Membrane topology prediction by hydropathy profile alignment: Membrane topology of the Na+-glutamate transporter GltS

Dobrowolski, A.J.; Sobczak-Elbourne, I.; Lolkema, J.S.

Published in:
Biochemistry

DOI:
10.1021/bi062275i

Publication date:
2007

Membrane Topology Prediction by Hydropathy Profile Alignment: Membrane Topology of the Na^+-Glutamate Transporter GltS[†]

Adam Dobrowolski, Iwona Sobczak-Elbourne,[‡] and Juke S. Lolkema*

Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

Received November 3, 2006; Revised Manuscript Received December 8, 2006

ABSTRACT: Structural classification of families of membrane proteins by bioinformatics techniques has become a critical aspect of membrane protein research. We have proposed hydropathy profile alignment to identify structural homology between families of membrane proteins. Here, we demonstrate experimentally that two families of secondary transporters, the ESS and 2HCT families, indeed share similar folds. Members of the two families show highly similar hydropathy profiles but cannot be shown to be homologous by sequence similarity. A structural model was predicted for the ESS family transporters based upon an existing model of the 2HCT family transporters. In the model, the transporters fold into two domains containing five transmembrane segments and a reentrant or pore-loop each. The two pore-loops enter the membrane embedded part of the proteins from opposite sides of the membrane. The model was verified by accessibility studies of cysteine residues in single-Cys mutants of the Na^+-glutamate transporter GltS of *Escherichia coli*, a member of the ESS family. Cysteine residues positioned in predicted periplasmic loops were accessible from the periplasm by a bulky, membrane-impermeable thiol reagent, while cysteine residues in cytoplasmic loops were not. Furthermore, two cysteine residues in the predicted pore-loop entering the membrane from the cytoplasmic side were shown to be accessible for small, membrane-impermeable thiol reagents from the periplasm, as was demonstrated before for the Na^+-citrate transporter CitS of *Klebsiella pneumoniae*, a member of the 2HCT family. The data strongly suggests that GltS of the ESS family and CitS of the 2HCT family share the same fold as was predicted by comparing the averaged hydropathy profiles of the two families.

Historically, hydropathy profiles of the amino acid sequence have played an important role in membrane protein research. In 1982, the proposal of Kyte and Doolittle for a simple method for displaying the hydropathic character of a protein resulted in the first widely used secondary structure prediction programs for membrane proteins (*1*). The alternating hydrophobic and hydrophilic regions in the profiles that correspond to the transmembrane α-helical segments (TMS)[1] and connecting loop regions of the protein, respectively, provided a membrane topology model of the protein that is a basic aspect of its structure. In the past, we have explored the idea that, in addition to secondary structure information, hydropathy profiles also contain tertiary structure information in the sense that a hydropathy profile is characteristic for a specific fold of the protein in the membrane (*2*). Within families of membrane proteins, the hydropathy profiles of different members are strikingly similar even though sequence identity may be as low as 20−25%. Apparently, the hydropathy profiles, like the 3D structures of homologous proteins are much better conserved than their amino acid sequence and, therefore, they report on the global fold of the proteins in a family. Examining hydropathy profiles provides a mechanism to identify distantly related membrane proteins even when sequence identity is too low to detect homology. This has led to the MemGen classification in which families of membrane proteins are grouped in structural classes by comparing the averaged hydropathy profiles of the families (*2*, *3*). Two classes of secondary transporters have been characterized in detail by the MemGen approach: class ST[3] groups 33 families that together contain 2051 sequences, and class ST[4], which contains 399 sequences distributed over two families (*3*, *4*). The relevance of a structural classification scheme for membrane proteins is evident when considering the high sequence diversity of, for example, secondary transport proteins, which results in over 100 families in the transporter classification (TC) system (*5*).

The MemGen classification system is not a membrane topology prediction method *per se*, but a major consequence of the approach is that all proteins in the different families in one class share the same fold, i.e., knowing the topology of one is knowing them all. This provides a mechanism to

[†] This work was supported by grants from The Netherlands Organization for Scientific Research (NWO).

* Corresponding author. Molecular Microbiology, Biological Centre, Kerklaan 30, 9751NN Haren, The Netherlands. E-mail: j.s.lolkema@rug.nl. Telephone: +31 50 363 2155. Fax: +31 50 363 2154.

[‡] Present address: University Centre for Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9700 AD Groningen, The Netherlands.

[1] Abbreviations: TMS, transmembrane segment; RSO, right-side-out; PMS, phenazine methosulfate; NEM, *N*-ethylmaleimide; AmdiS, 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid; MTSET, [2-(trimethylammonium)ethyl] methanethiosulfonate bromide; MTSES, (2-sulfonatoethyl) methanethiosulfonate; FM, fluorescein-5-maleimide.

10.1021/bi062275i CCC: $37.00
Published on Web 02/02/2007

validate the principle that hydropathy profiles represent a certain fold. Experimental support for the method came from an analysis of the membrane proteome of *Escherichia coli* in which the cellular location of the C-terminus of all membrane proteins was determined ("in" or "out") (*6*). The C-terminal location of all 19 *E. coli* proteins found in structural class ST[3] in the MemGen classification (covering seven different families) was correctly predicted based upon the structural model of the 2-hydroxycarboxylate transporter (2HCT) family, which is also in ST[3] (*7*).

In this study, we present the first detailed experimental support for the MemGen classification scheme. Prominent features of the structural model of the 2HCT family are two homologous domains that are oppositely oriented in the membrane (inverted topology) and each containing a reentrant or pore-loop structure (reviewed in *8*). The model was used to predict the membrane topology of the ESS family in class ST[3], a family of Na^+-dependent glutamate transporters. No significant (local) sequence identity is detected in a BLAST search (*9*) between members of the two families. Importantly, the MemGen model differs from the model predicted by TMHMM (*10*), which in a recent survey was shown to perform well relative to a number of other membrane topology predictors (*11*). Here, we demonstrate by experiment that the model based on the MemGen classification is the right model.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Growth Conditions, and GltS Constructs. *Escherichia coli* strains DH5α and ECOMUT2 (*12*, *13*) were routinely grown in Luria−Bertani broth (LB) medium at 37 °C under continuous shaking at 150 rpm. When appropriate, the antibiotics ampicillin and chloramphenicol were added at final concentrations of 50 and 30 μg/mL, respectively. All genetic manipulations were done in *E. coli* DH5α, while the GltS protein was expressed in *E. coli* ECOMUT2-harboring plasmid pBAD24 derivatives coding for wild type or cysteine mutants of GltS extended with six histidine residues at the N-terminus (His-tag). Production of the GltS proteins was induced by addition 0.1% arabinose when the optical density of the culture measured at 660 nm (OD_{660}) reached a value of 0.6. The Cys-less version of GltS and the single-Cys mutants of GltS were constructed by PCR using the QuickChange site-directed mutagenesis kit (Stratgene, La Jolla, CA). All mutants were sequenced to confirm the presence of the desired mutations (ServiceXS, Leiden, The Netherlands).

Transport Assays in RSO Membranes. *E. coli* ECOMUT2 cells expressing GltS variants were harvested from a 1 L culture by centrifugation at 10 000 × g for 10 min at 4 °C. Right-side-out (RSO) membrane vesicles were prepared by the osmotic lysis procedure as described (*14*). RSO membranes were resuspended in 50 mM KPi pH 7, rapidly frozen, and stored in liquid nitrogen. Membrane protein concentration was determined by the DC protein assay kit (Bio-Rad Laboratories, Hercules, CA). Uptake by RSO membranes was measured by the rapid filtration method in 50 mM KPi pH 6.0 containing 70 mM NaCl at 30 °C as described (*15*). The membranes were energized using the K-ascorbate/phenazine methosulfate (PMS) electron donor system (*16*).

Partial Purification of GltS Derivatives by Ni^{2+}-NTA Affinity Chromatography. *E. coli* ECOMUT2 cells expressing single-Cys mutants of GltS were harvested from a 200 mL culture by centrifugation at 10 000 × g for 10 min at 4 °C. The cells were washed once with 50 mM KPi pH 7.0 buffer and subsequently resuspended in 2 mL of the same buffer. The cells were broken by a Soniprep 150 sonicator operated at an amplitude of 8 μm by nine cycles consisting of 15 s ON and 45 s OFF. Cell debris and unbroken cells were removed by centrifugation at 9000 rpm for 10 min. Membranes were collected by ultracentrifugation for 25 min at 80 000 rpm at 4 °C in a Beckman TLA 100.4 rotor and washed once with 50 mM KPi pH 7.0. The His-tagged GltS derivatives were partially purified from the cytoplasmic membranes using Ni^{2+}-NTA affinity chromatography as follows. The membranes (4 mg/mL) were solubilized in 50 mM KPi pH 8, 400 mM NaCl, 20% glycerol, and 1% Triton X-100, followed by incubation on ice for 30 min. Undissolved material was removed by ultracentrifugation at 80 000 rpm for 25 min at 4 °C. The supernatant was mixed with Ni^{2+}-NTA resin (50 μL bed volume per 5 mg protein), equilibrated in 50 mM potassium phosphate pH 8.0, 600 mM KCl, 10% glycerol, 0.1% Triton X-100, 10 mM imidazole, and incubated overnight at 4 °C under continuous shaking. Subsequently, the column material was pelleted by pulse centrifugation and the supernatant was removed. The resin was washed with 10 volumes of equilibration buffer containing 300 mM KCl and 40 mM imidazole. The protein was eluted with half a bed volume of the washing buffer but containing 150 mM imidazole. The eluted fraction was stored at −20 °C until use.

Treatment with Thiol Reagents. Stock solutions of MTSES, MTSET, NEM, and AmdiS were prepared freshly in water. The treatment of the different reagents was stopped by addition of an equal concentration of dithiothreitol (DTT) in the case of NEM and AMdiS, and L-cysteine in case of the MTS reagents. The presence of DTT or L-cysteine did not affect the uptake rate in control experiments.

RSO membranes at a concentration of 1 mg/mL were treated for the indicated times and at the indicated temperatures with the thiol reagents in 50 mM KPi pH 7.0. Following treatment, RSO membranes were diluted twice into 50 mM KPi pH 5.0 containing 140 mM NaCl. The pH of the resulting suspension was 6.0, and the suspension was immediately used for uptake measurements.

E. coli ECOMUT2 cells expressing single-Cys mutants of GltS from a 400 mL culture were washed and resuspended in 4 mL of 50 mM KPi pH 7.0. Half of the cells were sonicated (see above) in the presence of 1 mM AMdiS, followed by incubation for 30 min at 30 °C and quenching with 1 mM DTT. The other half was treated with 1 mM AmdiS for the same time and at the same temperature, quenched with DTT, and then sonicated. Control experiments were done in an identical manner but omitting the AMdiS treatment. The GltS derivatives were partially purified using Ni-NTA affinity chromatography (see above) and treated with 0.1 mM FM for 5 min at 20 °C. The reaction was stopped with 0.5 mM DTT. Samples of 25 μL volume were mixed with SDS sample buffer and run on a 12% SDS-PAGE gel. Fluorescence of proteins labeled with FM was visualized on a Lumi-Imager F1 imager (Roche Diagnostic GmbH, Mannheim, Germany) by irradiation with UV light using a

520 nm filter. All samples containing FM were kept out of bright light until the gel was exposed. After exposure, the gel was stained with Coomassie brilliant blue (CBB) to compare the GltS protein levels. Quantification was done using the LumiAnalyst 3.1 software package supplied by Roche Diagnostic (Roche Diagnostic GmbH, Mannheim, Germany).

Computational Methods. A structural class in the MemGen classification contains a subset of the entries in the NCBI protein database (http://www.ncbi.nlm.nih.gov/entrez/) that are stored locally in the MemGen database (http://molmic35.biol.rug.nl/). Building of a structural class has been described before in detail (*3*, *4*) and involves a combination of BLAST searches (*9*), multiple sequences alignments (*17*), and hydropathy profile alignments (*2*). The [st324]ESS family in class ST[3] contains 96 entries, 76 of which represent unique proteins. A set of 27 sequences shows no pairwise sequence identity higher than 60% with any of the other sequences in the set, which prevents a bias toward very similar sequences. This set was used to construct the multiple sequence alignment and the family profile in Figure 1. The sequences are listed in the "multiple sequence alignment" section in the MemGen database.

Materials. The methanethiosulfonate (MTS) derivatives MTSET and MTSES were purchased from Anatrace Inc. (Ohio). NEM was purchased from Sigma-Aldrich BV (Zwijdrecht, The Netherlands), AMdiS and FM were purchased from Molecular Probes Europe BV (Leiden, The Netherlands).

RESULTS

Structural Model of GltS of Escherichia coli by MemGen. The Na^+-glutamate transporter GltS of *E. coli* is a member of the [st324]ESS family in the MemGen classification scheme. The ESS family and the 2HCT family ([st326]-2HCT) are found both in structural class ST[3] of the classification (*3*). A detailed structural model of the transporters in the latter family is available (*8*). Optimal alignment of the averaged hydropathy profiles of the two families resulted in a similarity score (*2*) of 0.8, indicating very similar profiles (Figure 1). The alignment allows for a projection of the 2HCT structural model on the ESS family. The transporters of the 2HCT family contain 11 transmembrane segments (TMS) with the N-terminus in the cytoplasm. Two times five segments (TMSs II−VI and TMSs VII−XI) form two homologous domains that have opposite orientations in the membrane. The two domains are connected by a large hydrophilic loop. In between the fourth and fifth TMS of each domain, the connecting loop folds back between the TMSs, forming a so-called reentrant or pore-loop. The pore-loop in the N-terminal domain (loop Vb) enters the membrane-embedded part from the periplasmic side, the one in the C-terminal domain (loop Xa) from the cytoplasmic side of the membrane. The two loops are believed to contact each other in the three-dimensional structure, where they would form the translocation path. The N-terminal TMS is not part of the two-domain structure and seems to form a separate domain by itself. The alignment of the two family hydropathy profiles shows that this N-terminal transmembrane segment is absent in the ESS transporters. The resulting model for GltS of *E. coli* consists of two domains containing five TMSs and one pore-loop structure each (Figure 1). Both the N- and C-termini are located in the periplasm. The model corresponds to the core structure of the ST[3] proteins (*18*).

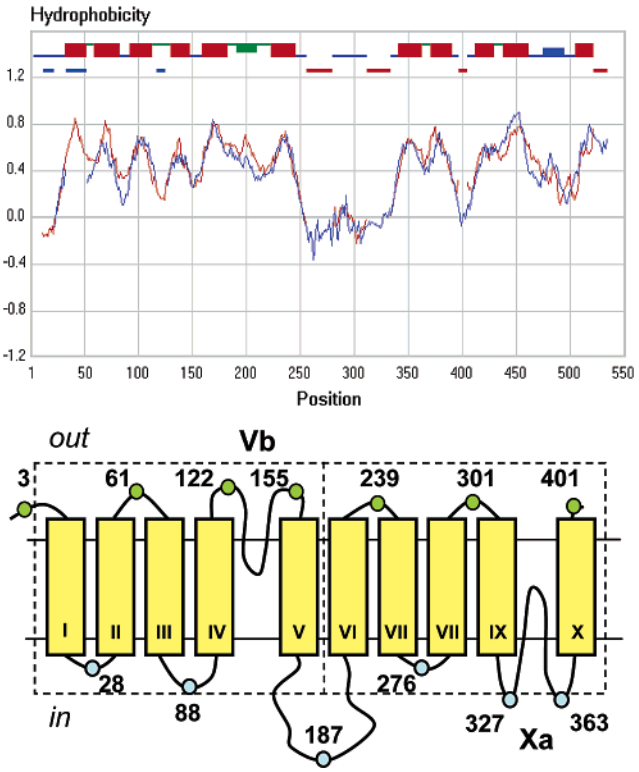

FIGURE 1: Topology model of the ESS family by MemGen. (top) Hydropathy profile alignment of the family profiles of the 2HCT family (red) and the ESS family (blue). The profiles of the 2HCT and ESS families represent 23 and 27 members, respectively, with pair wise sequence identities between 20 and 60%. SDS values (2) of the profiles were 0.114 and 0.117, respectively. The alignment resulted in a similarity score (2) of 0.8. The membrane topology model of the 2HCT family was indicated in the upper part. Transmembrane segments (red boxes), cytoplasmic loops (blue lines), periplasmic loops (green lines). Thickened parts of loop regions indicate the positions of pore-loops. Horizontal blue and red lines indicate positions of gaps introduced by the algorithm in the alignment in the blue and red profiles, respectively. (bottom) Membrane topology model of the ESS family based on the profile alignment. Dashed boxes indicate two homologous domains with inverted orientation in the membrane. Vb and Xa correspond to pore-loop structures. Circles in the loop regions indicate positions for which the location was determined. The numbers correspond to the residues in the GltS protein.

Cloning of the gltS Gene of E. coli and Construction of a Set of Single-Cys Mutants. The *gltS* gene of *E. coli* was cloned into vector pBAD24, yielding pBADHNGltS that codes for the GltS protein extended with six histidine residues at the N-terminus under control of the arabinose promoter. A Cys-less version of the GltS protein was constructed by mutating the four native cysteine residues one by one into serine residues. Transport activity of GltS with L-[^{14}C]-glutamate as the substrate was measured in right-side-out (RSO) membrane vesicles prepared from *E.coli* ECOMUT2 cells, which contain a copy of the *gltS* gene on the chromosome (*13*). Thus, the background activity was estimated in RSO membranes containing the Na^+-citrate transporter CitS of *Klebsiella pneumonia* from the 2HCT family that was produced from exactly the same expression system (pBADCitS) and which does not transport L-glutamate (*19*). RSO vesicles containing recombinant GltS showed at least

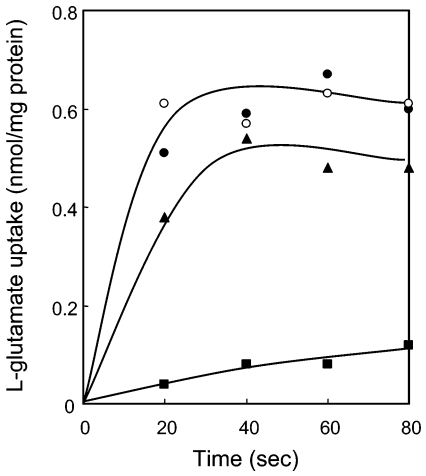

FIGURE 2: Glutamate uptake activity in RSO membranes. L-[^{14}C]-glutamate uptake was measured in RSO membrane vesicles expressing CitS of *K. pneumoniae* (■), GltS of *E. coli* before (●) and after (○) treatment with 1 mM NEM for 10 min at room temperature, and the Cys-less version of GltS (▲).

a five times higher uptake activity in comparison to RSO vesicles containing only endogenous GltS (Figure 2, ● and ■). RSO vesicles containing the Cys-less version of the GltS protein showed a comparable level of uptake activity as the wild type GltS protein. (Figure 2, ▲). A set of 15 single-Cys mutants was constructed, using site-directed mutagenesis and the Cys-less version of the *gltS* gene as a template, to probe the membrane topology of the protein. The positions of the mutations were selected in putative loop regions. Uptake activities of the single-Cys mutants measured in RSO membrane vesicles were comparable to the wild-type protein, indicating that the mutations did not significantly affect the folding of the proteins in the membrane (not shown). The positions of the cysteine residues introduced in the protein were indicated in Figure 1 and correspond to the following mutants: H3C, S28C, N61C, R88C, S117C, D122C, S145C, T155C, S187C, A239C S276C, S301C, R327C, R363C, and G401C.

Accessibility of the Cysteine Residues in the Single-Cys Mutants of GltS from Either Side of the Membrane. Accessibility of the cysteine residues introduced into the GltS protein from the water phase was shown by sonicating cells expressing the mutants in the presence and absence of 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMdiS), a membrane-impermeable, negatively charged maleimide derivative. Following sonication and quenching of the excess of the AMdiS reagent, the GltS single-Cys mutants were purified from the membrane by Ni-NTA affinity chromatography. To identify labeling with AMdiS in the first step, purified proteins were treated with the fluorescent thiol reagent fluorescein-5-maleimide (FM). Labeling with FM was detected by fluorescence imaging of the gel after SDS-PAGE. The upper rows (FM) in Figure 3 show the fluorescence image of the gel, and the bottom rows (CBB) show the same part of the gel after staining with Coomassie brilliant blue. All mutants that were not treated with AMdiS showed a clear fluorescent band when treated with FM in detergent solution. In contrast, all mutants, except S117C and S145C, were not labeled with FM after treatment with AMdiS during the sonication step, demonstrating the accessibility of the thiol group to AmdiS when the mutants reside

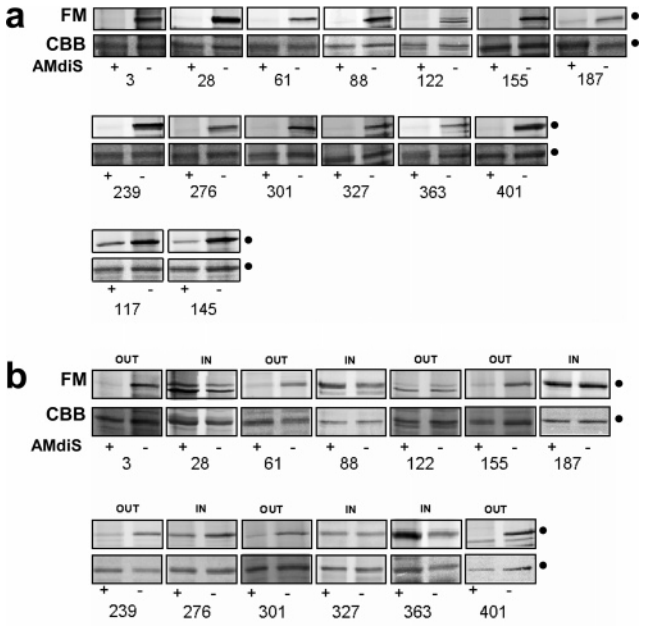

FIGURE 3: Accessibility of cysteine residues in single-Cys mutants of GltS by AMdiS. Single-Cys mutants were treated with (+) or without (−) 1 mM AMdiS for 30 min, followed by the addition of 1 mM DTT. (a) AMdiS was present during the breaking of the cells by sonication and 30 min thereafter. (b) The AMdiS treatment and quenching with DTT was done before breaking the cells. Following purification, the GltS proteins were treated with 0.1 mM FM for 5 min. Top rows labeled "FM" show the fluorescence image of the gel and those labeled "CBB" show the gel after Coomassie brilliant blue staining. Numbers indicate the position of cysteine residues in the corresponding single-Cys mutants. Dots mark the position of the GltS protein.

in the membrane. The protein bands in the Coomassie brilliant blue stained image show that the lack of fluorescence was not due to loss of the protein. Mutants S117C and S145C were only partly labeled with AMdiS in the first step, which suggests a restricted accessibility. These two mutants that reside both in the Vb region of the sequence were not considered further in this study.

Accessibility of the Cysteine Residues in Single-Cys Mutants of GltS from the Periplasmic Side of the Membrane. The sidedness of the cysteine residues introduced into the GltS protein was determined by treating whole cells with AMdiS before sonication. The AMdiS reagent cannot permeate the membrane and, therefore, can reach only residues located at the periplasmic side of the membrane. The labeling by AMdiS was identified as above by treating purified proteins with FM followed by fluorescence imaging. The results for mutants H3C, N61C, D122C, T155C, A239C, S301C, and G401C were very similar, as in the previous experiment. Treatment with AMdiS inhibited subsequent labeling of the purified protein by FM (Figure 3b, lanes +), indicating that AMdiS reacted at these sites in whole cells, indicating that these residues are located in the periplasm. Opposite results were obtained for mutants S28C, R88C, S187C, S276C, R327C, and R363C. The ratio of fluorescence intensity in AMdiS treated (+ lanes) and untreated (− lanes) samples was similar to the protein intensities of the bands in the Coomassie Brilliant Blue stained gel, indicating that AMdiS did not react with these sites in whole cells. Because the experiment above demonstrated that these sites are accessible from the water phase, it is concluded

that these residues are exposed to the cytoplasm. The two groups of mutants are presented as green and blue circles in the model in Figure 1. The results support the MemGen topology model. The Vb region in between TMSs IV and V and the Xa region in between TMSs IX and X are clearly located in the periplasm and cytoplasm, respectively.

Accessibility of Cytoplasmic Loop Xa from the Periplasm. Treatment of wild-type GltS, which contains four cysteine residues, three of which are located in transmembrane segments and one in the Xa region, with the membrane-permeable thiol reagent *N*-ethylmaleimide (NEM) did not affect glutamate uptake by RSO membranes (Figure 2, ● and ○). As expected, the activity of the Cys-less version of GltS was insensitive to the same treatment (not shown). More surprisingly, none of the single-Cys mutants was affected by NEM treatment. Apparently, labeling of the Cys residues did not affect the transport activity. It was shown before that labeling of two endogenous cysteine residues located in the pore-loop region Xa of CitS of *Klebsiella pneumoniae* in the 2HCT family resulted in decreased uptake activity, which emphasized the importance of this part of the sequence in catalysis (*20*) and, ultimately, resulted in the identification of the pore-loop structure (*15*). Three additional single-Cys mutants were constructed in GltS, yielding a total of five single-Cys mutants in the Xa region of the GltS protein: R327C, A339C, L347C, A355C, and R363C. Uptake activity measured in RSO membranes revealed a comparable level of activity as wild-type GltS. Treatment of the RSO vesicles with NEM, which is nonpolar and membrane-permeable, reduced uptake activity for the middle three single-Cys mutants (A339C, L347C, A355C) down to 30−40%, while the activity of the two flanking mutants (R327C and R363C) was not affected (Figure 4a). About 10−20% of the uptake activity in RSO vesicles is due to the chromosomal copy of wild type *gltS* gene, which is not sensitive to the thiol reagent (Figure 2). Increasing the NEM concentration and time of treatment did not result in lower activity of the single-Cys mutants, suggesting that the labeled molecules still have residual activity. The same observation was made for the endogenous cysteines in the Xa region of CitS of *Klebsiella pneumoniae* (*15*, *20*). Treatment of RSO membranes containing the five single-Cys mutants with membrane-impermeable AMdiS did not significantly decrease the uptake activity, which is in line with a cytoplasmic location of the Cys residues (Figure 4b).

The methanethiosulfonate (MTS) derivatives 2-(trimethylammonium)ethyl methanethiosulfonate bromide (MTSET) and sodium (2-sulfonatoethyl) methanethiosulfonate (MTSES) represent small thiol-reactive reagents that react with cysteine residues in proteins to form mixed disulfides (*21*). MTSET and MTSES differ in the charge of the groups attached to the reactive MTS moiety; MTSET is positively charged, and MTSES is negatively charged, which makes them membrane-impermeable. Treatment of RSO membranes with these reagents has been shown before to compromise the energy-generating system in the membranes to a certain level (*20*). Accordingly, treatment of RSO membranes containing wild-type and Cys-less GltS with MTSET and MTSES resulted in a decrease of 20−40% of the uptake activity (not shown). A similar reduction in uptake activity was observed for mutants R327C, L347C, R363C in the case of MTSES and for mutants R327C and R363C in the case of MTSET. However, treatment of mutants A339C and A355C with MTSES or MTSET resulted in a significantly lower residual activity (Figure 4c,d). The activities were down to the level observed after treatment with NEM. Treatment of mutant L347C with MTSET resulted in slightly less inactivation, while the same mutant was not affected by MTSES. The results demonstrate that residues in the Xa region in the cytoplasmic loop between TMS IX and X are accessible from the periplasmic side of the membrane for small, but not for more bulky membrane-impermeable reagents. The same observation was made for the endogenous Cys residues in the Xa region of CitS of *K. pneumoniae* in the 2HCT family (*15*).

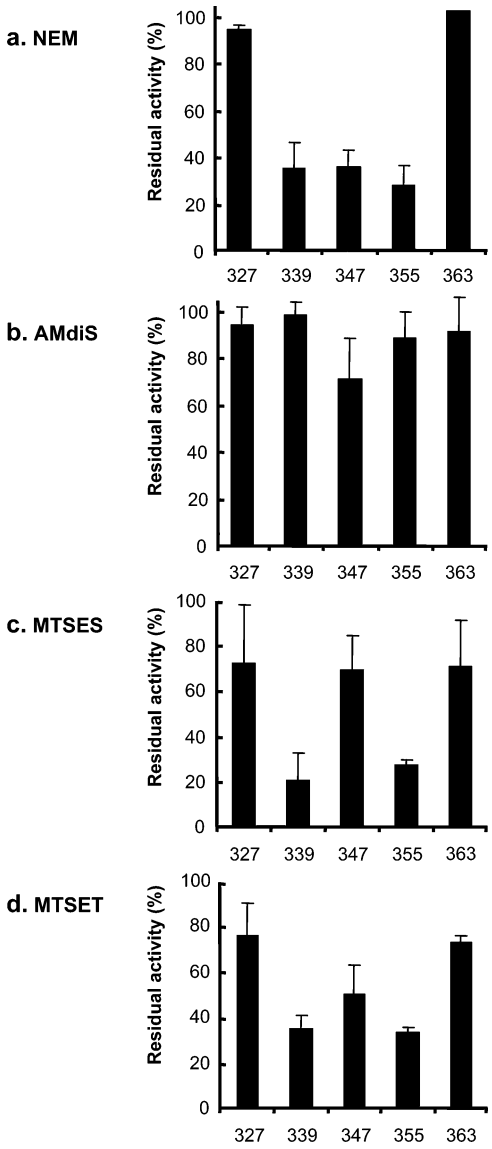

FIGURE 4: Inactivation of single-Cys mutants of GltS by various thiol reagents. RSO membrane vesicles containing single-Cys mutants R327C, A339C, L347C, A355C, and R363C were treated at room temperature with 1 mM NEM for 10 min (a), 0.25 mM AMdiS for 20 min (b), 10 mM MTSES for 20 min (c), and 1 mM MTSET for 10 min (d). Residual uptake activity was plotted as the percentage of the rate obtained with untreated membranes. Bars represent the average and standard deviation obtained from at least three independent measurements.

DISCUSSION

The glutamate transporter GltS of *E. coli* is the only characterized member of the ESS family of secondary transporters. The *gltS* gene product was shown be a L-glutamate transporter that also has affinity, albeit much lower, for D-glutamate, α-methylglutamate, and homocysteate (*13*, *22*−*24*). Glutamate transport was observed only in the presence of Na^+, irrespective of pH, suggesting an obligatory coupling of L-glutamate and Na^+ translocation. GltS is one of three L-glutamate transporters coded on the genome of *E. coli* K12. The [st324]ESS family in the MemGen classification consists of 76 unique members, all from bacterial origin, mostly from the γ subdivision of the Proteobacteria. The sequences in the family show a remarkably narrow size distribution, with most of the sequences consisting of 400−410 residues. Because no significant sequence similarity could be detected with members of other families, the ESS family is not in any of the known superfamilies in the transporter classification (TC) system (2.A.27 ESS) (*5*). In the MemGen classification, the ESS family is in structural class ST[3] together with 32 other families of membrane proteins.

The structural model for the members of the ESS family presented in Figure 1 is based on a different approach than models produced by predictors like TMHMM (*10*). While the latter rely on general features of membrane proteins, the MemGen method is basically homology modeling, porting information between families in the same structural class. Therefore, the MemGen method relies on the correct classification of families in structural classes. The MemGen method allows for the incorporation of structural details like domain structure and pore-loop structures, while TMHMM is restricted to loops and transmembrane segments. In a recent analysis in which the performance of transmembrane helix predictors was tested on a database of high-resolution structures of membrane proteins (*11*), hidden Markov model based approaches were shown to perform well for proteins consisting of transmembrane helices and loops, but less so when the protein contained structural elements like pore-loops. For most sequences in the ESS family, TMHMM predicts the putative pore-loop in the second domain (Xa, Figure 1) to be a transmembrane segment, while this is the case for about half of the putative pore-loops in the first domain (Vb). This results in models with 9−12 TMS for the different members. TMHMM predicts nine transmembrane segments for GltS of *E. coli*, three in the N-terminal half and six in the C-terminal half. The experimental data presented here is according the 10 TMS model predicted by the MemGen method. Cysteine residues placed in the N- and C-termini (H3C and G401C) of the GltS protein of *E. coli* were both localized in the periplasm by the accessibility of the bulky, membrane-impermeable thiol reagent AMdiS. The same was observed for five other cysteine residues in mutants N61C, D122C, T155C, A239C, and S301C, suggesting the presence of four periplasmic loops. The periplasmic location of both cysteine residues at positions 122 and 155 in region Vb indicates that the region is not transmembrane. Six cysteine residues in single-Cys mutants S28C, R88C, S187C, S276C, R327C, and R363C were not accessible for AMdiS in whole cells, while they were accessible from the water phase, which indicates a cytoplasmic location and five cytoplasmic loops. Again, the cytoplasmic location of the cysteine residues in R327C and R363C shows that the Xa region is not transmembrane.

Pore-loop structures are commonly found in channel proteins like the well- studied K^+ channels (*25*) and aquaporins (*26*), where they function as selectivity filters. The recently reported crystal structure of a glutamate transporter homologue of the archaeon *Pyrococcus horikoshi* revealed that they may be essential features in secondary transporters as well (*27*). The pore-loop structure in the Xa region of the transporters of the 2HCT family is based mostly on experimental studies of the Na^+-citrate transporter CitS of *K. pneumoniae* (*15*, *20*) and, to a lesser extent, the citrate/malate transporter CimH of *Bacillus subtilis* (*28*). Identification followed from the accessibility of sites in the cytoplasmic loop by small water-soluble thiol reagents from the periplasmic side of the membrane. The access pathway is believed to be the translocation pathway of substrate and co-ions through the protein (*8*, *15*). Using similar criteria, the present study identifies the pore-loop structure in the corresponding Xa region in the GltS transporter of the ESS family. Labeling of cysteine residues in the Xa region in three single-Cys mutants of GltS, A339C, L347C, and A355C with membrane-permeable NEM resulted in reduced transport activity. The same sites were not accessible to membrane-impermeable AMdiS when added at the periplasmic side of the membrane in RSO membranes, but two of the three, A339C and A355C, were clearly accessible for the small, membrane-impermeable MTSES and MTSET reagents in the same experimental system. The third mutant, L347C, that was inactivated by NEM was clearly less reactive with MTSES and MTSET, which suggests that not all positions of the Xa region are equally accessible to all reagents. No experimental evidence for the pore-loop in the N-terminal domain is available for any of the proteins in the 2HCT family or any other family in structural class ST-[3]. The pore-loop is based on weak homology between the N- and C-terminal halves of the proteins in ST[3] detected by bioinformatics tools (*18*). The pore-loop regions in both domains are characterized by the well-conserved sequence motif GGXG, which is believed to be at the vertex of the loops (*8*). The motif is present in the putative pore-loop sequence in the N-terminal domain of GltS as GGHG at positions 136−139.

The folding of the GltS protein in the membrane in 10 TMS and the identification of the pore-loop structure in between TMS IX and X strongly indicates a similar fold for the members of the ESS and 2HCT families. Apart from yielding a detailed structural model for the members of the ESS transporter family, the experimental data presented in this study supports the MemGen classification method, which is based on hydropathy profile alignment.

REFERENCES

1. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol. 157*, 105−132.
2. Lolkema, J. S., and Slotboom, D. J. (1998) Estimation of structural similarity of membrane proteins by hydropathy profile alignment, *Mol. Membr. Biol. 15*, 33−42.
3. Lolkema, J. S., and Slotboom, D. J. (2003) Classification of 29 families of secondary transport proteins into a single structural class using hydropathy profile analysis, *J. Mol. Biol. 327*, 901−909.

4. Lolkema, J. S., and Slotboom, D. J. (2005) Sequence and hydropathy profile analysis of two classes of secondary transporters, *Mol. Membr. Biol. 22*, 177−189.
5. Saier, M. H., Jr. (2000) A functional-phylogenetic classification system for transmembrane solute transporters, *Microbiol. Mol. Rev. 64*, 354−411.
6. Daley, D. O., Rapp, M., Granseth, E., Melén, K., Drew, D., and von Heijne, G. (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome, *Science 308*, 1321−1323.
7. Lolkema, J. S. (2006) Domain structure and pore-loops in the 2-hydroxycarboxylate transporter family, *J. Mol. Microbiol. Biotechnol. 11*, 318−325.
8. Sobczak, I., and Lolkema, J. S. (2005) The 2-hydroxycarboxylate transporter family: physiology, structure, and mechanism, *Microbiol. Molec. Biol. Rev. 69*, 665−695.
9. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search, *Nucleic Acids Res. 25*, 3389−3402.
10. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *J. Mol. Biol. 305*, 567−580.
11. Cuthbertson, J. M., Doyle, D. A., and Sansom, S. P. (2005) Transmembrane helix prediction: a comparative evaluation and analysis, *Prot. Eng. Des. Sel. 18*, 295−308.
12. Gaillard, I., Slotboom, D. J., Knol, J., Lolkema, J. S., and Konings, W. N. (1996) Purification and reconstitution of the glutamate carrier GltT of the thermophilic bacterium *Bacillus stearothermophilus*, *Biochemistry 35*, 6150−6156.
13. Tolner, B., Ubbink-Kok, T., Poolman, B., and Konings, W. N. (1995) Cation selectivity of the L-glutamate transporters of *Escherichia coli*, *Bacillus stearothermophilus*, and *Bacillus caldotenax*: dependence on the environment in which the proteins are expressed, *Mol. Microbiol. 18*, 123−133.
14. Kaback, H. R. (1974) Transport in isolated bacterial membrane vesicles, *Methods Enzymol. 31*, 698−709.
15. Sobczak, I., and Lolkema, J. S. (2004) Alternating access and a pore-loop structure in the Na^+-citrate transporter CitS of *Klebsiella pneumoniae*, *J. Biol. Chem. 279*, 31113−31120.
16. Konings, W. N., Barnes, E. N., Jr., and Kaback, H. R. (1971) Mechanisms of active transport in isolated membrane vesicles. 2. The coupling of reduced phenazine methosulfate to the concentrative uptake of β-galactosides and amino acids, *J. Biol. Chem. 246*, 5857−5861.
17. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
18. Lolkema, J. S., Sobczak, I., and Slotboom, D. J. (2005) Secondary transporters of the 2HCT family contain two homologous domains with inverted membrane topology and trans re-entrant loops, *FEBS J. 272*, 2334−2344.
19. Bandell, M., Ansanay, V., Rachidi, N., Dequin, S., and Lolkema, J. S. (1997) Membrane potential-generating malate (MleP) and citrate (CitP) transporters of lactic acid bacteria are homologous proteins. Substrate specificity of the 2-hydroxycarboxylate transporter family, *J. Biol. Chem. 272*, 18140−18146.
20. Sobczak, I., and Lolkema, J. S. (2003) Accessibility of cysteine residues in a cytoplasmic loop of CitS of *Klebsiella pneumoniae* is controlled by the catalytic state of the transporter, *Biochemistry 42*, 9789−9796.
21. Akabas, M. H., Stauffer, D. A., Xu, M., and Karlin, A. (1992) Acetylcholine receptor channel structure probed in cysteine-substitution mutants, *Science 258*, 307−10.
22. Kalman, M., Gentry, D. R., and Cashel, M. (1991) Characterization of the *Escherichia coli* K12 gltS glutamate permease gene, *Mol. Gen. Genet. 225*, 379−386.
23. Deguchi, Y., Yamato, I., and Anraku, Y. (1989) Molecular cloning of gltS and gltP, which encode glutamate carriers of *Escherichia coli* B, *J. Bacteriol. 171*, 1314−1319.
24. Deguchi, Y., Yamato, I., and Anraku, Y. (1990) Nucleotide sequence of gltS, the Na+/glutamate symport carrier gene of *Escherichia coli* B, *J. Biol. Chem. 265*, 21704−21708.
25. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbi, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998) The structure of the potassium channel: molecular basis of K+ conduction and selectivity, *Science 280*, 69−77.
26. Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J., and Stroud, R. M. (2000) Structural determinants of water permeation through aquaporin-1, *Science 290*, 481−486.
27. Yernool, D., Boudker, O., Jin, Y., and Gouaux, E. (2004) Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*, *Nature 431*, 811−818.
28. Krom, B. P., and Lolkema, J. S. (2003) Conserved residues R420 and Q428 in a cytoplasmic loop of the citrate/malate transporter CimH of *Bacillus subtilis* are accessible from the external face of the membrane, *Biochemistry 42*, 467−474.

BI062275I