at most sites was well tolerated. All the mutants in the loops or the turns were stable and only one mutant out of 10 in the β-strands (KdgM-R204C) was unstable. This method is less destabilizing for the protein than the introduction of epitopes, which is sometimes used for the determination of outer-membrane protein topology: insertions into the β-strands often destabilize the protein [19,20]. Furthermore, epitopes are detected by bulky antibodies or proteases. The analysed protein may impose constraints on the localization and accessibility of the inserted sequences and impair their recognition by the antibody or the protease, leading to erroneous conclusions [20]. The small PEO-MAB molecule is more likely to be able to interact with all the surface-exposed cysteine residues. This approach should become a method of choice to study the topology of cysteine-less outer-membrane proteins.

To define the transmembrane segments of our topological model, we looked for residues conserved in KdgM homologues and for regions predicted to have amphipathic properties [17]. These two criteria are good indicators of β-strand regions. In our model, 90 % of the residues conserved in more than seven out of ten of the aligned proteins are in the transmembrane segments. Similarly almost all the regions showing amphipathy are in cross-membrane segments. The only exception is loop L3. However, this loop is very short and may consist essentially of a β-turn. Moreover, cysteine labelling experiments confirmed that Ser-72 is surface-exposed on loop L3.

The small size of KdgM compared to other porins led us to suppose that this protein may contain fewer β-strands. Our model contains 14 transmembrane segments, while the *E. coli* general porins OmpC and OmpF have 16, the *E. coli* maltoporin and sucrose porins have 18 β-strands, and the model proposed for many other porins contains 16 transmembrane segments [18,21,22]. The results obtained by electron crystallography with OmpG showed a barrel smaller than OmpF, suggesting it could also possess 14 strands, although computer modelling suggested 16 strands [10,12]. The *Fusobacterium nucleatum* FomA porin could also be formed by 14 β-strands [23]. Thus, KdgM may be a smaller barrel than the known porins but other families of porin could have the same size. The KdgM substrate is oligogalacturonates. The question is, could the pore of a 14-strand β-barrel be wide enough to allow the passage of oligosaccharides? These saccharides look similar in size and shape to maltodextrins, the substrates of LamB. The size of the LamB pore is restricted by loop L3 that folds back inside the barrel and allows the selection of the substrate. In consequence, the size of the LamB pore is smaller than that of the 16-stranded OmpF porin [3]. Moreover the diameter of a barrel is defined not only by the number of strands that constitute it but also by the angle of these strands with the axis of the molecule: for example, OmpLA, which has 12 strands, has no channel but the TolC trimer, with the same number of β-strands, has a wide barrel radius [24,25]. Thus, it is possible that the 14-stranded KdgM barrel could conduct oligogalacturonate molecules.

KdgM has only one relatively long loop, L6. In other porins loops have precise functions. OmpF loop L2 connects one subunit to its neighbour by latching into its channel [6]. Monomeric KdgM does not need such a latch. In LamB, loop L3 folds into the lumen but the channel entrance is further restricted by residues from loops L1 and L6 [3]. A smaller diameter KdgM barrel may not need loops to restrict its size. In all the porins crystallized, loop L3 was found to constrict the lumen of the pore. However, the KdgM loop L3 is very short and only loop L6 could play this role. The deletion of loop L6 allows for entry of maltopentaose into the bacteria, indicating a role of this loop in the control of the size of the lumen. It will be interesting to analyse the role played by this loop in the selection of the substrate (oligogalacturonides versus oligomaltosides).

In conclusion, the KdgM-topology study reveals that, with its smaller number of β-strands, the reduced size of the loops and a possible replacement of constriction loop L3 by loop L6, KdgM possesses a new type of topology for porins. Determination of the three-dimensional structure of the protein should confirm these results.

We thank Dr T. Schirmer for his help with topology prediction, Dr Nicole Cotte-Pattat, Dr Sylvie Reverchon, Dr Vladimir Shevchik and Dr Valerie James for reading the manuscript. This work was supported by grants from the Centre National de la Recherche Scientifique and from the ministère de l'Education et de la Recherche.

REFERENCES

1 Koebnik, R., Locher, K. P. and van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. Mol. Microbiol. **37**, 239–253
2 Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius J. N. and Rosenbusch, J. P. (1992) Crystal structures explain functional properties of two *E. coli* porins. Nature (London) **358**, 727–733
3 Schirmer, T., Keller, T. A., Wang, Y. F. and Rosenbusch, J. P. (1995) Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution. Science (Washington, D. C.) **267**, 512–514
4 Forst, D., Welte, W., Wacker, T. and Diederichs, K. (1998) Structure of the sucrose-specific porin ScrY from *Salmonella typhimurium* and its complex with sucrose Nat. Struct. Biol. **5**, 37–46
5 Van Gelder, P., Saint, N., van Boxtel, R., Rosenbusch, J. P. and Tommassen, J. (1997) Pore functioning of outer membrane protein PhoE of *Escherichia coli*: mutagenesis of the constriction loop L3. Protein Eng. **10**, 699–706
6 Phale, P. S., Philippsen, A., Kiefhaber, T., Koebnik, R., Phale, V. P., Schirmer, T. and Rosenbusch, J. P. (1998) Stability of trimeric OmpF porin: contributions of the latching loop L2. Biochemistry **37**, 15663–15670
7 Dutzler, R., Wang, Y.-F., Rizkallah, P. J., Rosenbusch, J. P. and Schirmer, T. (1995) Crystal structures of various maltooligosaccharides bound to maltoporin reveal a specific sugar translocation pathway. Structure **4**, 127–134
8 Pajatsch, M., Andersen, C., Mathe, S. A., Böck, A., Benz, R. and Engelhardt, H. (1999) Properties of a cyclodextrin-specific, unusual porin from *Klebsiella oxytoca*. J. Biol. Chem. **274**, 25159–25166
9 Fajardo, D. A., Cheung, J., Ito, C., Sugawara, E., Nikaido, H. and Misra, R. (1998) Biochemistry and regulation of a novel *Escherichia coli* K-12 porin protein, OmpG, which produces unusually large channels. J. Bacteriol. **180**, 4452–4459
10 Conlan, S., Zhang, Y., Cheley, S. and Bayley, H. (2000) Biochemical and biophysical characterization of OmpG: a monomeric porin. Biochemistry **39**, 11845–11854
11 Blot, N., Berrier, C., Hugouvieux-Cotte-Pattat, N., Ghazi, A. and Condemine, G. (2002) The oligogalacturonate-specific porin KdgM of *Erwinia chrysanthemi* belongs to a new porin family. J. Biol. Chem. **277**, 7936–7944
12 Behlau, M., Mills, D. J., Quader, H., Kühlbrandt, W. and Vonck, J. (2001) Projection structure of the monomeric porin OmpG at 6 Å resolution. J. Mol. Biol. **305**, 71–77
13 Pripilov, A., Phale, P. S., van Gelder, P., Rosenbusch, J. P., and Koebnik, R. (1998) Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from *E. coli*. FEMS Microbiol. Lett. **163**, 65–72
14 Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
15 Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature (London) **227**, 680–685
16 Simonet, V., Malléa, M., Fourel, D., Bolla, J.-M. and Pagès, J.-M. (1996) Crucial domains are conserved in enterobacteriaceae porins. FEMS Microbiol. Lett. **136**, 91–97
17 Schirmer, T. and Cowan, S. W. (1993) Prediction of membrane-spanning beta-strands and its application to maltoporin. Protein Sci. **2**, 1361–1363
18 Bina, J., Bains, M., Hancock, R. E. W. (2000) Functional expression in *Escherichia coli* and membrane topology of porin HopE, a member of a large family of conserved proteins in *Helicobacter pylori*. J. Bacteriol. **182**, 2370–2375
19 Wong, K. K. Y. and Hancock, R. E. W. (2000) Insertion mutagenesis and membrane topology model of the *Pseudomonas aeruginosa* outer membrane protein OprM. J. Bacteriol. **182**, 2402–2410
20 Guédin, S., Willery, E., Tommassen, J., Fort, E., Drobecq, H., Locht, C., and Jacob-Dubuisson, F. (2000) Novel topological features of FhaC, the outer membrane transporter involved in the secretion of the *Bordetella pertussis* filamentous hemagglutinin. J. Biol. Chem. **275**, 30202–30210
21 Huang, H., Jeanteur, D., Pattus, F. and Hancock, R. E. W. (1995) Membrane topology and site-specific mutagenesis of *Pseudomonas aeruginosa* porin OprD. Mol. Microbiol. **16**, 931–941
22 Mobasheri, H., Ficht, T. A., Marquis, H., Lea, E. J. A. and Lakey, J. H. (1997) Brucella Omp2a and Omp2b porins: single channel measurements and topology prediction. FEMS Microbiol. Let. **155**, 23–30
23 Puntervoll, P., Ruud, M., Bruseth, L. J., Kleivdal, H., Hogh, B. T., Benz, R. and Jensen, H. B. (2002) Structural characterization of the fusobacterial non-specific porin FomA suggests a 14-stranded topology, unlike the classical porins. Microbiology **148**, 3395–3403
24 Snijder, H. J., Ubarretxena-Beladia, I., Blaauw, M., Kalk, K. H., Verhij, H. M., Egmont, M. R., Dekker, N. and Dijkstra, B. W. (1999) Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. Nature (London) **405**, 914–919
25 Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C. (2000) Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. Nature (London) **405**, 914–919

Received 3 January 2003/19 February 2003; accepted 25 February 2003
Published as BJ Immediate Publication 25 February 2003, DOI 10.1042/BJ20030027