Dynamic Membrane Topology of the *Escherichia coli* β-Glucoside Transporter BglF

Received for publication, September 22, 2004, and in revised form, March 7, 2005
Published, JBC Papers in Press, March 8, 2005, DOI 10.1074/jbc.M410896200

Sharon Yagur-Kroll and Orna Amster-Choder‡

From the Department of Molecular Biology, The Hebrew University-Hadassah Medical School, P. O. Box 12272, Jerusalem 91120, Israel

The *Escherichia coli* BglF protein, a permease of the phosphoenolpyruvate-dependent phosphotransferase system, catalyzes transport and phosphorylation of β-glucosides. In addition, BglF regulates bgl operon expression by controlling the activity of the transcriptional regulator BglG via reversible phosphorylation. BglF is composed of three domains; one is hydrophobic, which presumably forms the sugar translocation channel. We studied the topology of this domain by Cys-replacement mutagenesis and chemical modification by thiol reagents. Most Cys substitutions were well tolerated, as demonstrated by the ability of the mutant proteins to catalyze BglF activities. Our results suggest that the membrane domain contains eight transmembrane helices and an alleged cytoplasmic loop that contains two additional helices. The latter region forms a dynamic structure, as evidenced by the alternation of residues near its ends between face-in and face-out states. We suggest that this region, together with the two transmembrane helices encompassing it, forms the sugar translocation channel. BglF periplasmic loops are close to the membrane, the first being a reentrant loop. This is the first systematic topological study carried out with an intact phosphotransferase system permease and the first demonstration of a reentrant loop in this group of proteins.

The *Escherichia coli* BglF protein (EII*β*gl), an enzyme II of the phosphoenolpyruvate-dependent carbohydrate phosphotransferase system (PTS), catalyzes concomitant transport and phosphorylation of β-glucosides across the cytoplasmic membrane (1). In addition to its ability to phosphorylate its sugar substrate, BglF phosphorylates the transcriptional regulator BglG in the absence of β-glucosides and dephosphorylates P-BglG upon addition of β-glucosides to the growth medium (2, 3). By controlling the phosphorylation state of BglG, BglF controls the dimeric state of BglG and, thus, its ability to bind RNA and antiterminate transcription of the bgl operon (4).

BglF dimerizes spontaneously in the membrane, and its dimeric form can catalyze all the above mentioned activities (5). Several other enzymes II of PTS were shown to regulate activity of transcription factors. Two representative examples are SacX, a sucrose permease from *Bacillus subtilis*, which regulates the activity of the BglG homologue SacY by reversible phosphorylation (6, 7), and PstG, a glucose permease from *E. coli*, which regulates the activity of the global regulator Mic by membrane sequestration (8, 9).

The phosphate flux in PTS starts with a phosphoryl group, donated by phosphoenolpyruvate, which is passed through the general PTS enzymes, enzyme I, and HPr, to the various sugar-specific permeases. Like many other PTS permeases, BglF is composed of three domains: the hydrophilic A and B domains (IIA*β*gl and IIB*β*gl) and the hydrophobic C domain (IIC*β*gl) (reviewed in Refs. 10–12). The domains of BglF are covalently linked to one another in the order BCA. The A domain is phosphorylated by HPr; the phosphate is then transferred to the B domain and subsequently to β-glucosides; the membrane-spanning C domain presumably forms the sugar translocation channel and at least part of the sugar-binding site. Cys-24, residing in the B domain, was shown to be responsible for the phosphorylation of both the incoming sugar and the BglG protein and for the dephosphorylation of P-BglG (13, 14). The mechanism that ensures correct delivery of the phosphoryl group to the right entity, sugar or protein, by the same active site is not known. A clue to this mechanism seems to lie in the different domain requirements of the distinct functions; whereas the B domain is sufficient for BglG phosphorylation, both the B and C domains, connected in the order BC, are required for the sugar-stimulated functions, i.e. sugar phosphotransfer and P-BglG dephosphorylation (15, 16). Catalysis of the sugar-stimulated functions involves specific interaction(s) between the active site-containing B domain and the integral membrane C domain (17).

To understand better how BglF catalyzes its different activities, we investigated its membrane topology. To examine the way BglF traverses through the membrane, we constructed a set of single-Cys BglF variants by site-directed mutagenesis, and we assessed the exposure and the sidedness of the planted cysteines, in whole cells and in membrane vesicles with an inside-out orientation, by chemical modification with thiol-specific reagents. Of the 36 constructed single-Cys mutants, only one was not expressed. Because the BglF-active site is a cysteine residue (Cys-24), single-Cys mutants are not active. Hence, we constructed a parallel set of di-Cys mutants, each containing Cys-24 in the B domain and a single cysteine in the membrane C domain. We used the di-Cys mutants to test the effect of the Cys replacements on BglF activities. All mutants catalyzed phosphorylation and dephosphorylation of BglG, and except for one, they were all capable of β-glucosides phosphotransfer, implying that they retained proper folding. The topol-
Dynamic Membrane Topology of BglF

...transport and phosphorylate these sugars, which were then cleaved by the products of the constitutively expressed chromosomal bglA gene and the plasmid-encoded bglB gene, respectively. Hence, the formation of red colonies on MacConkey-agar plates and MacConkey-sal cin plates was indicative of the ability of the di-Cys BglF mutants to catalyze β-glucosidases phosphotransfer.

To estimate quantitatively the ability of the BglF variants to phosphotransfer β-glucosides, the assay described by Schaeffer (21) was carried out with minor modifications. HMS174(DE3) cells expressing the di-Cys BglF mutants were grown in M63 minimal medium, supplemented with 0.4% succinate as a carbon source, at 37 °C. Whole cells reached an A600 of 0.4–0.5, isopropyl β-D-galactopyranoside was added to a final concentration of 1 mM. After 30 min, cultures were put on ice for 20 min, and their absorbance was measured (A600). Cells (1 ml) were pelleted by centrifugation, washed, and resuspended in 0.9 ml of 0.075 M phosphate buffer, pH 7.5, containing 1 mM MgSO4. The cells were scratched into the buffer and incubated for 30 min at 37 °C. The reaction was stopped by the addition of 0.5 ml of 2 M Na2CO3. Cells were pelleted by centrifugation, and the amount of p-nitrophenol was estimated by measuring absorbance at 410 nm. The units of enzyme activity were calculated using the following equation: units = [A600 × 1000]/(A600 × 1 ml × 10 min).

Cysteine Accessibility to Thiol Reagents in Whole Cells—Cysteine accessibility to thiol reagents was assayed essentially as described by Ninio et al. (22) with some modifications. HMS174(DE3) cells expressing single- and di-Cys BglF variants were grown in LB medium supplemented with 0.1 mg/ml ampicillin at 30 °C. When cells reached an A600 of 0.3, expression of the BglF variants was induced by the addition of isopropyl β-D-galactopyranoside to a final concentration of 0.1 mM. After 5 h, 4 aliquots of 1 ml of culture were pelleted and resuspended in 200 μl of a final concentration of 20 μM Tris-NaCl buffer (150 mM NaCl, 15 mM Tris, pH 7.5), pelleted again, and resuspended in 200 μl of Tris-NaCl buffer. Freshly made MTS, MTSET, or MTSES, at final concentrations of 1, 1, and 2.5 mM, respectively, were added, each to 1 aliquot. The fourth aliquot served as a control and hence was not treated with MTS reagents. After 10 min of incubation at either 25 or 4 °C, as indicated in the text, cells were pelleted, washed with Tris-NaCl buffer and pelleted again. The cells were lysed by resuspending them in 100 μl of 30 mM Tris-HCl, pH 8.0, 30% sucrose, 10 mM EDTA, pH 8.0, and 50 μg/ml lysozyme and incubating them for 15 min at 37 °C. After the addition of 600 μl 15 mM MgSO4 containing 5 μg/ml DNase, incubation was continued for 15 min at 37 °C. Membranes were pelleted by centrifugation (14,000 rpm for 15 min at 4 °C) and resuspended in 100 μl of 0.5 M sucrose-20 μM Tris-HCl, pH 7.5, 10 mM MgSO4 containing 1 mg/ml lysozyme. Membranes were solubilized by the addition of 0.5 ml of 0.5% Triton X-100 to a final concentration of 0.8% and gentle rotation of the tubes for 15 min at room temperature. The BglF variant proteins were purified to near-homogeneity by using nickel-chelate chromatography as follows. To each aliquot, 300 μl of Tris-NaCl buffer containing 30 mM imidazole and 50 μl of nickel-nitritolactric acid beads (previously washed three times with DM and 30 mM imidazole) were added, and the tubes were rotated for 1 h. Subsequently, the beads were washed with denaturation buffer (Tris-NaCl buffer containing 6 M urea and 0.5% SDS), pelleted, and resuspended in 100 μl of denaturation buffer containing 0.25 mM fluor- escin-5-maleimide (5-FM, dissolved in Me2SO). Samples were incubated for 30 min at 37 °C. The protein content in each lane was evaluated by staining the same gel with Coomassie Brilliant Blue. The Coomassie-stained profile was recorded with ImageMaster VDS-CL imaging system (white light, no filter). The amounts of the fluorescent BglF bands were normalized to the amounts of the stained bands using the TINA 2.0 software. The percentage of labeling of each mutant protein with 5-FM after treatment with MTS reagents was calculated from the ratio of the intensity of the band of interest to the total intensity of the band. The results obtained with the untreated samples ruled out the possibility that free cysteines are oxidized in this procedure after cell lysis.

Preparation of Membrane Vesicles with an Inside-out Orientation and Cysteine Accessibility Analyses—HMS174(DE3) cells expressing single-
Cys BglF variants were grown and induced as described above, and inside-out (ISO) membrane vesicles were prepared essentially as described before (29). Cells were washed with 50 mm Tris-HCl, pH 7.5, 5 mm MgCl₂, and 0.2 mM 4-(2-aminoethyl)benzene-sulfonyl fluoride and resuspended in 10 mM Tris-HCl, pH 7.5, 140 mM choline chloride, 0.25% sucrose, 5 mM MgSO₄, and 5 μg/ml DNase. The suspension was passed once through a French press cell operated at 10,000 pounds/square inch. Unbroken cells were removed by centrifugation at 5,000-10,000 × g for 15 min at 4 °C, and membranes were collected from the supernatant by centrifugation at 100,000 × g for 1 h at 4 °C. The membranes were frozen in liquid nitrogen and stored at −80 °C in 50 mM potassium phosphate, pH 7.4, 5 mM MgSO₄, and 10% glycerol at −20 °C. The membranes were then resuspended in 50 mM potassium phosphate, pH 7.4, 5 mM MgSO₄, and 10% glycerol at −20 °C.

Membrane Topology Prediction Methods—Five topology prediction methods, TMHMM (24), HMMTOP (25), MEMSAT (26), TOPPRED (27), and PHD (28), were used in their single-sequence mode. All adjustable parameters were left at their default values. All these methods are available on-line.

RESULTS

Prediction of Transmembrane Segments and Construction of Single- and Di-cysteine BglF Mutants—We used five topology prediction algorithms, HMMPHMM, TMHMM, TOPRED, MEMSAT, and PHD (see "Experimental Procedures"), to predict the membrane topology of BglF (Fig. 1, first five lines). The extent of agreement between these frequently used prediction methods, provided by a simple "majority vote" approach, was shown to be a good indication for the reliability of the predicted topology of E. coli inner membrane proteins is close to one when four or more of these methods agree (29). We made use of these predictions to choose residues, which were predicted with a high degree of certainty to face either the cytoplasm or the periplasm, and others, on which there is no consensus between the predictions, and we mutated them to cysteines in a background of a Cys-less BglF derivative. In addition, we constructed variants that contain each of the native cysteines in BglF as a single cysteine, including one in the B domain and one in the linker connecting the C and A domains, which are presumed to be inside the cell. This collection of these single-Cys BglF variants was expected to provide useful tools for establishing the operational parameters for our experimental approach and for investigating the actual membrane topology of BglF.

All together, 36 single-Cys BglF derivatives (with single cysteines between positions 97 and 480) were engineered as histidine-tagged proteins. The expression of only one mutant, T130C, was not detected after affinity purification or metabolic labeling with [³⁵S]methionine (not shown). All other 35 single-Cys mutants were expressed in levels comparable with wild-type BglF, with the exception of variants with single-Cys at positions 141, 297, 313, 433 and 434, which were produced at slightly reduced levels (50–70% of wild-type level) (not shown). All single-Cys BglF variants were found in the membrane fraction after cell fractionation (not shown). The collection of single-Cys BglF mutants, whose expression was detected, was used to study the topology of the membrane C domain by cysteine accessibility analyses, as described below.

Because the BglF active site, which phosphorylates the incoming β-glucosides and the BglG protein and also accepts the phosphate from P-BglG, is a cysteine residue (Cys-24), single-Cys BglF mutants are not active. Hence, we constructed a parallel set of di-Cys mutants, each containing Cys-24, in the B domain, plus one of the 35 cysteines described above to test the effect of the cysteine substitution in the C domain on BglF activity. The pattern of expression of the di-Cys variants was the same as that of the single-Cys mutants (not shown). The protocol of chemical modification by thiol-specific reagents was repeated with the di-Cys mutants. Most significantly, the
results obtained with the two sets of mutants led to the same conclusions, as detailed below, indicating that a replacement of Cys-24 in the B domain by a serine does not affect the arrangement of the BglF derivatives in the membrane.

**Catalytic Activities of the Cys-replacement BglF Mutants**—We assayed the ability of the di-Cys BglF mutants, containing the active site cysteine in the hydrophilic B domain, and a single cysteine in the C domain, to catalyze the various activities of BglF, i.e. sugar phosphotransfer and BglG phosphorylation and dephosphorylation. As a preliminary qualitative assay for sugar phosphotransfer, cells expressing the different di-Cys BglF variants were grown on MacConkey indicator plates containing the β-glucoside salicin or arbutin. Cells expressing functional BglF proteins catalyze uptake of salicin and arbutin; subsequent metabolism of these sugars leads to acidification of the medium and the appearance of red colonies. Cells expressing inactive BglF mutants form white colonies on MacConkey indicator plates, whereas mutants with an intermediate activity grow as pink colonies. Of the 35 di-Cys mutants, only one, in position 432, formed white colonies on the indicator plates, and five others gave pink colonies (Table II). All other mutants grew as red colonies, indistinguishable from cells expressing wild-type BglF or the variant containing Cys-24 as a single cysteine (Cys-24 only).

Hence, we concluded that all the di-Cys mutants, excluding one, retained at least some ability to transport and phosphorylate β-glucosides.

To estimate quantitatively the ability of the di-Cys BglF mutants to catalyze sugar phosphotransfer, we tested the ability of cells expressing them to transport and phosphorylate the pNPG. Hydrolysis of P-pNPG generates a yellow color, whose intensity can be quantitated by measuring the absorbance at 410 nm (see “Experimental Procedures”). The results are presented in Table II. Cells expressing the V432C variant, which formed white colonies on MacConkey salicin and MacConkey arbutin plates, were completely inactive in the pNPG assay. Of the five Cys replacements that led to growth of pink colonies on MacConkey salicin/arbutin plates, four exhibited partial activity, ranging from 42 to 74% activity, compared with wild-type BglF and to Cys-24 only, and the fifth demonstrated normal activity (90% activity). Notably, one of the four is a replacement of Ile-433 (67% activity), which is adjacent to Val-432, whose replacement generated an inactive permease. Another residue, Thr-368, whose replacement with cysteine decreased the phosphotransferase activity of BglF (42% activity), is the third residue in the GLYE motif, which is conserved in PTS permeases and suggested to be involved in sugar transport and phosphorylation (see “Discussion”). All other di-Cys BglF derivatives demonstrated a very high activity (79–125% activity).

Our results indicate that the capacity of most di-Cys mutants to transport and phosphorylate β-glucosides is near or equal to that of the wild-type protein.

To test the ability of the di-Cys BglF derivatives to reversibly phosphorylate BglG, we expressed them in strain MA200-1. This strain carries a defective bglF gene and a chromosomal fusion of the bglF promoter and terminator to the lacZ gene (19). Expression of the lacZ gene depends on transcriptional antitermination by BglG. Because of the mutation in the bglF gene, BglG is not negatively regulated by phosphorylation in MA200-1, and hence, lacZ is expressed constitutively leading to growth of red colonies on MacConkey lactose plates. Introduction of a plasmid encoding wild-type BglF into MA200-1 results in inductive expression of lacZ. This is manifested by the growth of white colonies on MacConkey lactose plates in the absence of β-glucosides, due to phosphorylation of BglG by the plasmid-encoded BglF, and the formation of red colonies in the presence of β-glucosides, due to dephosphorylation of BglG by BglF. As can be seen in Table II, all the di-Cys BglF derivatives gave the same phenotype as wild-type BglF, i.e. the mutants act as BglG kinases in the absence of β-glucosides and as BglG phosphatases in the presence of the sugar.

Taken together, all BglF mutants are active as BglG kinases and phosphatases, and except for V432C, they all catalyze sugar phosphotransfer. The fact that most mutations were well tolerated indicates that replacement of the native amino acids at these positions by cysteines does not cause structural perturbations. The discrepancy between the negative effect of few Cys substitutions on sugar phosphotransfer and the lack of effect of these substitutions on BglG reversible phosphorylation suggests that these substitutions affect sites involved in β-glucosides recognition or binding rather than BglF proper folding.

### Table I

**Prediction of transmembrane helices in BglF**

The numbers correspond to amino acid positions in the BglF protein.

| HMMTOP | TMHMM | PHDhtm | TOPRED | MEMSAT | This work |
|--------|-------|--------|--------|--------|----------|
| 102–121 | 100–119 | 108–125 | 100–120 | 100–124 | 100–117 |
| 142–164 | 139–161 | 144–161 | 140–160 | 140–164 | 144–161 |
| 171–190 | 168–190 | 170–190 | 172–190 | 173–190 |
| 211–230 | 205–227 | 209–227 | 211–231 | 203–227 | 208–230 |
| 245–264 | 234–256 | 246–265 | 245–265 | 244–264 | 246–263 |
| 273–292 | 260–282 | 286–304 | 287–307 | 273–293 | 274–296 |
| 299–320 | 287–309 | 300–317 | 300–317 |
| 325–346 | 324–346 | 325–345 | 324–345 |
| 355–374 | 353–373 | 355–375 |
| 383–400 | 381–400 | 385–402 | 381–401 | 383–399 |
| 407–424 | 384–399 | 404–424 | 406–422 | 413–430 |
| 429–448 | 426–448 | 427–446 | 429–449 | 429–449 | 437–454 |

No. of TMs 12 10 7 11 12 8
The activity of di-Cys BglF variants as sugar phosphotransferases and as BglG-negative regulators

The background for all di-Cys BglF proteins was a BglF variant containing the active site cysteine as a sole cysteine (Cys-24 only).

| BglF mutants | Phenotype on MacConkey-arbutin plates | Phenotype on MacConkey-salicin plates | % Activity in pNPG assay | Phenotype on MacConkey-acetate plates |
|--------------|--------------------------------------|--------------------------------------|---------------------------|-------------------------------------|
| Wild type    | Red                                  | Red                                  | 100                       | White                               |
| Vector only  | White                                | White                                | 1                         | Red                                 |
| C24S        | White                                | White                                | 2                         | Red                                 |
| Cys-less     | White                                | White                                | 0                         | Red                                 |
| Cys-24 only  | Red                                  | Red                                  | 101                       | Red                                 |
| L125C + Cys-24 | Red                               | Red                                  | 95                       | White                               |
| L125C + Cys-24 | Red                               | Red                                  | 87                       | White                               |
| E136C + Cys-24 | Red                               | Red                                  | 93                       | White                               |
| S138C + Cys-24 | Red                               | Red                                  | 96                       | White                               |
| Y141C + Cys-24 | Red                               | Red                                  | 79                       | White                               |
| A163C + Cys-24 | Red                               | Red                                  | 90                       | White                               |
| N192S + Cys-24 | Red                               | Red                                  | 125                      | White                               |
| A196C + Cys-24 | Red                               | Red                                  | 101                      | White                               |
| A198C + Cys-24 | Red                               | Red                                  | 100                      | White                               |
| L201C + Cys-24 | Red                               | Red                                  | 94                       | White                               |
| Cys-227 + Cys-24 | Red                               | Red                                  | 93                       | White                               |
| S240C + Cys-24 | Red                               | Red                                  | 86                       | White                               |
| Cys-24 + S240C | Red                               | Red                                  | 90                       | White                               |
| Cys-24 + S271C | Pink                               | Pink                                 | 90                       | White                               |
| A291C + Cys-24 | Red                               | Red                                  | 97                       | White                               |
| W297C + Cys-24 | Pink                               | Pink                                 | 74                       | White                               |
| Cys-31S + Cys-24 | Red                               | Red                                  | 85                       | White                               |
| L320C + Cys-24 | Red                                  | Red                                  | 103                      | White                               |
| Cys-346 + Cys-24 | Red                               | Red                                  | 93                       | White                               |
| A350C + Cys-24 | Red                                  | Red                                  | 90                       | White                               |
| L364C + Cys-24 | Red                                  | Red                                  | 99                       | White                               |
| T388C + Cys-24 | Pink                               | Pink                                 | 42                       | White                               |
| Cys-387 + Cys-24 | Red                               | Red                                  | 102                      | White                               |
| A400C + Cys-24 | Red                                  | Red                                  | 92                       | White                               |
| S406C + Cys-24 | Red                                  | Red                                  | 98                       | White                               |
| L409C + Cys-24 | Pink                               | Pink                                 | 74                       | White                               |
| S411C + Cys-24 | Red                                  | Red                                  | 89                       | White                               |
| V432C + Cys-24 | White                              | White                                | 1                        | Red                                 |
| I433C + Cys-24 | Pink                               | Pink                                 | 67                       | White                               |
| G434C + Cys-24 | Red                                  | Red                                  | 97                       | White                               |
| Cys-441 + Cys-24 | Red                               | Red                                  | 92                       | White                               |
| Cys-480 + Cys-24 | Red                               | Red                                  | 94                       | White                               |

* Sugar phosphotransfer was studied in strain HMS174(DE3), which carries a cryptic bgl operon on its chromosome.
* The ability of di-Cys BglF variants to negatively regulate BglG activity, depending on β-glucoside addition to the growth medium, was assayed in MA200-1, which carries a bgl'-lacZ transcriptional fusion on its chromosome and a defective bglF gene.
* 0.5% arbutin or 0.4% salicin were added to the growth medium when indicated.
* To assay salicin phosphotransfer, pZGM-B encoding the bglB gene, whose product hydrolyzes phosphorylated salicin, was introduced into the cells.
* Activity was calculated as following: (% Activity = Activity / Activity [control] × 100). Wild type activity was defined as 100%. The values represent the average of at least three independent measurements. Standard deviation ranged from 0 to 16% activity.

which fluorescent labeling of the cysteine residue is blocked by prior incubation of the cells with the impermeable but nonfluorescent MTSET and MTSES reagents are expected to locate extracellularly. Labeling of intracellular-facing residues should not be affected by this maneuver. In contrast, fluorescent labeling of cysteine in both the intracellular and extracellular milieu is expected to be inhibited after incubating the cells with the permeable, nonfluorescent MTSEA reagent. Fluorescent labeling at positions where the cysteine is either in a nonaqueous environment or is engaged in a disulfide bond is not expected to be blocked by any of the MTS reagents.

To establish the operational parameters that are suitable to distinguish between cytoplasm- and periplasm-facing residues in our experimental system, we first studied the accessibility of three positions, whose subcellular localization is predicted with a high probability (see Fig. 1), i.e. Cys-24 and Leu-97, both expected to be cytoplasmic, the first being the active site that resides in the hydrophilic B domain and the second linking the B and C domains that precedes the first predicted TM, and Leu-125, which is predicted to locate to the first periplasmic loop. The results are presented in Fig. 2. Incubation of cells expressing BglF variants with single cysteines in positions 24 and 97 at 25 °C for 10 min with the permeable, nonfluorescent reagents MTSET or MTSES at a final concentration of 1 mM did not result in inhibition of 5-FM labeling, i.e. the intensity of the fluorescent labeling of these proteins is comparable with that obtained when these cells were not pretreated with an MTS reagent (Fig. 2A, compare lanes 2 and 3 with lane 1 and lanes 6 and 7 with lane 5). Incubation of cells expressing these two BglF variants with 2.5 mM of the permeable, nonfluorescent MTSEA reagent resulted in complete inhibition of the fluorescent labeling with 5-FM (Fig. 2A, lanes 4 and 8). Different results were obtained with the single-Cys L125C variant. Incubation with all three MTS reagents resulted in inhibition of 5-FM labeling of this variant (Fig. 2A, compare lanes 14–16 with lane 13). Hence, cysteines planted in these three positions exhibited the accessibility patterns expected from their predicted location, i.e. a cytoplasmic
location for Cys-24 and Leu-97 and a periplasmic location for Leu-125. These results clearly demonstrate that under these conditions the membrane-impermeant MTS reagents do not modify intracellular residues, when applied to the extracellular side of whole cells, whereas the MTSEA compound, which is known to penetrate the membrane, readily reacts with these residues. The Cys-less BglF variant was not labeled in our experiments (Fig. 2A, lane 21). Increasing the concentration of the impermeable MTSET reagent did not affect the results (Fig. 2B). Increasing the concentration of 5-FM did not affect the nature of the results, nor did it increase the intensity of the fluorescent bands (Fig. 2C). Hence, the lowest concentrations of the thiol reagents, which are sufficient for determining membrane orientation, were chosen for subsequent analyses.

Next, we compared the accessibility of the cysteines in the di-Cys mutants, L97C+Cys-24 and L125C+Cys-24, which contain the active site cysteine in the B domain in addition to the Cys replacement in the C domain, with that of the respective single-Cys variants L97C and L125C. Labeling of the faced-in cysteine in position 97 was not affected by the presence of Cys-24, i.e., L97C+Cys-24, like L97C, was not blocked by prior treatment with the permeable MTSEA reagent (Fig. 2A, compare lanes 9–12 with lanes 5–8). As for L125C+Cys-24, it demonstrated the same bias in its reaction with the MTS reagents as L125C, although the inhibition of 5-FM labeling caused by the impermeable MTS reagents was to a lesser extent (Fig. 2A, compare lanes 17–20 with lanes 13–16). The lower inhibition of labeling of L125C+Cys-24 compared with L125C was expected from the presence of an internal cysteine in position 24, which is inaccessible to the impermeable reagents, in addition to faced-out cysteine in position 125 in the di-Cys variant (Fig. 2A, lanes 2 and 3). These results suggested that the presence of Cys-24 in the hydrophilic B domain allows for the assessment of membrane orientation of cysteines in the C domain, whereas the absence of Cys-24 does not affect the membrane topology of the C domain. Henceforth, we studied the exposure of various Cys replacements with both sets of mutants, the single- and the di-Cys, and the results with the two collections led to the same conclusions.

Using the experimental conditions established for the Cys-24, L97C, and L125C variants, the accessibility of all positions, which were replaced by cysteines, was assessed. The percentage of inhibition of 5-FM labeling by the MTS reagents was estimated as described under “Experimental Procedures.” The analysis allowed us to distinguish between four groups of Cys replacements, presented in Fig. 3, A–D, which demonstrated the following characteristics. For group A, Cys residues at positions 227, 251, 291, 296, 387, and 441 were inaccessible to both the permeable MTSEA reagent and the impermeable MTSET reagent, as indicated by the lack of inhibition of 5-FM labeling following prior treatment with both reagents. The unreactive nature of these residues indicates that they are not exposed to an aqueous environment, suggesting that they are buried either inside the membrane core or in an otherwise apolar environment. All positions in the other three groups were accessible to the permeable MTSEA reagent, as indicated by the complete inhibition (90–100%) of 5-FM labeling of cysteines planted in these positions after treating the cells with MTSEA. Therefore, only the results with and without prior treatment by the impermeable MTSET reagent are shown in Fig. 3, B–D. For group B, Cys residues at positions 141, 201, and 271 behaved like the cysteine in position 125 and were accessible to both the permeable and impermeable MTS reagents, respectively, as indicated by the efficient inhibition of 5-FM labeling following prior treatment with these reagents. This pattern of behavior suggests that these positions are exposed to solvent from the periplasmic side of the membrane. For group C, Cys residues at positions 136, 138, 163, 192, 196, 240, 346, 350, 364, 368, 377, and 480 behaved like the cysteines in positions 24 and 97, i.e., they were accessible only to the permeant but not to the impermeant MTS compounds. Hence, these cysteines are solvent-exposed but are not freely accessible from the periplasm. For group D, Cys residues at positions 198, 297, 313, 320, 400, 406, 409, 411, 433, and 434 were accessible to the permeant MTSEA reagent but demonstrated intermediate accessibility to the impermeant MTSET reagent. Whereas prior treatment by MTSET led to 77% and 8% average inhibition of 5-FM labeling at group B and group C positions, respectively, positions in group D demonstrated 20–40% inhibition (297, 313, and 411 were in the lower range and the rest in the higher range; the extent of inhibition was estimated as described under “Experimental Procedures”). Hence, group D positions are slightly accessible from the periplasm, but their accessibility mode resembles more the pattern observed with the internal group C than with the external group B positions, although it is more equivocal. Moreover, whereas the results with most variants were reproducible (5–10% deviations), the results with several group D positions (320, 406, 409, 433, and 434) were not very consistent (43%). Levels of inhibition by MTSET were observed in different experiments (the accessibility of each of the 34 positions was tested 4–12 times, with positive and negative controls included in each experiment). Preincubation of cells expressing the Cys replacement mutants with MTSES yielded very similar results to those obtained with MTSET, excluding two group D positions, 313 and 320, whose labeling was efficiently blocked by prior treatment with MTSES (not shown) but only partially blocked by MTSET (Fig. 3D). In no case did the results obtained with the di-Cys derivatives lead to different conclusions than those deduced from the experiments with the single-Cys replacements. However, because the distinction between mutants from groups C and D is less straightforward with the di-Cys variants and because of the additional decrease in MT-
SET inhibition caused by the presence of Cys-24 (see above), results obtained with single-Cys mutants are shown.

To examine the possibility that inaccessibility of any of the group A variants to MTSEA was because of the formation of a disulfide bond between neighboring molecules, we used the reducing agent tris(2-carboxyethyl)phosphine (TCEP) to reduce putative disulfides. In contrast to the negligible reactivity of TCEP with maleimide compounds, which allows for their simultaneous addition (31, 32), we found that TCEP reacts stoichiometrically with the MTSEA reagents, and their concurrent addition in equimolar amounts leads to quenching of the MTSEA compounds (not shown). Hence, we incubated the cells with 2.5 mM TCEP for 5 min, washed them, added twice the molarity of TCEP with maleimide compounds, and continued with our regular protocol described above. Labeling with 5-FM of all six group A mutants was not inhibited by this procedure (data not shown). Therefore, the inaccessibility of group A positions to MTSEA is not due to the engagement of the cysteines in disulfide bonds but is probably due to their burial in a hydrophobic, nonaqueous environment.

**Accessibility of Positions with Intermediate Access to MTSET at 4 °C**—Because of the inconsistency of the results with several group D Cys replacements, we repeated the protocol for the estimation of cysteine accessibility under conditions that minimize backbone thermal protein motion (33, 34). To this end, whole cells expressing the BglF variants were preincubated in the absence or presence of the membrane-impermeable reagent MTSET at 4 °C. Following cell lysis and affinity purification of the protein or some membrane leakiness permitted at or facilitated by these regions that is enhanced by thermal motion (see “Discussion”). Because the organism is likely to proliferate and consume carbohydrates at temperatures closer to 25 than to 4 °C, the dynamics of these residues at 25 °C might be more relevant to its existence in nature.

**Effect of Ligand on Cys Accessibility**—To explore further the role of the group D residues in BglF conformation and function, the effect of the β-glucoside salicin on the accessibility of cysteines in those positions to MTSET at 25 °C was tested. To this end, whole cells containing single-Cys BglF derivatives were incubated with MTSET posterior to incubation of the cells with or without salicin. Following labeling of the purified and denatured BglF proteins with 5-FM, fluorescence was detected as described above. As group D positions show intermediate and inconsistent accessibility to MTSET, we considered only dramatic changes in the accessibility to the reagent. Such a change was manifested by L409C. The 5-FM labeling intensity of this variant was consistently and reproducibly high when MTSET was added after preincubation with salicin (Fig. 3F, right lane), whereas the intensity observed when salicin was not added was intermediate to low and inconsistent as described above (a common result is presented in Fig. 3F, left lane). Hence, salicin


Accessibility and deduced location of the positions in BglF that were mutated to cysteines

The abbreviations used are as follows: WC, whole cells; ISO, membrane vesicles with inside-out orientation; IF, interface; ND, not determined.

| Position | Accessibility to MTSEA | Accessibility to MTSET | Deduced location | Remarks |
|----------|------------------------|------------------------|-----------------|---------|
|          | in WC                  | in ISO                 |                  |         |
| Cys-24   | +                      | –                      | ND               | Cytoplasm Active site |
| Leu-97   | +                      | –                      | Cytoplasm        |
| Leu-125  | +                      | –                      | Periplasm        |
| Glu-136  | +                      | –                      | ND               |
| Ser-138  | +                      | –                      | In reentrant loop |
| Tyr-141  | +                      | –                      | Periplasm        |
| Ala-163  | +                      | –                      | ND               |
| Asn-192  | +                      | –                      | Periplasm        |
| Ala-196  | +                      | –                      | ND               |
| Ala-198  | +                      | ±                      | ND               |
| Leu-201  | +                      | –                      | Periplasm        |
| Cys-227  | –                      | –                      | ND               |
| Ser-240  | +                      | –                      | Cytoplasm        |
| Cys-251  | –                      | –                      | TM V             |
| Ser-271  | +                      | –                      | Periplasm        |
| Ala-291  | –                      | –                      | ND               |
| Phe-296  | –                      | –                      | TM VI            |
| Trp-297  | +                      | +                      | Cytoplasm        |
| Cys-313  | +                      | ±^d                    | Cytoplasm        |
| Leu-320  | +                      | ±^d, e                  | Cytoplasm/periplasm | In motion |
| Cys-346  | +                      | –                      | Cytoplasm        |
| Ala-350  | +                      | –                      | Cytoplasm        |
| Leu-364  | +                      | –                      | Cytoplasm        |
| Thr-368  | +                      | –                      | Cytoplasm        |
| Leu-377  | +                      | –                      | Cytoplasm        |
| Cys-387  | –                      | –                      | helix i          |
| Ala-400  | +                      | ±                      | Cytoplasm        |
| Ser-406  | +                      | ±^e                    | Cytoplasm/periplasm | In motion |
| Leu-409  | +                      | ±^e                    | Cytoplasm/periplasm | Blocked by ligand |
| Ser-411  | +                      | ±                      | Cytoplasm        |
| Ile-433  | +                      | ±                      | ND               |
| Gly-434  | +                      | ±                      | Periplasm        |
| Cys-441  | –                      | –                      | ND               |
| Cys-480  | +                      | –                      | Cytoplasm        |

^a Permeable reagent.
^b Impermeable reagent.
^c Based on experimental results and HMMTOP prediction.
^d Accessible to the impermeable reagent MTSES in whole cells.
^e Accessible to MTSET at 4 °C in whole cells.

Prevented access to position 409. Notably, L409C was accessible to MTSET at 4 °C in whole cells (see Fig. 3E). Hence this position is accessible from the periplasm, at least to a certain extent, to both thiol reagents and β-glucosides.

Membrane Topology Prediction—The new version of the HMMTOP prediction method allows the user to submit experimental information, i.e. to define the localization of certain residues or segments (35). The additional information is incorporated into the topology prediction by a conditional probability, greatly improving the prediction accuracy. We took advantage of this tool, and we submitted our results about the accessibility of the various positions in BglF to cell-penetrating and nonpenetrating MTS reagents in whole cells to this software. Before the submission of our experimental data, the HMMTOP algorithm predicted that the C domain of BglF contains 12 TMs (see Fig. 1 and Table I). The integration of the experimental data (summarized in Table III) and the theoretical considerations applied by the algorithm led to several surprising predictions regarding the deduced subcellular localization of certain residues in BglF and their surroundings.

A model that integrates all our experimental data obtained with whole cells, by simply classifying MTSET-accessible positions as facing out and MTSET-inaccessible as facing in, could not be suggested (see “Discussion”). One obstacle was that very few positions could be defined a priori as external because of their unrestricted accessibility to permeable and impermeable MTS reagents from the periplasm (see Table III). Consequently, all topology predictions forced us to assign periplasmic location to residues shown to be little or not accessible to MTSET. The preferred model (Fig. 5), which assembles the experimental results and the consensus predictions utmost, assumes that several residues (Glu-136, Ser-138, Asn-192, Ala-196, Ala-198, Ile-433, and Gly-434) reside in periplasmic loops but are barely or not at all exposed to the periplasm. However, their accessibility to MTSEA indicates that they are secluded in a polar environment, i.e. residing in the membrane interface or aqueous crevices within the membrane rather than in the hydrocarbon core (see “Discussion”). Moreover, the first predicted periplasmic loop might be a reentrant loop, as positions near its ends (125 and 141) are accessible to MTSET, whereas positions in its middle (136 and 138) are not.

The topology of the region between residues 297 and 411 was hard to predict. This region is predicted to form two TMs by the prediction methods (designated i and ii in Fig. 1). The inaccessibility of Cys-387, positioned at the heart of the alleged helix ii, to all MTS reagents supports this prediction. However, if these helices traverse the membrane, the segment connecting them is supposed to face the periplasm, a prediction that is not supported by the complete lack of exposure of all Cys replacements in this segment to MTSET in whole cells (Fig. 3C). The
membrane vesicles to MTSET. The proteins were purified, denatured, labeled with 5-FM, and analyzed as described in Fig. 2A. BglF variants with single cysteines locating to the cytoplasm (L97C) or the periplasm (L201C). C, Cys replacements in the first periplasmic loop. C, Cys replacements in the big cytoplasmic loop.

![Diagram](https://via.placeholder.com/150)

**FIG. 4. Accessibility of the single-Cys BglF variants in ISO membrane vesicles to MTSET.** ISO membranes expressing single-Cys BglF variants were incubated for 2 min, at 25 °C, with or without the freshly prepared membrane-impermeable MTSET reagent at a final concentration of 0.01 mM. The proteins were purified, denatured, labeled with 5-FM, and analyzed as described in Fig. 2A. BglF variants with single cysteines locating to the cytoplasm (L97C) or the periplasm (L201C). C, Cys replacements in the first periplasmic loop. C, Cys replacements in the big cytoplasmic loop.

**TABLE 1. Fluorescence and Coomassie staining of MTSET labeling.**

| Treatment | Fluorescence | Coomassie |
|-----------|--------------|-----------|
| Control   | +            | -         |
| MTSET     | +            | -         |

**Cysteine Accessibility to MTSET in ISO Membrane Vesicles**—To clarify the point of whether the first periplasmic loop is a reentrant loop and to further explore the topology of the region between residues 297 and 411, we studied the accessibility of positions within these regions using membrane vesicles with inside-out orientation. To this end, ISO membrane vesicles prepared from cells expressing BglF variants with Cys replacements in the two regions under investigation were incubated in the absence or presence of the impermeable reagent MTSET at 25 °C. The BglF protein variants were subsequently purified, and labeling of the denatured proteins by the fluorescent reagent 5-FM was monitored as described for whole cells. The L97C and L201C BglF variants served as controls for this set of experiments. As shown in Fig. 4A, incubation of ISO membranes prepared from the L97C variant with the impermeable, nonfluorescent reagent MTSET resulted in complete inhibition of the fluorescent labeling with 5-FM, whereas incubation of L201C ISO membranes did not result in inhibition of 5-FM labeling. Hence, cysteines planted in these positions exhibited the accessibility patterns expected from their predicted location and validated the results obtained with whole cells, i.e., a cytoplasmic location for Leu-97 and a periplasmic location for Leu-201. Cys-251, which did not react with both permeable and impermeable MTS reagents in whole cells, was also inaccessible to MTSET in ISO vesicles (not shown), demonstrating that under these conditions MTSET does not interact with residues that are buried in a nonaqueous environment. These results demonstrate that, under the conditions employed with ISO membranes, the membrane-impermeant MTSET reagent modifies only intracellular residues. Indeed, a cysteine planted in position 480, which resides in the linker that connects BglF membrane domain to the hydrophilic A domain, was modified very efficiently in ISO vesicles (not shown).

By using the same experimental conditions, we studied the accessibility of the four cysteines planted in the predicted first periplasmic loop. The results are presented in Fig. 4B. Cys residues at positions 125 and 141 behaved like the cysteine in position 201, as indicated by the lack of inhibition of 5-FM labeling following prior treatment with MTSET. Hence, these positions, shown to be accessible to this reagent from the periplasm in whole cells, are inaccessible to it from the cytoplasmic side of the membrane. The Cys residue at positions 136 was partially blocked by MTSET in ISO vesicles. An important result was obtained with position 138, which was efficiently blocked by MTSET in ISO membranes (68% inhibition). Hence, position 138, shown in whole cells not to be accessible to MTSET from the periplasm, is accessible to this reagent from the cytoplasm. The finding that a loop residue, which is accessible from the cytoplasm, is flanked by two residues, within the same loop, with accessibility from the periplasm demonstrates that this is an extracellular loop that reenters the membrane.

Next, we studied the accessibility of the cysteines planted in the region between residues 297 and 411 to MTSET in ISO membrane vesicles. As can be seen in Fig. 4C, three patterns of behavior were observed. Positions 320, 350, 377, and 409 were very accessible to MTSET in ISO membranes (70–80% inhibition of 5-FM labeling). Positions 346, 364, 368, 400, and 406 (exemplified by position 400 in Fig. 4C) were partially accessible to MTSET in ISO membranes (25–45% inhibition of 5-FM labeling). Positions 297, 313, and 411 (exemplified in Fig. 4C by position 297) were very poorly (313 and 411) or not at all (297) accessible to MTSET in ISO membranes. Hence, at 25 °C, most positions looked upon in this region were fully or partially exposed to the cytoplasm (as observed in ISO membranes), but either transiently (on both sides of this region) or not at all (in the middle of the region) to the periplasm (as observed in whole cells). Taken together, the accessibility mode of this region precludes its description as containing two TM helices connected by a periplasmic loop and suggests that it is largely cytoplasmic, albeit dynamic.

**DISCUSSION**

The BglF membrane protein catalyzes concomitant transport and phosphorylation of β-glucosides across the cytoplasmic membrane and regulates the activity of the transcriptional regulator BglG by reversible phosphorylation (2, 3, 36). We report here a study of BglF membrane topology in whole cells and in membrane vesicles with inside-out orientation. The first step was to plant cysteine residues in the membrane domain and its environs in a Cys-less BglF variant. Out of 36 single-Cys variants, the expression of only one, T130C, was not detected. All the others were expressed at a level similar to wild-type BglF. The activity of the vast majority of the expressed BglF mutants was comparable with that of wild-type BglF, demonstrating that almost all substitutions were neither residues essential for BglF activity nor of positions whose replacement leads to structural perturbations. Substitution of Asn-192 by cysteine even improved sugar phosphorylation (125% activity). One substitution that affected BglF activity was T368C, which reduced the protein ability to transport and/or phosphorylate β-glucosides to 42% of wild-type activity. Replacement of Thr-368 was expected to affect sugar phosphorylation by BglF, as it is part of the GLX motif suggested to play a role in PTS carbohydrates phosphorylation (37). Substitution of V432C completely abolished the protein activity as a sugar phosphorylase catalyzing sugar phosphorylation. Replacement of the residue following it, Ile-433, also affected sugar phosphorylation (67% activity). Both V432C and I433C variants were active in BglG regulation. Hence, this region is critical for the activity of BglF as a sugar phosphorylase. This notion is supported by the finding that residues immediately preceding Ile-432 are involved in β-glucosides recognition.2

To assess the membrane sidedness of the planted cysteines, we used three small, flexible, and charged MTS reagents that react rapidly and specifically with sulphydryl groups of cytos-
Dynamic Membrane Topology of BglF

...results with whole cells and with membrane vesicles with inside-out orientation.

The Topological Model Deduced for BglF—We submitted our results about the accessibility of the various positions in BglF to MTS reagents to the HMMTOP prediction method, which allows the user to define the localization of certain residues or segments (35). The experimental results imposed some constraints that could not be met by the prediction method, probably due to the drawbacks of these algorithms, the main one is their ability to predict only the architecture of the segments that putatively form the transmembrane helices. Information about other segments, such as loops, in transmembrane proteins/domains is indirectly inferred from the number, position, and orientation of the TMS. This indirectly deduced information is relevant only to the subcellular localization of the other elements, and not to their folding, and even this is not very reliable, as reentrant loops cannot be predicted nor can different types of pores/channels be anticipated. Powerful prediction methods for the nonmembranous elements are not available.

To resolve some of the conflicts between the experimental data, obtained with whole cells, and the structural criteria imposed by the algorithm, we carried out experiments with ISO membrane vesicles. The assembly of all experimental data (summarized in Table III) and structural constraints underlying the algorithm, such as hydrophobicity profile, minimal TM length, etc., yielded the topological model shown in Fig. 5. Our results are compatible with the existence of eight TMs in the BglF membrane domain, which are connected by four periplasmic and three cytoplasmic loops. We demonstrated that the first periplasmic loop is a reentrant loop. The third cytoplasmic loop is rather sizeable. We suggest that it contains two additional helices and that it forms a dynamic structure, with segments near its termini (marked by a dotted line in Fig. 5) flipping between faced-in and -out states (see below).

Several lines of data support this model as follows. The predicted model conforms to the positive-inside rule, which states that loops that do not translocate across the membrane contain more positively charged residues than do translocated loops (43). Orientation of membrane proteins is presumably determined by the asymmetric distribution of positively charged residues in cytoplasmic and translocated loops (reviewed in Ref. 44). In BglF, positively charged residues are preferentially found in the cytoplasmic A and B domains and in the three cytoplasmic loops of the C domain (two in the first loop, three in the second loop, and seven in the third loop). Only two positively charged residues are found in periplasmic loops, of which one seems to be embedded in the membrane.

We have shown before that native cysteine(s) in the C domain can form a disulfide bond with Cys-24, the active site cysteine residing in the hydrophilic B domain, in the presence of β-glucosides (17). In an attempt to identify the cysteine that bonds with Cys-24, we studied the reactivity of BglF native cysteines using the thiol-specific alkylating agent 5-iodoacetamidofluorescein. The only reactive cysteines, besides Cys-24, were Cys-346, located at the center of the big cytoplasmic loop in the C domain, and Cys-480 in the A domain. The five other native cysteines Cys-227, Cys-231, Cys-313, Cys-387, and Cys-441, all located in the C domain, were not reactive. According to the model presented in Fig. 5, Cys-227, Cys-251, and Cys-441 are buried in TMs, and Cys-387 locates to an α-helix suggested to participate in forming the sugar channel. Thus, the results obtained with the thiol modifier 5-iodoacetamidofluorescein, predicting burial of most native cysteines, are in agreement...
with the model presented here, which is based on a different experimental approach. Moreover, the only cysteine in the C domain that reacted with 5-iodoacetamidofluorescein, Cys-346, is predicted here to be located in a dynamic region and is presumably involved in creating the sugar translocation channel. Hence, formation of a sugar-dependent disulfide bond between Cys-346 and Cys-24 can get the active site cysteine to the vicinity of the sugar and bring about the observed coupling between sugar phosphorylation and translocation.

The data that we submitted to HMMTOP did not include the positions that are not exposed to an aqueous environment (Fig. 3A). Yet, the model (Fig. 5) that best fits all the other requirements placed five of the six group A positions (Cys-227, Cys-251, Ala-291, Phe-296, and Cys-441) within TMs anyhow, and we assume that they are buried in the hydrocarbon core of the membrane.

Finally, the deduced topological model is supported by preliminary results with cell-penetrating cross-linking reagents and di-cysteine BglF mutants, which demonstrated proximity of Cys-24 to positions located solely in the cytoplasm, most of them in the big cytoplasmic loop.3

**The Periplasmic Loops Are in Close Association with the Membrane**—Of the four periplasmic loops, the third and the fourth are very short (10 residues or less). Short loops are thought to be in close association with the phospholipids, whereas long loops can form a well-defined structure in the cytosol or in the periplasm with only their ends interacting with the membrane. The bilayer fabric of the membrane has two chemically distinct regions, the hydrocarbon core and the interfaces. The combined thickness of the interfaces is equal to the 30-Å thickness of the hydrocarbon core (45). The thickness of a single interface (15 Å) can easily accommodate an α-helix parallel to the membrane plane or a short unstructured loop. Out of the six residues that compose the fourth periplasmic loop, two (Ile-433 and Gly-434) demonstrate intermediate and irreversible accessibility to MTSET at 25 °C, suggesting that they are partially buried in the membrane. The fact that they are accessible to MTSEA suggests that they are not buried in the hydrocarbon core but rather in the interface. Similar results were observed with the residue preceding them, Val-432, but are not reported here, as Cys replacement of this residue resulted in a protein defective in sugar phosphorylation.

Our results demonstrate that the first periplasmic loop, which is more than 20 residues long, is a reentrant loop, with positions near both ends facing the periplasm and residues in the middle being accessible from the cytoplasm. All these positions are accessible to MTSEA, indicating that this loop is in an aqueous environment. Because the tip of reentrant loops appears not to reach the level of the phospholipids head groups on the other side of the membrane, based on the solved structure of potassium channel (46), this loop might locate to a water-filled crevice within the membrane. Reentrant loops have been documented in a number of channels and transporters, where they have been implicated in permeation (e.g., Refs. 46–50). Results from our laboratory suggest that residues in this loop are involved in sugar recognition.7 It might be that in the three-dimensional structure of BglF, this loop is part of or associated with the sugar translocation channel, which seems to consist mainly of helices VI and VII and the big cytoplasmic loop.

The second periplasmic loop, which is 17 amino acids long, seems to be partially embedded in the membrane. This hypothesis is based on our experiments in whole cells, in which positions in the N-terminal half of this loop (192 and 196) were inaccessible to MTSET, a position in the middle (198) showed partial accessibility to MTSET, and a position in the second half of this loop (201) was exposed to this impermeable reagent (Table III). As all these positions were accessible to MTSEA, the N-terminal half is supposedly buried in the membrane interface. Although this loop contains a large amount of charged and polar residues and, hence, is not predicted to form an α-helix, the hypothesis that helices III and IV, together with the second periplasmic loop that connects them, form an irregular α-helix cannot be completely ruled out. Because of the restrictions imposed by prediction algorithms, among them is the length of TMs, irregular structures cannot be predicted. However, many helices are apparently irregular, i.e., they do not span the membrane in a straight way. As a matter of fact, the crystal structure of LacY shows that out of 12 TMs, only 3 helices appear to be straight, and all the others are arched, kinked, or S-shaped (51). The possibility that the second periplasmic loop is also a reentrant loop cannot be ruled out, and its consideration requires substitutions of more residues in this region by cysteines.

In summary, although large extra-membranous domains do not seem to be much influenced by their attachment to the bilayer, and can essentially be viewed as membrane-tethered globular proteins (52), this does not seem to be the case with BglF periplasmic loops. Our experimental evidence indicates that not only the short periplasmic loops, but also the long ones, are closely associated with the membrane. Hence, BglF sequences do not seem to be much exposed to the periplasm. An accepted hypothesis for membrane proteins, which are implicated in transport, is that sequences facing the periplasm are involved in ligand recognition. The observation that some of BglF periplasmic loop residues that associate with the membrane are seemingly embedded in the interface suggests that they are accessible to the sugar ligand. Moreover, as β-glucosidases are amphipathic in their nature, their recognition might be preferably mediated by residues embedded in the interface.

The Large Cytoplasmic Loop Participates in Forming the Sugar Translocation Channel—The large cytoplasmic loop connecting putative TMs VI and VII is predicted, based on the hydrophobicity profile, to be a trans-membranous region. Although our data suggest that this region is largely cytoplasmic, it is consistent with the notion that segments within this loop fold into α-helices i and ii (see above). Had these helices been regular TMs, the segment connecting them would have been expected to face the periplasm. However, all cysteine residues planted in this segment (in positions 346, 350, 364, 368, and 377) are inaccessible from the periplasm (Fig. 3C) and are either completely or partially accessible from the cytoplasm (Fig. 4C). The possibility that this segment is a reentrant loop, which enters the membrane twice from the periplasmic face, leaving the central part in the cytoplasm, is not supported by out data, as we did not find residues in this segment that face the periplasm. However, to investigate this possibility further, a more systematic Cys scanning of this segment is necessary.

The Cys accessibility analyses in whole cells suggest that the regions encompassing helix i and ii alternate between face-in and -out states at 25 °C. The behavior of this region was more consistent at a lower temperature or in the presence of the sugar substrate. However, even at 25 °C, two positions (313 and 320) that demonstrated partial accessibility to MTSET were completely exposed to another impermeable reagent MTSES. The model, drawn in Fig. 5, which takes the inconsistent behavior of these regions into consideration, shows schematically that these regions (marked by dotted lines) and helices i and ii are capable of moving in and out the membrane, thereby causing the neighboring helices VI and VII to slightly change their membrane insertion. Helix VI is the one on which there is total lack of agreement.
between the prediction methods (Fig. 1). The amino acid sequence in this region may allow for substantial changes in the boundaries and the positioning of the helix. Helix VII is predicted only by three out of five algorithms. The observation that positions at the periplasmic tail of helix VII (433 and 434) are partially and inconsistently accessible to MTSET at 25 °C supports the notion that the entire region is characterized by conformational flexibility. The finding that position 409 is inaccessible to MTSET in the presence of β-glucosides can be interpreted either as ligand sequestration of this position or as reflecting ligand-induced conformational alterations.

A model that can account for all these results is that the big cytoplasmic loop, the helices encompassing it, and the fourth periplasmic loop are involved in forming a gated channel that allows for occasional interaction of impermeable MTS reagents with individual positions at the gate. The inconsistent nature of this interaction could also be explained by reorientation of an empty sugar-binding site from one surface of the membrane to the other that is pertinent to the catalytic cycle of the permease. Minimizing thermal protein motion by lowering the temperature facilitates the interaction, whatever the mechanism is. Helices i and ii may serve as building blocks of the sugar channel and facilitate its positioning in or near the membrane. As this region exceeds in size BglF cytoplasmic domains A and B, it is reasonable to assume that it forms a functional domain. The finding that Val-432 and Ile-433 at the tail of the TM VII are essential for sugar phosphotransfer supports the involvement of this region in sugar transport. When envisioning the sugar channel, one should keep in mind that BglF dimerizes in the membrane spontaneously (5). Hence, the channel is probably formed by at least two or, based on work with other PTS permeases (see below), more monomers.

**Implications for PTS Permeases Topology and Sugar Transport**—Attempts to crystallize PTS sugar permeases yielded crystals with poor diffraction. Projection maps of two-dimensional crystals, obtained for the E. coli mannitol and glucose permeases, indicate the presence of dimers centered around a 2- and 3-fold axis, respectively (53, 54). However, the level of the resolution does not provide novel topological and structural information. Hence, despite the extensive study of PTS permeases, the nature of the sugar translocation channel is not known.

Our study is the first detailed topological study with an intact PTS permease. Previous systematic investigations of PTS permeases employed strategies to assess the membrane orientation of reporter epitopes, PhoA and β-galactosidase, fused to sequential C-terminal deletions or nested deletions of PTS permeases (reviewed in Refs. 10 and 55). These studies, which were supported by random linker insertion mutagenesis (56), suggested the existence of six to eight TMs and the presence of one large (more than 80 residues) hydrophilic loop at the cytoplasmic side of the membrane (57). The amphipathic leader immediately preceding TM1 and the first two TMs were suggested to be involved in targeting and correct insertion of the N-terminal part of the C domain into the membrane (11, 58, 59). The C-terminal TM was also suggested to be implicated in stable anchoring of the permease in the membrane (60). The existence of reentrant loops has not been decisively demonstrated in other PTS permeases before. A recent publication (61) raises the possibility that the first cytoplasmic loop in the mannitol permease, originally suggested to contain only six TMs and two large cytoplasmic loops, may fold back into the membrane. Although this region in BglF and in the mannitol permease show marked differences, i.e. in BglF it is longer and predicted with high probability to contain two TMs (Fig. 1, TM III and TM VI), as opposed to one TM in EIEmtl (61), it is possible that in both permeases it spans the membrane irregularly (see the discussion on BglF second periplasmic loop above). The possible role of the reentrant loop described here for BglF is not known.

The big cytoplasmic loop is relatively preserved, although some studies placed a region corresponding to part of this loop in the membrane (62). Several residues in the big cytoplasmic loop appear to have an important role in sugar binding and/or phosphorylation, in particular the highly conserved GIxxE motif (reviewed in Refs. 10 and 55). This motif was suggested to play an important role in substrate binding and translocation (63). We substituted the third and least conserved residue in this motif. The resulting T368C variant demonstrated only 42% phosphotransfer activity, highlighting the importance of this position for sugar phosphotransfer. As this region is neither periplasmic nor permanently trans-membranous, it has been suggested that it projects up into the channel formed by TMs (55, 64), or alternatively serves as a cap or gate for such a channel. Studies with the mannitol permease suggested that the when the mannitol-binding site is loaded by the sugar, it converts from a periplasm-facing to a cytoplasm-facing orientation (reviewed in Ref. 10). The inward- and outward-facing states are presumably related through a state in which the mannitol-binding site is occluded (65). Newly predicted topological models for the mannitol permease suggest that, similarly to BglF, the big cytoplasmic loop contains two α-helices (61). Based on all the above, it seems reasonable that residues in BglF that face the periplasm are involved in substrate recognition and confer specificity, whereas the big cytoplasmic loop and the TMs encompassing it are involved in sugar translocation. The interaction of residues in the big cytoplasmic loop with the Cys-24 active site (17) provides an explanation for the coordination between the processes of sugar transport and phosphorylation.

**Acknowledgments**—We thank Carmit Miyara and Yuval Nevoh for technical assistance. We thank Galya Mondere-Rothkoff for the gift of plasmid pZGM-B. We appreciate the helpful discussions with Galya Mondere-Rothkoff, Liat Fux, and Noa Gordon. We acknowledge Shimon Shuldiner, Eitan Bibi, Baruch Kanner, Etana Padan, and Shai Apikin for helpful discussions. We thank Hannah Rahamimoff for suggestions on the use of TCEP. We are grateful Liat Fux and Galya Mondere-Rothkoff for critical reading of the manuscript.

**REFERENCES**

1. Fox, C. F., and Wilson, G. (1968) Proc. Natl. Acad. Sci. U. S. A. 59, 988–995.
2. Amster-Choder, O., Houman, F., and Wright, A. (1989) Cell 58, 847–855.
3. Amster-Choder, O., and Wright, A. (1990) Science 249, 540–542.
4. Amster-Choder, O., and Wright, A. (1992) Science 257, 1395–1398.
5. Chen, Q., and Amster-Choder, O. (1998) Biochemistry 37, 8714–8723.
6. Crutz, A. M., Steinmetz, M., Ayerich, S., Roselyne, R., and Le Cog, D. (1999) J. Bacterial. 172, 1043–1050.
7. Idelson, M., and Amster-Choder, O. (1998) J. Bacterial. 180, 660–666.
8. Lee, S., Boos, W., Bouché, J., and Plumbridge, J. (2000) EMBO J. 19, 5353–5361.
9. Tanaka, Y., Kimata, K., and Aiba, H. (2000) EMBO J. 19, 5344–5352.
10. Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993) Microbiol. Rev. 57, 543–594.
11. Lengeler, J. W., Jahreis, K., and Wehmeier, U. F. (1994) Biochim. Biophys. Acta 1188, 1–29, and references therein.
12. Beshard, G. T., and Bress, J. (1999) Biochim. Biophys. Acta 1422, 73–104.
13. Chen, Q., Arents, J. C., Bader, R., Postma, P., and Amster-Choder, O. (1997) EMBO J. 16, 4617–4627.
14. Chen, Q., Postma, P., and Amster-Choder, O. (2000) J. Bacterial. 182, 2033–2036.
15. Chen, Q., and Amster-Choder, O. (1998) Biochemistry 37, 17040–17047.
16. Chen, Q., and Amster-Choder, O. (1999) J. Bacterial. 181, 462–468.
17. Chen, Q., Nusebuh-Shochat, A., and Amster-Choder, O. (2001) J. Biol. Chem. 276, 44751–44756.
18. Siebold, C., Flukiger, K., Butler, R., and Erni, B. (2001) FEBS Lett. 504, 104–111.
19. Mahadevan, S., Reynolds, A. E., and Wright, A. (1987) J. Bacterial. 169, 2570–2578.
20. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Gene (Amst.) 77, 51–59.
21. Schaefler, S. (1967) *J. Bacteriol.* 93, 254–263
22. Ninio, S., Elbaz, Y., and Schuldiner, S. (2004) *FEBS Lett.* 562, 193–196
23. Rosen, B. P. (1986) *Methods Enzymol.* 125, 328–336
24. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) *J. Mol. Biol.* 305, 567–580
25. Tumyady, G. E., and Simon, I. (1998) *J. Mol. Biol.* 283, 489–506
26. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1994) *Biochemistry* 33, 3038–3049
27. von Heijne, G. (1992) *Protein Sci.* 1, 1704–1718
28. Rost, B., Fariselli, P., and Casadio, R. (1996) *FEBS Lett.* 386, 267–269
29. Nilsson, J., Persson, B., and von Heijne, G. (2000) *Protein Sci.* 9, 567–580
30. Nilsson, J., Persson, B., and Von Heijne, G. (2002) *Protein Sci.* 11, 2974–2980
31. Getz, E. B., Xiao, M., Chakrabarty, T., Cooke, R., and Selvin, P. R. (1999) *Anal. Biochem.* 273, 567–580
32. Ren, X., Kasir, J., and Rahamimoff, H. (2001) *J. Biol. Chem.* 276, 9572–9579
33. Venkatesan, P., Liu, Z., Hu, Y., and Kaback, H. R. (2000) *Biochemistry* 39, 10649–10655
34. Zhang, W., Hu, Y., and Kaback, H. R. (2000) *Biochemistry* 39, 10649–10655
35. Tusnady, G. E., and Simon, I. (2001) *J. Mol. Biol.* 305, 567–580
36. Schnetz, K., and Rak, B. (1990) *Proc. Natl. Acad. Sci. U. S. A.* 87, 5074–5078
37. Lengeler, J. W., Titgemeyer, F., Vogler, A. P., and Wohrl, B. M. (1990) *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 326, 489–504
38. Akabas, M. H., Stauffer, D. A., Xu, M., and Karlin, A. (1992) *Science* 258, 307–310
39. Stauffer, D. A., and Karlin, A. (1994) *Biochemistry* 33, 6840–6849
40. Roberto, D. D., Lewis, S. D., Bailou, D. P., Olson, S. T., and Shafer, J. A. (1986) *Biochemistry* 25, 5095–5101
41. Karlin, A., and Akabas, M. H. (1998) *Methods Enzymol.* 293, 123–145
42. Holmgren, M., Liu, Y., Xu, Y., and Yellen, G. (1996) *Neuropharmacology* 35, 797–804
43. von Heijne, G., and Gavel, Y. (1988) *Eur. J. Biochem.* 174, 671–678
44. von Heijne, G. (1994) *Annu. Rev. Biophys. Biomol. Struct.* 23, 167–196
45. White, C. H., Lolkema, G. K., Robillard, G. T., and Brisson, A. (1999) *J. Biol. Chem.* 276, 32395–32398
46. Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998) *Science* 280, 69–77
47. MacKinnon, R. (1995) *Neuron* 14, 889–892
48. Slotboom, D. J., Rognes, T., Krogan, N. J., and Lolkema, J. S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 14282–14287
49. Grunewald, M., Menaker, D., and Kanner, B. I. (2002) *J. Biol. Chem.* 277, 26074–26080
50. Gasol, E., Jimenez-Vidal, M., Chillaron, J., Zorzano, A., and Palacin, M. (2004) *J. Biol. Chem.* 279, 31228–31236
51. Abramson, J., Smirnova, I., Kasho, V., Geshke, G., Kaback, H. R., and Iwata, S. (2003) *Science* 301, 610–615
52. von Heijne, G. (1996) *Protein Sci.* 5, 1704–1718
53. Koning, R. I., Keegstra, W., Oostergetel, G. T., Schuurman-Wolters, G., Robillard, G. T., and Brissin, A. (1999) *J. Mol. Biol.* 287, 845–851
54. Zhuang, J., Gutknecht, R., Flugger, K., Hasler, L., Erni, B., and Engel, A. (1999) *J. Mol. Biol.* 326, 89–96
55. Jacobson, G. R., and Saraseni-Richards, C. (1993) *J. Bioenerg. Biomembr.* 25, 621–626
56. Beutler, R., Kaufmann, M., Ruggiero, F., and Erni, B. (2000) *Biochemistry* 39, 3745–3750
57. Sugiyama, J. E., Mahmoodian, S., and Jacobson, G. R. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 9603–9607
58. Yamada, Y., Chang, Y. Y., Daniels, G. A., Wu, L. P., Tomich, J. M., Yamada, M., and Saier, M. H., Jr. (1995) *J. Biol. Chem.* 270, 17863–17871, and references therein
59. Aboulwafa, M., Chung, Y. J., Wai, H. H., and Saier, M. H., Jr. (2003) *Microbiology* 149, 763–771, and references therein
60. Briggs, C. E., Khandekar, S. S., and Jacobson, G. R. (1992) *Res. Microbiol.* 143, 139–149
61. Verloop, E. B., Bulterman, J. B., Schuurman-Wolters, G. K., Deursen, M. J., and Poolman, B. (2005) *J. Mol. Biol.* 346, 889–894
62. Buhr, A., and Erni, B. (1993) *J. Biol. Chem.* 268, 11599–11603
63. Saraseni-Richards, C. A., and Jacobson, G. R. (1997) *J. Bacteriol.* 179, 1135–1142
64. Lengeler, J. W. (1990) *Biochim. Biophys. Acta* 1018, 155–159
65. Lolkema, J. S., Dijkstra, D. S., and Robillard, G. T. (1992) *Biochemistry* 31, 5514–5521