Topology of RbsC, the Membrane Component of the *Escherichia coli* Ribose Transporter†

Jeffrey B. Stewart and Mark A. Hermodson*

Department of Biochemistry, Purdue University, West Lafayette, Indiana

Received 9 April 2003/Accepted 12 June 2003

The topology of RbsC, the membrane component of the ribose transporter in *Escherichia coli*, has been determined by using 34 single-cysteine mutants and a modified fluorescence labeling technique designated multiplex labeling. This technique gives topology, expression, and localization information for a membrane protein from a single batch of bacterial cells. The results indicate that RbsC contains 10 transmembrane-spanning helices, with the N and C termini being in the cytosol. This topology matches predictions from the latest prediction programs and the topology of the similar, recently crystallized membrane protein BtuC.

Over 300 families of membrane transporters have been identified and divided into four main classes based on their functions and energy coupling mechanisms (11). The first class contains channels and porins, which transport solutes via an energy-independent facilitated diffusion mechanism. Members of the second class, secondary transporters, catalyze uniport, antiport, or symport of solutes and utilize a carrier-mediated process such as the proton motive force to drive the reaction. Primary active transporters constitute a third class and use a primary energy source, such as ATP, for the uptake or extrusion of molecules. The fourth class has group translocators that phosphorylate their substrate during translocation.

A survey was recently conducted on 18 prokaryotic organisms, and it was found that a third of all membrane transport proteins are ATP-binding cassette (ABC) transporter proteins (9). These primary active transporters utilize the energy from the hydrolysis of ATP to move substrates into or out of the cell and are divided into 50 families. The ribose transport system in *Escherichia coli* involves an ABC protein complex classified in the second carbohydrate uptake transporter 2 (CUT2) family. Members of the CUT2 family differ from CUT1 protein complexes in that the two ABCs of the transport complex in the CUT2 proteins are joined in the same polypeptide chain. A functional ABC protein or protein complex commonly consists of two transmembrane domains and two cytosolic ABCs.

Mutations in ABC proteins can have a dramatic impact on an organism. For example, mutations in the ion channel cystic fibrosis transmembrane regulator were found to be responsible for cystic fibrosis (10). Altered expression of an ABC protein can lead to undesirable properties, such as multiple drug resistance (MDR) (2).

The ribose transport complex in *E. coli* is expressed from the *rbs* operon, encoding six proteins, RbsDACBKR (1, 16). RbsA is the cytosolic ATP-binding portion of the membrane complex, presumed to provide the energy for transport of ribose

into the cell. RbsC is the membrane-embedded portion of the membrane complex and possibly provides a channel through the membrane for ribose. RbsB is a ribose binding protein found in the periplasm. RbsK is a cytosolic ribokinase that phosphorylates ribose once it is in the cell, and RbsR is the repressor protein which regulates the operon. The function of RbsD has yet to be determined. The ribose transport system in *E. coli* is a multicomponent complex structurally similar to the single component cystic fibrosis transmembrane regulator and MDR proteins. The MDR homologue of this system would thus be the complex of RbsA and RbsC.

Although the specific mechanism for ribose membrane transport by RbsABC is not known, there is a general pathway whereby ribose enters the periplasm through pores in the outer membrane and is bound by RbsB, which then binds to the RbsAC complex. The energy released when ATP is bound and hydrolyzed by RbsA on the cytoplasmic side of the membrane is then used to bring ribose into the cell, where it is released for phosphorylation by RbsK.

Y. Park and C. Park recently published a proposed topology of RbsC using alkaline phosphatase fusions (8). The secreted protein alkaline phosphatase (PhoA) was used as a reporter for domain exposure to the periplasm since fusions resulting in localization of the PhoA portion to the periplasm have higher PhoA activity than fusions exposed to the cytosol (7). The topology proposed from these fusions has six transmembrane-spanning helices (TMs), with the N- and C-terminal ends being in the cytosol. In contrast, 10 TMs were found in the crystal structure of the membrane component of the *E. coli* vitamin B₁₂ transporter, BtuC (6), also an ABC transporter. A sequence alignment of RbsC and BtuC reveals 21% identity. While this level of identity is in the “gray area” for the determination of homology, it would not be surprising if there were also similarities in the proteins’ structures.

A topological model for RbsC can also be constructed by using transmembrane prediction programs. The dense alignment surface (DAS) method (3), which identifies 10 TMs (Table 1), was used to choose the initial model for the work reported in this paper. As seen in Table 1, TM 8 is predicted to be 6 amino acids long rather than the generally accepted length of 15 to 30 amino acids for a TM. This problem is likely due to the program’s use of scoring a target sequence based on

* Corresponding author. Mailing address: Department of Biochemistry, Purdue University, 175 S. University St., West Lafayette, IN 47907. Phone: (765) 494-1637. Fax: (765) 494-7897. E-mail: hermodson@purdue.edu.

† Journal paper no. 17113 of the Purdue Agricultural Experiment Station.

TABLE 1. Transmembrane predictions for RbsC

TM no.	Prediction program			
	DAS ^a		HMM ^b	
	Amino acids	Length	Amino acids	Length
1	23–36	14	24–43	20
2	58–72	15	48–67	20
3	74–93	20	74–93	20
4	95–118	24	98–117	20
5	127–140	14	124–143	20
6	176–187	12	170–190	21
7	222–236	15	221–240	20
8	259–264	6	249–268	20
9	277–288	12	275–292	18
10	304–316	13	297–316	20

^a The DAS method was from <http://www.sbc.su.se/~miklos/DAS/> and reference 3.
^b The HMM method was from <http://www.enzim.hu/hmmtop/html/submit.html> and reference 14.

sequences with known structures rather than on properties of the target sequence itself. A more advanced program using the hidden Markov model (HMM) to search for differences in the amino acid distributions in membrane and nonmembrane regions (14) also identifies 10 TMs (Table 1), but the domain lengths are more uniform and are similar in length and composition to the structures of other membrane proteins. Results from this program were used in the final model. However,

there are some older programs based only on hydrophobicity that propose from 6 to 12 TMs (data not shown). Experimental data to support topological models can be obtained by several approaches (for a review, see reference 15), and cysteine mutagenesis (5) is one technique that has been used successfully to support many models. A modified cysteine mutagenesis protocol designated multiplex labeling (13) was used on RbsC, and the results are reported here to provide experimental evidence to support a 10-TM model for RbsC.

MATERIALS AND METHODS

Cell lines. Standard DH5α and BL21(DE3) cell lines were used for plasmid preparation and protein expression, respectively. An *rbsC* deletion mutant cell line was used for testing activity (1).
Mutant preparation. Mutants were then prepared by using a Quikchange site-directed mutagenesis kit (Stratagene, La Jolla, Calif.) according to instructions included in the kit. Complementary primers designed according to the recommendations of the manufacturers were used along with a cysteineless *rbsC* template.
Activity. The *rbsC* mutants were assayed using a variant of the swarming and growth assay, which is referred to as the ring assay (4). An *rbsC* deletion mutant *E. coli* strain (1) was transformed with a plasmid carrying the mutant protein and grown overnight at 37°C. Cells were then spiked onto a tryptone agarose plate and allowed to grow at 37°C until a ring formed halfway to the edge of the plate. Cells from the edge of the ring were then transferred to a maltose minimal agarose plate and allowed to grow at 37°C until a ring formed halfway to the edge of the plate. Cells from the edge of the ring were then transferred to a ribose minimal agarose plate (0.1 mM ribose) and allowed to grow at 37°C until the wild-type RbsC ring formed halfway to the edge of the plate. The ring diameters of the other samples were then measured and compared to the ring diameters of

FIG. 1. Outline of multiplex labeling protocol. The letters *a* to *d* indicate the four different labeling methods used on the same batch of cells: in vivo (*a*), in vitro (*b*), vesicle (*c*), and SDS-vesicle (*d*).

FIG. 2. Multiplex labeling gel for the single-cysteine N157C, E19C, and V309C RbsC mutants. Shown is a fluorescence scan of a gel of three mutant RbsC proteins produced from plasmids as described in the text and processed according to the multiplex labeling protocol outlined in Fig. 1. The lowercase letters used in Fig. 1 indicate the four different labeling methods: in vivo (a), in vitro (b), vesicle (c), and SDS-vesicle (d). Shown are cysteineless His-tagged RbsC with the N157C replacement (lanes 2, 5, 8, and 11), cysteineless His-tagged RbsC with the E19C replacement (lanes 3, 6, 9, and 12), and cysteineless His-tagged RbsC with the V309C replacement (lanes 4, 7, 10, and 13). Purified, fluorescein-labeled His-RbsC is shown in lane 1 for reference.

cells carrying a plasmid encoding wild-type RbsC to give the activity as a percentage of that of wild-type RbsC.

Multiplex labeling. Single-cysteine variants of RbsC were labeled with fluorescein-5-maleimide (F5 M; Molecular Probes, Eugene, Oreg.) following a protocol called multiplex labeling (13), outlined in Fig. 1. Cells carrying the expression vector for RbsC variants were grown to the stationary phase and then induced. It was observed that protein production from the vector while the cells were in the stationary phase resulted in good production of RbsC while greatly

reducing the backgrounds of other proteins. Four 1-ml aliquots were removed after 3 h of induction. To the first aliquot, 12 μ l of 20 mM F5 M in dimethylformamide was added in the dark and allowed to react for 15 min at ambient temperature. The cells were quenched with 100 μ l of 100 mM glutathione and centrifuged at $4,000 \times g$ for 5 min to pellet the cells. The cells were washed to remove free label and then lysed with sodium dodecyl sulfate (SDS) (an equal amount of $2\times$ SDS-polyacrylamide gel electrophoresis loading buffer was added) and sonication (20 pulses at a 70% duty cycle). This sample was designated the

FIG. 3. Fluorescence intensities of the N157C, E19C, and V309C proteins on the SDS-PAGE gel described in the legend to Fig. 2. Intensity traces of the N157C and E19C protein lanes of the gel in Fig. 2 were created with the gel analysis program ImageQuant from Molecular Dynamics. Fluorescence intensity is plotted versus distance from an arbitrary point near the leading edge of the gel; thus, bands on the gel moved from right to left. RbsC is at about 340 pixels.

TABLE 2. Normalized fluorescence intensities of RbsC mutants

RbsC mutation	Normalized fluorescence intensity ^a	Avg intensity $\pm$ SD
T42C	25.2	21.7 ± 3.9
N45C	27.0	
N48C	24.2	
N146C	15.5	
N157C	26.0	
G167C	17.6	
E238C	22.8	
A245C	17.0	
G252C	20.9	
D256C	19.5	
S298C	23.3	
E19C	1.66	
V55C	1.14	
V95C	3.17	
A97C	2.80	
V124C	1.93	
H192C	1.10	1.98 ± 0.69
A201C	1.36	
N205C	2.65	
V216C	1.90	
K272C	2.36	
L292C	1.28	
N318C	2.38	
T69C	0.156	
I72C	0.167	
S78C	0.319	
T83C	0.136	
A109C	0.206	
I111C	0.087	
A120C	0.108	
None (wild type; C227)	0.247	0.17 ± 0.10
V262C	0.079	
I275C	0.384	
L279C	0.070	
V309C	0.082	

^a Band intensities were first quantified by using a program included with the PhosphorImager and then normalized by taking the ratio of the intensity to the intensity of an internal reference.

in vivo-labeled sample (*a* in Fig. 1). The second aliquot of cells was lysed with SDS and sonication (as described above) and frozen overnight. This sample was then labeled with F5 M as described for the in vivo sample and frozen. This sample was designated the in vitro-labeled sample (*b* in Fig. 1). The third aliquot of cells was sonicated to produce vesicles and then centrifuged at 4,000 $\times g$ for 40 min. The supernatant was transferred to another vial and labeled with F5 M as described for the in vivo sample, except that twice the amount of F5 M was used. This sample was then centrifuged at 100,000 $\times g$ for 10 min, and the supernatant was discarded. The pelleted membranes were solubilized with SDS (as described above) and frozen. The sample was designated the vesicle-labeled sample (*c* in Fig. 1). The fourth aliquot was processed in a way similar to that used for the third aliquot, except that the labeling by F5 M was performed after the addition of SDS. This sample was designated the SDS-vesicle-labeled sample (*d* in Fig. 1).

Finally, all four samples were sonicated again to lyse the membranes completely and solubilize all the proteins, and the samples were then loaded onto an SDS-PAGE gel. After electrophoresis, the gel was scanned on a PhosphorImager in the fluorescence mode with standard fluorescein settings. A fluorescence-scanned gel for cells expressing the single-cysteine N157C, E19C, and V309C RbsC mutants is shown in Fig. 2.

Fluorescence. Gels were scanned while they were still in the glass plates on a Typhoon 8600 PhosphorImager (Molecular Dynamics) in fluorescence mode.

The excitation wavelength was 532 nm with a green laser, and a 526-nm-wavelength filter was employed for detection.

RESULTS AND DISCUSSION

Multiplex labeling is similar to a method described previously (5) in that it uses SDS-PAGE gels of fluorescein-labeled single-cysteine mutants to propose membrane protein topology. The advantages of multiplex labeling include the ability to collect data from four separate experiments with the same batch of cells. The first experiment with multiplex labeling resulted from the discovery that we could selectively label periplasmic cysteines of membrane proteins in *E. coli* cells immediately after their removal from a shaker. The periplasm is an oxidizing environment, so most native proteins will have oxidized cysteines, resulting in a very low background. Further experiments were added to the protocol to show the total expression of each variant (in vitro labeling), to determine which variants had cysteine residues exposed to either the periplasm or the cytoplasm (vesicle labeling), and to verify the plasma membrane location of the RbsC variants (SDS-vesicle labeling).

The first step in the basic cysteine mutagenesis technique was to create an active, cysteineless variant of the target protein, RbsC. To facilitate production and future purification of mutant proteins, the *rbsC* gene had previously been cloned into a composite plasmid based on pBR322 with a T7 promoter and an N-terminal extension encoding a His tag of 10 consecutive histidines and a factor X_a cleavage site (MGH₁₀SSGHIEGR ↓ H[RbsC]) (16). The resulting plasmid was designated pH14. The single cysteine at position number 227 in wild-type RbsC was then mutated to an alanine. This plasmid encoding the cysteineless RbsC was designated pHCa and was used as a template for the preparation of the remaining 34 mutants.

The activity of the pHCa mutant in an *rbsC* mutant cell line was 80% of that of the wild-type protein when it was assayed with a swarming and growth assay (4). This assay has been used in various forms for many years to test the ability of *E. coli* strains to grow on minimal media. Some cells are spotted onto the center of a soft agar minimal medium plate and incubated. If the cells are motile with an intact chemotaxis system and can grow in the minimal media, they will grow outward, forming a white ring. The diameter of this ring serves as a measure of cell viability and motility. For RbsC the assay is used to see how well mutant versions of the protein can complement an *rbsC* *E. coli* deletion mutant strain that is deficient in the ability to utilize 0.1 mM ribose as a carbon source (1). Out of the 34 mutant proteins that were prepared, only the N45C, N48C, D256C, V262C, and L292C proteins were inactive. These inactive mutant proteins were also labeled and used for topology determination since the multiplex labeling results showed that they were expressed and transported to the membrane properly.

Multiplex labeling of stationary-phase cells expressing single-cysteine mutants of RbsC yielded four types of samples (*a*, *b*, *c*, and *d* in Fig. 1) with different fluorescent band intensities on SDS-PAGE gels. The first type of sample, designated in vivo labeling, gave topology information based on how fluorescent band intensity varied with the position of the cysteine

FIG. 4. Topological model of RbsC based on multiplex labeling data and the HMM-based TM prediction program.

relative to the membrane. Mutant RbsC proteins with a cysteine embedded in the membrane did not label (Fig. 2, lane 4), whereas mutants with a cysteine in the cytosol or near the membrane edge labeled at about the same intensity as that of background labeling (Fig. 2, lane 3). In contrast, if the cysteine was in the periplasm, the band was at least 20 times more intense (Fig. 2, lane 2; see Fig. 3 for a comparison of the intensities of lanes 2 to 4). The differences among band intensities were thus classified into three categories for model construction: strong, weak, and not present.

The second type of sample labeled all cysteines in the cell, which gave a measure of protein expression (Fig. 2, lanes 5 to 7). The third and fourth types of samples contained vesicles (Fig. 1, *c* and *d*). As in the *in vivo* labeling experiment, cysteines embedded in the membrane were not labeled in vesicles (Fig. 2, lane 10) but were labeled once the vesicles were disrupted (Fig. 2, lane 13). Cysteines present in the periplasm or cytosol were equally labeled in the vesicle experiment due to the randomization of vesicle orientation occurring during sonication (12) (Fig. 2, lanes 8 and 9). There was also some variation in intensity in the vesicle labeling bands resulting from the position of the cysteine relative to the membrane, but the variations were not as dramatic as in the *in vivo* labeling experiment. These variations are included in the model as strong, weak, and not present.

Of the 34 single-cysteine mutant proteins that were made, 12 did not label in the *in vivo* or vesicle experiments, but they were expressed and transported to the membrane as seen in the *in vitro* and SDS-vesicle experiments. Thus, these mutants have cysteines embedded in their membranes. Eleven of the mutant proteins were strongly labeled in the *in vivo* experiments. Thus, these proteins have cysteines in the periplasm. The remaining

mutant proteins were weakly labeled in the *in vivo* experiments, suggesting that they have cysteines in their cytosol. However, two of these proteins were also weakly labeled in the vesicle experiment (the V55C and A97C mutant proteins). This finding suggests that these two mutant proteins have in their periplasm cysteines which have reduced access to the label. For instance, the cysteines may be near the edge of the membrane. After many topologies were tried, the only one that best fit all the data (*in vivo* data are summarized in Table 2) was the 10-TM model seen in Fig. 4. This figure also has the calculated data from the transmembrane prediction program HMM (14) (results shown in Table 1 and as red-colored amino acids in Fig. 4).

The mutants were chosen based on an initial model predicted by the DAS TM prediction program (3) (results shown in Table 1), and mutations of amino acids in each of the predicted loops were made on both sides of the membrane. As more mutants were made and tested with the multiplex labeling protocol, it became clear that the DAS model was incorrect, and a new model based on HMM was used instead. As seen by the red-colored amino acids in Fig. 4, this model fits the data very well. However, the T42C, E238C, G252C, D256C, and S298C mutant proteins were predicted by the HMM program to be in the membrane, but they labeled strongly, suggesting that they are in the periplasm instead. Since these mutants were predicted to be at the periplasmic ends of the TMs, it is perhaps not surprising that they have easy access to the label, and they may end up actually being within the membrane but accessible to the periplasm.

The similarity of the prediction, labeling data, and the published crystal structure of the related protein BtuC suggests that the actual topology of RbsC has 10 rather than 6 TMs as

published previously (8). The gene fusion technique used to produce the six-TM model produces hybrid proteins, likely giving rise to misleading data. Multiplex labeling, on the other hand, required only single-cysteine replacement of a cysteine-less template. The fact that labeling was done on cells in growth media probably also led to more reliable and consistent data.

Even though there is a 21% sequence identity between RbsC and BtuC and both proteins are ABC transporters, the identity is uniformly distributed throughout the sequences. It was not readily apparent how to map the RbsC sequence onto the BtuC structure.

The clustering of the inactive mutants on the periplasmic side of the membrane suggests that there is a close interaction between the binding proteins RbsB and RbsC and that RbsC may function as more than just a channel for ribose, since the replacement of a single amino acid with a cysteine does not seem to be a significant enough change to block a channel.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of PHS grants GM 54993 and T32GM08296. We also acknowledge the support of the Purdue Cancer Center Core Grant NCI CCSG.

REFERENCES

1. Barroga, C. F., H. Zhang, N. Wajih, J. H. Bouyer, and M. A. Hermodson. 1996. The proteins encoded by the *rbs* operon of *Escherichia coli*: I. Overproduction, purification, characterization, and functional analysis of RbsA. *Protein Sci.* **5**:1093–1099.
2. Chen, C., J. E. Chin, K. Ueda, D. P. Clark, I. Pastan, M. M. Gottesman, and I. B. Roninson. 1986. Internal duplication and homology with bacterial transport proteins in the *mdr1* (P-glycoprotein) gene from multidrug-resistant human cells. *Cell* **47**:381–389.
3. Cserző M., E. Wallin, I. Simon, G. von Heijne, and A. Elofsson. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng.* **10**:673–676.
4. Hazelbauer, G. L., C. Park, and D. M. Nowlin. 1989. Adaptational “crosstalk” and the crucial role of methylation in chemotactic migration by *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86**:1448–1452.
5. Jones, P. C., A. Sivaprasadarao, D. Wray, and J. B. Findlay. 1996. A method for determining transmembrane protein structure. *Mol. Membr. Biol.* **13**:53–60.
6. Locher, K. P., A. T. Lee, and D. C. Rees. 2002. The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* **296**:1091–1098.
7. Manoil, C., and J. Beckwith. 1986. A genetic approach to analyzing membrane protein topology. *Science* **233**:1403–1408.
8. Park, Y., and C. Park. 1999. Topology of RbsC, a membrane component of the ribose transporter, belonging to the AraH superfamily. *J. Bacteriol.* **181**:1039–1042.
9. Paulsen, I. T., L. Nguyen, M. K. Sliwinski, R. Rabus, and M. H. Saier, Jr. 2000. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J. Mol. Biol.* **301**:75–100.
10. Riordan, J. R., J. M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, and J. L. Chou. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**:1066–1073.
11. Saier, M. H., Jr. 2000. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol. Mol. Biol. Rev.* **64**:354–411.
12. Short, S. A., and R. Kaback. 1975. Localization of D-lactate dehydrogenase in native and reconstituted *Escherichia coli* membrane vesicles. *J. Biol. Chem.* **250**:4291–4296.
13. Stewart, J. B. 2003. Multiplex labeling: a new method for determination of membrane protein topology. Ph.D. thesis. Purdue University, West Lafayette, Ind.
14. Tusnády, G. E., and I. Simon. 1998. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol.* **283**:489–506.
15. van Geest, M., and J. S. Lolkema. 2000. Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol. Mol. Biol. Rev.* **64**:13–33.
16. Zaitseva, J., H. Zhang, R. A. Binnie, and M. Hermodson. 1996. The proteins encoded by the *rbs* operon of *Escherichia coli*: II. Use of chimeric protein constructs to isolate and characterize RbsC. *Protein Sci.* **5**:1100–1107.