

Membrane topology of the Mep/Amt family of ammonium transporters

Gavin H. Thomas,¹ Jonathan G. L. Mullins² and Mike Merrick^{1*}

¹Department of Molecular Microbiology, John Innes Centre, Colney Lane, Norwich, Norfolk NR4 7UH, UK.

²Department of Biology and Health Sciences, Faculty of Science, Technology and Design, University of Luton, Luton, Beds LU1 3JU, UK.

Summary

The Mep/Amt proteins constitute a new family of transport proteins that are ubiquitous in nature. Members from bacteria, yeast and plants have been identified experimentally as high-affinity ammonium transporters. We have determined the topology of AmtB, a Mep/Amt protein from *Escherichia coli*, as a representative protein for the complete family. This was established using a minimal set of AmtB–PhoA fusion proteins with a complementary set of AmtB–LacZ fusions. These data, accompanied by an *in silico* analysis, indicate that the majority of the Mep/Amt proteins contain 11 membrane-spanning helices, with the N-terminus on the exterior face of the membrane and the C-terminus on the interior. A small subset, including *E. coli* AmtB, probably have an additional twelfth membrane-spanning region at the N-terminus. Addition of PhoA or LacZ α -peptide to the C-terminus of *E. coli* AmtB resulted in complete loss of transport activity, as judged by measurements of [¹⁴C]-methylammonium uptake. This C-terminal region, along with four membrane-spanning helices, contains multiple residues that are conserved within the Mep/Amt protein family. Structural modelling of the *E. coli* AmtB protein suggests a number of secondary structural features that might contribute to function, including a putative ammonium binding site on the periplasmic face of the membrane at residue Asp-182. The implications of these results are discussed in relation to the structure and function of the related human Rhesus proteins.

Introduction

Movement of ammonium across biological membranes is an important process in many organisms, including humans, plants and bacteria (Knepper *et al.*, 1989; Kleiner, 1993). The identification of high-affinity ammonium transport proteins in *Saccharomyces cerevisiae* and *Arabidopsis thaliana* has stimulated renewed interest in this process (Marini *et al.*, 1994; Ninnemann *et al.*, 1994). These proteins, which have been named Mep/Amt, form a new family of transporters with over 30 currently identified members from all three domains of life (Saier *et al.*, 1999) (Fig. 1). They also exhibit significant sequence similarity to the Rhesus proteins of human erythrocytes (Marini *et al.*, 1997). The ability of the Mep/Amt proteins to transport the ammonium analogue [¹⁴C]-methylammonium has been demonstrated in a number of organisms in addition to *S. cerevisiae* and *A. thaliana*. These include *Azorhizobium caulinodans*, *Azospirillum brasilense*, *Azotobacter vinelandii*, *Corynebacterium glutamicum*, *Escherichia coli*, *Klebsiella pneumoniae*, *Rhizobium etli* and *Synechocystis* sp. PCC 6803 (Siewe *et al.*, 1996; van Dommelen *et al.*, 1998; Michel-Reydellet *et al.*, 1998; Meletzus *et al.*, 1998; Montesinos *et al.*, 1998; Soupene *et al.*, 1998; Taté *et al.*, 1998; Jack *et al.*, 1999). Despite being widely characterized genetically and functionally, no structural analysis has yet been reported for this new family of transport proteins.

Current members of the Mep/Amt family vary in length from around 400–450 amino acids, although some members have extended C-terminal regions increasing their length to about 600 amino acids. They have been predicted to be integral membrane proteins of around 45–50 kDa containing between 10 and 12 transmembrane (TM) helices (Siewe *et al.*, 1996; Montesinos *et al.*, 1998). In order to determine the secondary structure of this family of transporters, we have chosen to investigate the topology of the *E. coli* AmtB protein. This is the only Mep/Amt family member in this organism and is essential for [¹⁴C]-methylammonium transport (Soupene *et al.*, 1998). It is encoded by the *amtB* gene, which is the second gene in the *glnKamtB* operon (van Heeswijk *et al.*, 1996). As expected for a proposed high-affinity ammonium transporter, it is only expressed when the amount of ammonium in the growth medium is limiting (Atkinson and Ninfa, 1998). This control is dependent on the nitrogen regulatory (Ntr) system, which induces

Accepted 11 May, 2000. *For correspondence. E-mail mike.merrick@bbsrc.ac.uk; Tel. (+44) 1603 50749, Ext. 2749; Fax (+44) 1603 450018.

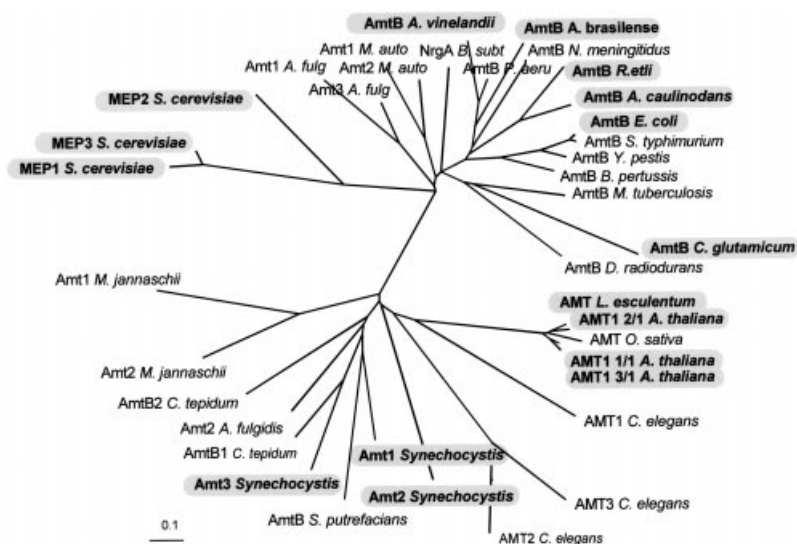


Fig. 1 Phylogenetic analysis of the Mep/Amt family of ammonium transporters. Protein sequences were truncated at their N- and C-termini to remove highly variable extensions of the proteins. The program PHYLIP was used to produce a distance matrix tree, which was displayed in TREEVIEW. Sequences that have been experimentally determined to transport [^{14}C]-methylammonium are boxed. *M. auto*, *Methanobacterium thermoautotrophicum*; *A. fulg*, *Archaeoglobus fulgidus*; *B. subt*, *Bacillus subtilis*, *P. aero*, *Pseudomonas aeruginosa*.

expression of the *glnKamtB* operon approximately 2000-fold during nitrogen-limited growth (Atkinson and Ninfa, 1998). This regulatory pattern is also consistent with previous studies examining the regulation of [^{14}C]-methylammonium transport in *E. coli* (Servín-González and Bastarrachea, 1984; Jayakumar *et al.*, 1986). The gene preceding *amtB* encodes the GlnK protein, which has recently been identified as an additional component of the Ntr system in *E. coli* (van Heeswijk *et al.*, 1996; Atkinson and Ninfa, 1998; 1999). The co-expression of a *glnK*-like gene with an *amtB* gene is a conserved feature of many eubacterial and archaeal genomes. The identification of conserved gene pairs in microbial genomes indicates that their products are likely to interact physically (Dandekar *et al.*, 1998). Therefore, a physical interaction between AmtB and GlnK might be expected, although this has not yet been confirmed experimentally (Thomas *et al.*, 2000).

The 44.5 kDa *E. coli* AmtB protein contains 428 amino acid residues and has been predicted to have 12 TM

helices (van Heeswijk *et al.*, 1996). In this study, we determine the membrane topology of this protein and, from this, deduce a general topology model for the Mep/Amt protein family. We also investigate the effects on transport activity of fusing protein domains to the C-terminus of AmtB and use sequence analysis to identify regions of the protein that might be involved in the transport process.

Results

Hydropathy analysis of the Mep/Amt family

A collection of Mep/Amt sequences was obtained from published sequences, sequences in EMBL/GenBank/DBJ and those in uncompleted microbial genome sequences (see http://alife.ulb.ac.be/~bandre/Mep_List.htm). These include phylogenetically distinct family members from *Caenorhabditis elegans*, plants, yeast, Eubacteria and Archaea (Fig. 1). The structural conservation between these proteins was investigated by analysing their hydropathy profiles to determine whether a global structure could be assigned for the complete family (Lolkema and Slotboom, 1998). Average hydropathy analysis of 23 members of the Mep/Amt family demonstrated a conserved structure of 11 hydrophobic regions (Fig. 2) with a low structure divergence score of 0.116. A small subset of the family, including *E. coli* AmtB, contains proteins with an additional hydrophobic segment at the N-terminus that might constitute a 12th membrane-spanning helix. Gaps in the alignment are almost always found in regions of the proteins that are predicted to be hydrophilic loops interconnecting the helices (Fig. 2).

To investigate the structure of the Mep/Amt proteins, we chose to use the AmtB protein from the bacterium

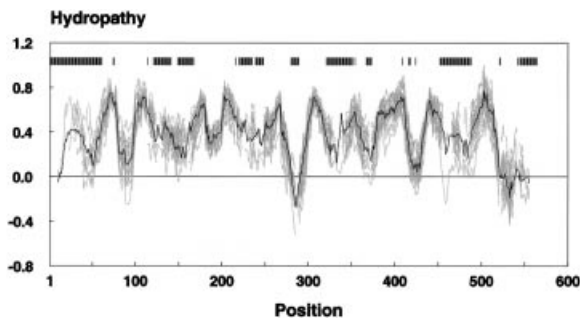


Fig. 2 Average hydropathy profile of the Mep/Amt proteins. The bold line represents the family hydropathy profile, whereas the thin lines are the individual profiles for the selected family members. Vertical bars at the top of the profile indicate the positions of gaps in the multiple sequence analysis.

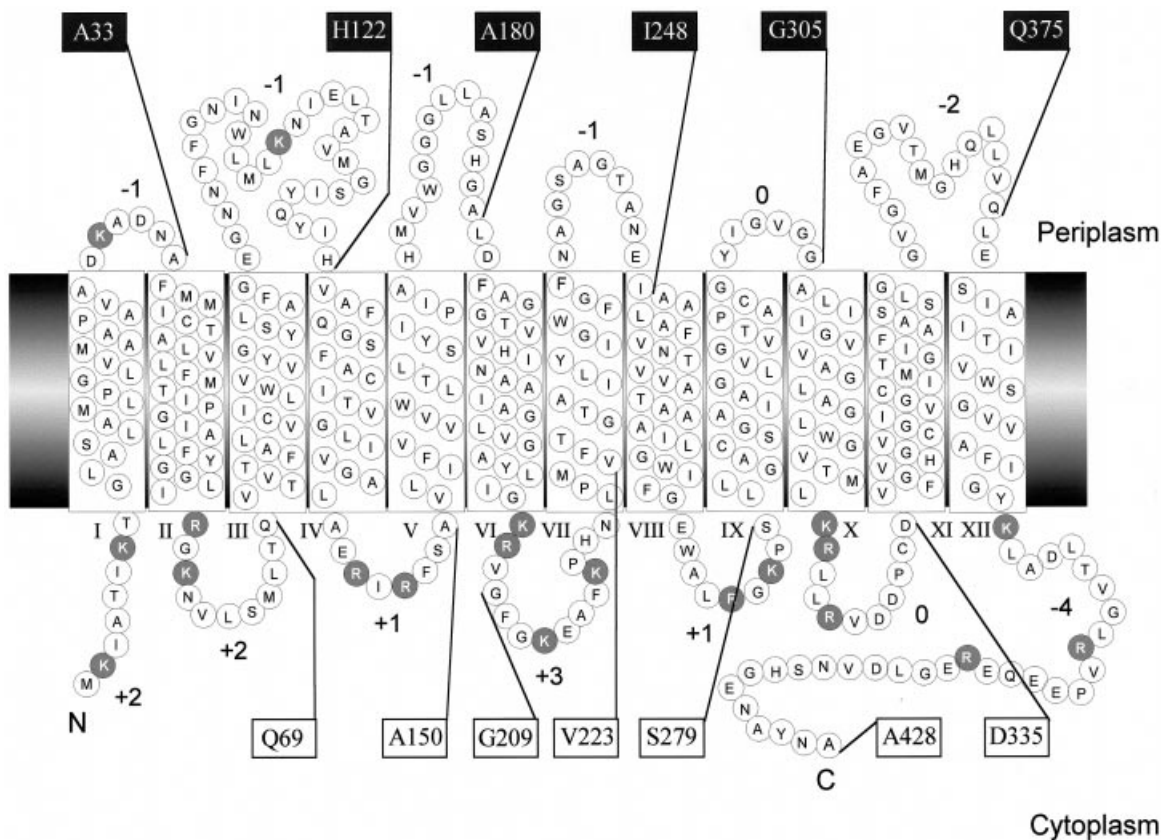


Fig. 3. Proposed topological model of the *E. coli* AmtB protein. Positively charged residues are shown in black, and the overall charge of each loop is indicated. Black boxes indicate positions of AmtB–PhoA fusions that were blue on indicator plates (between 1400 and 7160 AP units), whereas white boxes correspond to fusions that were white on indicator plates (between 20 and 120 units). The numbers in the boxes are the positions of each fusion within AmtB.

E. coli, as it is the only member of the Mep/Amt family in this bacterium. Also, *E. coli* is very amenable to genetic methods of topology determination.

Generation of a topological model for the *E. coli* AmtB protein

The sequence of *E. coli* AmtB was analysed using topology prediction programs that make assignments based on a combination of hydrophathy and topogenic signals located within the primary amino acid sequence. We used MEMSAT (Jones *et al.*, 1994), TMPRED (Hofmann and Stoffel, 1993), TMAP (Persson and Argos, 1997), TOPPRED2 (von Heijne, 1992), HMMTOP (Tusnady and Simon, 1998) and TMHMM (Sonnhammer *et al.*, 1998), of which the last two methods use recently developed hidden Markov models (HMM) for predicting topology. The outputs of MEMSAT, TOPPRED2, HMMTOP and TMHMM gave similar results; they all predict 12 TM helices and an orientation with both the N- and C-termini in the cytoplasm ($N_{IN}C_{IN}$). This topological model is presented in Fig. 3 and uses the helix ends predicted by MEMSAT. The other two

programs, TMPRED and TMAP, do not predict helix IX within the protein and, consequently, predict an $N_{IN}C_{OUT}$ topology. The positively charged residues in the *E. coli* AmtB protein are also illustrated in Fig. 3. After helix IV, these are found exclusively in regions of the protein that are predicted to be cytoplasmic loop regions. The absence of helix IX, as predicted by TMPRED and TMAP, would reverse this distribution along the rest of the protein, so we favoured the 12 TM helix $N_{IN}C_{IN}$ alternative as an initial model. The majority of the putative TM helices are linked by short loop regions. Application of the algorithms of Deleage and Roux (1987) and Chou and Fasman (1979) predicts that most of these short loops have a high propensity to form β -turns, a common feature of linkers in membrane proteins (data not shown).

Testing the model using PhoA fusions

To determine experimentally the topology of the *E. coli* AmtB protein in the cytoplasmic membrane, we used a genetic approach of constructing gene fusions between *amtB* and *phoA*, which encodes the location-specific

reporter protein alkaline phosphatase (AP) (Traxler *et al.*, 1993). Alkaline phosphatase characteristically exhibits high activity when located in the periplasm. As the topology modelling generally favoured a single model with a maximum of 12 helices, we used the approach of Boyd *et al.* (1993), in which a series of fusions are constructed to residues around the C-termini of each predicted hydrophilic region of the protein. The sites of these fusion junctions are illustrated in Fig. 3. We also constructed an additional fusion at the start of the loop between helices VI and VII (residue Gly-209) to determine whether the properties of this fusion differed significantly from a fusion at the start of TM VII (residue Val-223).

Thirteen specific polymerase chain reaction (PCR) products containing regions of *amtB* were cloned in frame into the *phoA* vector pSK4158, and transformed into strain ET8000 (Paulsen *et al.*, 1995). The activity of each fusion protein was initially assessed on indicator plates. Bacteria containing AmtB-PhoA fusions displayed an alternating blue and white colour throughout the full length of AmtB (Table 1).

AP activities were determined for all AmtB-PhoA fusion proteins, and the alternating pattern of blue and white colonies observed on indicator plates was reflected in the calculated PhoA activities (Fig. 3 and Table 1). Colonies that were blue on indicator plates give AP activities between 1400 and 7160 units of activity, while those that resulted in white colonies had activities between 20 and 120 units. There is a clear separation of these two classes of fusion, over 10-fold in all cases, which can be seen in Fig. 3 overlaid on the topology model. The high-activity fusions all have fusion junctions that are predicted to be periplasmic, whereas all the low-activity fusions have fusion junctions that are predicted to be cytoplasmic.

To complement this data, we also used Western blotting with anti-AP antibody to distinguish between fusions with a periplasmic or cytoplasmic location. This method is based on the observation that the amount of immunoreactive products on a Western blot is much

Table 1. AP activities and β -galactosidase activities determined for the AmtB-PhoA/LacZ fusion proteins.

Position of fusion in AmtB	Colour on M9Gln XP indicator plates	PhoA activity (Miller units ^a)	LacZ activity (Miller units)
No fusion	White	10 (\pm 1)	20 (\pm 10)
Ala-33	Blue	7160 (\pm 1390)	2480 (\pm 925)
Gln-69	White	26 (\pm 4)	1290 (\pm 255)
His-122	Blue	3070 (\pm 290)	740 (\pm 130)
Ala-150	White	48 (\pm 5)	2920 (\pm 290)
Ala-180	Blue	2290 (\pm 380)	25 (\pm 5)
Gly-209	White	105 (\pm 40)	ND
Val-223	White	120 (\pm 5)	2730 (\pm 215)
Ile-248	Blue	2770 (\pm 665)	545 (\pm 240)
Ser-279	White	30 (\pm 5)	2240 (\pm 440)
Gly-305	Blue	1400 (\pm 420)	1950 (\pm 490)
Asp-335	White	20 (\pm 8)	990 (\pm 305)
Gln-375	Blue	3200 (\pm 750)	390 (\pm 215)
Ala-428	White	50 (\pm 15)	30 (\pm 10)

All assays were performed in the *E. coli* strain ET8000 and are averages from three independent duplicate experiments. All values are rounded to three significant figures.

a. As defined by Manoil (1991).

greater for periplasmic fusions than for cytoplasmic fusions. This results from the greater stability of the active periplasmically located PhoA, either as fusion protein or as cleaved PhoA, over the inactive cytoplasmically located PhoA, which is misfolded and rapidly degraded (Geest and Lolkema, 2000). Again, the same pattern was seen as with the AP activities (Fig. 4A) with the exception of the full-length fusion (A428), which accumulates to a much greater level than the other cytoplasmically located fusions. Despite this fusion protein being more stable, perhaps because of the presence of the complete AmtB protein, the corresponding AP activity is low, suggesting that the attached PhoA moiety is located in the cytoplasm. The Western blot also illustrates that a band of the correct size for each fusion protein can be identified, demonstrating that all fusions are being expressed. The shortest fusion accumulated to a much greater degree than the other fusions, to the extent that it could be visualized on

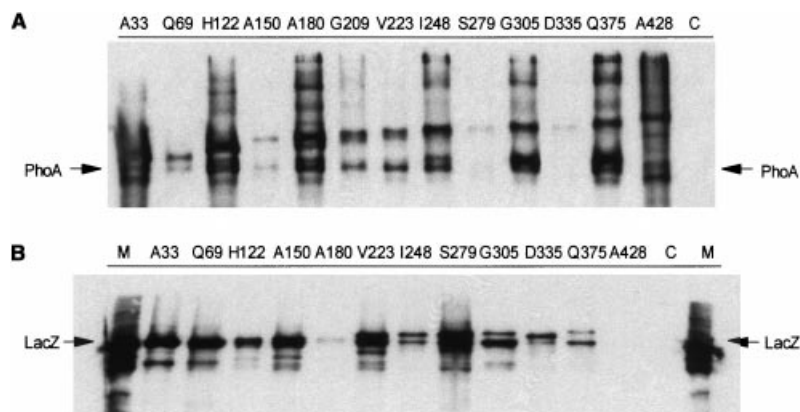


Fig. 4. Western blots of whole-cell extracts from ET8000 transformed with either (A) *amtB-phoA* fusions or (B) *amtB-lacZ* fusions. The positions of the fusion junction are indicated above each lane. On both membranes, the lane labelled C contained an extract from untransformed ET8000. Arrows indicate the positions of free PhoA and LacZ.

Coomassie blue-stained whole-cell extracts that had been separated by SDS-PAGE. This phenomenon has been seen in other studies and is thought to be related to differential mRNA stability during the synthesis of each fusion protein (Sarsero and Pittard, 1995).

The unambiguous nature of the AP fusion data clearly agrees with the predicted topology. It also eliminates the alternative structure suggested by the topology prediction programs TMPRED and TMAP.

Testing the model using LacZ fusions

Although the PhoA fusion data confirmed our *in silico* topology model, we felt that a complementary data set using LacZ fusions would be desirable. We first used a system that had been used previously to complement the pSK4158-derived *phoA* fusion data (Paulson *et al.*, 1995). This involved cloning the same set of PCR products used for the *phoA* study into the *lacZ α* gene of pUC118. These were then transformed into DH5 α and assayed for β -galactosidase. The resulting fusions at positions A428 and D335 were blue on indicator plates, whereas all other fusions were white. The amount of β -galactosidase activity in the fusions at positions G305, D335, Q375 and A428 was determined and, despite the D335 and A428 fusions giving positive results on indicator plates, no activity could be detected. This weak complementation of strain DH5 α with the AmtB-LacZ α -peptide fusions may be the result of the location of the α -peptide in proximity to the membrane, where it is physically constrained from interacting with the rest of LacZ.

Therefore, we cloned the same set of PCR products into plasmid pNM480, which generates in frame fusions to full-length LacZ (Minton, 1984). When the resulting series of plasmids was transformed into ET8000 and grown on indicator plates, all fusions yielded blue colonies. The β -galactosidase activities were quantified in ET8000 (Table 1), and Western blotting was used to assess the amount of LacZ-immunoreactive polypeptides that accumulated in these cells. The presence of a number of these fusion proteins can be seen in Fig. 4B and, in each case, there is an additional band that migrates at the same position as β -galactosidase.

These engineered AmtB-LacZ fusions gave activities that were complementary to the PhoA fusions, except for the fusions at Ala-33, Gly-305 and Ala-428, which required further interpretation. The fusion at Ala-33 gives high levels of both AP and β -galactosidase activity; however, more weight is put on the AP data, as the PhoA moiety requires active translocation across the membrane to give activity (Sarsero and Pittard, 1995; Geest and Lolkema, 2000). The fusion at Gly-305 also yields high AP and high β -galactosidase activity; however, examination of the Western blot (Fig. 4B) reveals that the predominant

LacZ-containing polypeptide in this strain is free LacZ. This suggests that most of the fusion protein is being degraded and that the majority of the β -galactosidase activity in this sample can be ascribed to the degradation product. The fusion at the C-terminus of AmtB (Ala-428) gave no detectable signal, which agrees with the β -galactosidase assay data, suggesting that, for some reason, this construct is not expressing a fusion protein.

We also attempted to generate random *amtB-lacZ* fusions using infection and selection with the transposon Tn*lacZ*n (Manoil and Bailey, 1997). After eight independent infection and selection procedures, we located seven different fusion positions within the *amtB* gene that had detectable β -galactosidase activity. DNA sequencing across the fusion junctions revealed that only three of these were in frame with *amtB*. However, when products of these three active in frame fusions were analysed by Western blotting of whole-cell extracts with anti-LacZ antibody, bands were only identified that corresponded to free LacZ (data not shown). This indicated that the fusion junctions selected were in positions in which the LacZ protein was either being independently translated or being rapidly cleaved from AmtB during or after insertion of the hybrid protein into the membrane.

Subcellular location of the Ala-33-PhoA fusion

The low activity of the PhoA fusion at Ala-428, combined with the clear appearance of the intact fusion protein on a Western blot (Fig. 4A), provides evidence that the C-terminus of the protein is on the cytoplasmic side of the membrane. The location of the N-terminus is also proposed to be on the inside of the cell, but this is difficult to determine using fusions to short N-terminal fragments of AmtB. We have established that fusion of PhoA after the first predicted TM helix (Ala-33) has a high level of AP activity, which demonstrates that this region provides sufficient topogenic information to target PhoA efficiently to the periplasmic side of the membrane. The addition of residues that include TM II (Gln-69) locate PhoA back on the cytoplasmic side, resulting in low AP activity. We can deduce the location of the N-terminus by localizing the position of the Ala-33-PhoA fusion within the cell. If the fusion is membrane bound, a genuine TM helix is being formed, which means that the N-terminus must be on the inside. Alternatively, if the N-terminus is on the outside, then the first 33 residues cannot form a TM helix and are probably being used as a signal peptide by PhoA for its localization to the periplasm, where it will have high activity.

To determine whether the Ala-33-PhoA fusion is membrane bound or periplasmic, we prepared periplasmic extracts and sphaeroplasts from a strain expressing the Ala-33 fusion. We made similar fractions from cells

expressing longer fusions at His-122, Asp-335 and Gln-375 as controls. The PhoA-reactive polypeptides within these fractions were determined by Western blotting (Fig. 5). The three controls illustrate that accumulation of the membrane-bound fusion proteins results in release of some PhoA into the periplasm when the fusion junction is periplasmic (Fig. 5A, compare H122 and Q375 with D335). However, in all cases, the full-length fusions were only found in the sphaeroplasts. In cells expressing the Ala-33 fusion, there are some PhoA-reactive bands in the periplasm, but the majority of the signal is retained in the sphaeroplasts (Fig. 5B). This distribution can be seen most clearly on a Coomassie stain of the membrane (Fig. 5B).

Additionally, the AP activities of the two subcellular extracts were determined using the same quantities of extracts used in Fig. 5 ($\approx 25 \mu\text{g}$ of protein). Each of the fusions with periplasmic fusion junctions, including the Ala-33 fusion, had > 10 -fold more AP activity in the sphaeroplasts than in the periplasm. Hence, the majority of these fusion proteins appear to be membrane bound, from which we can deduce that the first 33 residues form a genuine TM helix and that the N-terminus is cytoplasmic.

Ability of AmtB and hybrid proteins to transport [^{14}C]-methylammonium

To investigate whether full-length AmtB could function with an additional protein domain at its C-terminus and what effect the progressive C-terminal deletion of its TM helices would have on activity, we first constructed an *E. coli* strain containing a deletion of the *amtB* gene. An unmarked deletion of *amtB* was introduced onto the chromosome of strain ET8000, generating strain GT1001. The loss of *amtB* in this strain resulted in the complete loss of the ability of washed bacterial suspensions to transport and assimilate the ammonium analogue [^{14}C]-methylammonium (data not shown). This is in agreement with similar data reported by Soupene *et al.* (1998). Transport activity is restored to over 200% of wild-type levels by complementation of GT1001 with a plasmid containing *amtB* (pGT14). Transformation of GT1001 with a plasmid expressing full-length fusions (Ala-428) of AmtB to both PhoA and LacZ α resulted in no detectable restoration of [^{14}C]-methylammonium assimilation (data not shown). Similarly, no restoration of transport activity was seen with any of the shorter fusions at positions Gly-305, Asp-335, Asn-375 and Ala-428. This demonstrates that the insertion of a large protein domain at the C-terminus of this protein has a dramatic effect on its ability to transport [^{14}C]-methylammonium.

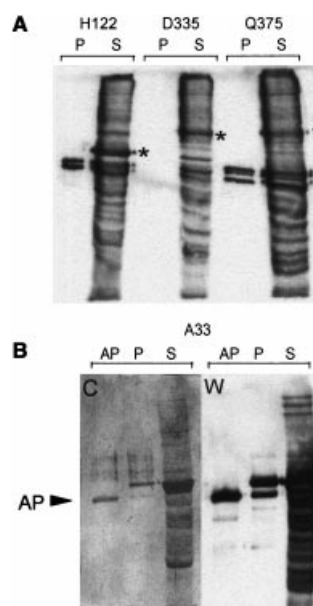


Fig. 5. A. Western blot of periplasms (P) and sphaeroplasts (S) prepared from ET8000 expressing *amtB-phoA* fusions at positions His-122, Asp-335 and Gln-375. The stars represent the positions of the intact fusion protein.

B. Extracts from ET8000 expressing the *amtB-phoA* fusion at position Ala-33 stained with Coomassie blue (C) or Western blotting with anti-PhoA antibody (W). Both images are of the same membrane, which included purified AP (AP) as a relative size marker.

Conservation and variphobicity of amino acids within the Mep/Amt family

Multiple sequence alignment of 16 Mep/Amt proteins known to be able to catalyse transport of [^{14}C]-methylammonium (identified in Fig. 1) reveals that only 27 amino acids are conserved in all these sequences (Fig. 6). These 16 sequences are from phylogenetically distinct sources, including plants, yeast and eubacteria, which give a good representation of the Mep/Amt family (Fig. 1). No Archaeal sequences were included, as none of these has yet been functionally characterized. The conserved residues cluster in four main regions of the protein. Fifteen of these residues are found in three pairs of adjacent helices, I/II, V/VI and IX/X (see *Discussion* for new numbering of TMs). Eight of the remaining 11 are located on the cytoplasmic side of the membrane, whereas of the three that are periplasmic, two are located directly adjacent to helices V/VI. Ten of the 16 conserved residues that are within TM helices are glycine residues. In contrast, six of the eight conserved cytoplasmic residues are charged.

Variphobicity analysis based on the 16 sequences with known transport ability is summarized in Table 3, with specific amino acid positions given for the *E. coli* AmtB protein. Variphobic residues in the transmembrane helices were identified as those residues that were not

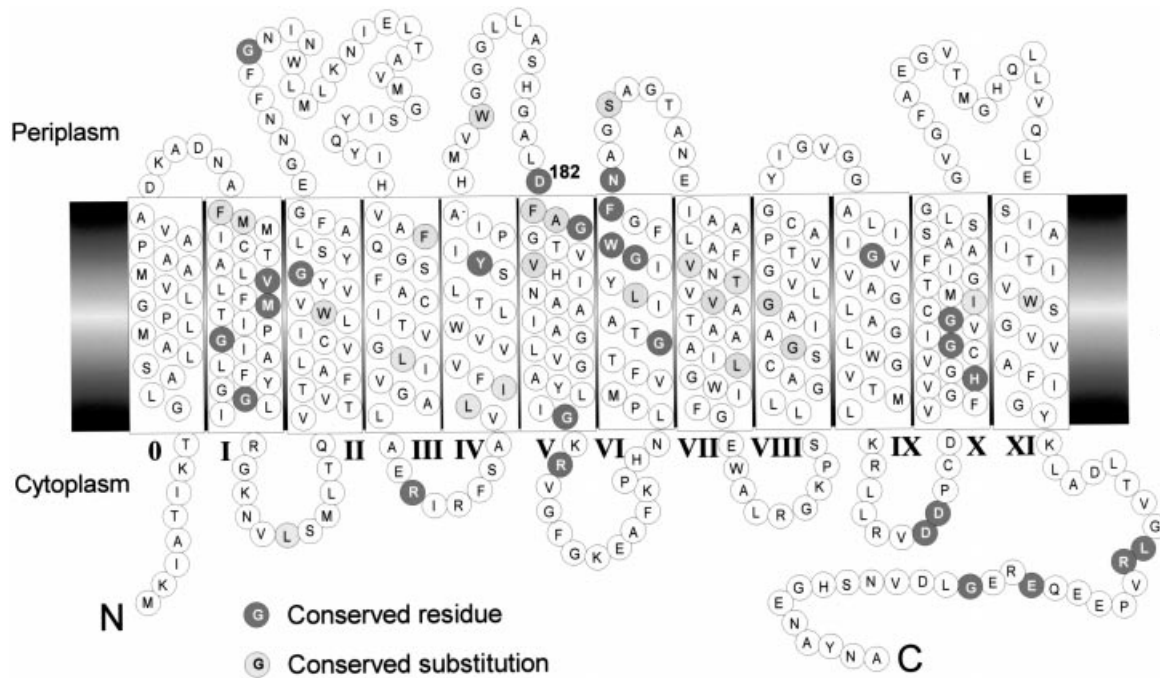


Fig. 6. Topological model of the *E. coli* AmtB protein overlaid with conserved residues within the Mep/Amt family. These residues were identified using a subset of Mep/Amt proteins (see *Experimental procedures*), which were aligned using CLUSTALX. Black balls represent residues that are completely conserved in this analysis, whereas grey balls are those in which a conservative substitution is found. The location of a putative ammonium binding site at Asp-182 is indicated.

conserved in identity across different members of the Mep/Amt proteins, but that always maintained a hydrophobic nature. The variphobicity index suggests that a number of helices, namely TM I, III, V, VI, IX and X, reside in a generally buried environment, with less interaction with the lipid bilayer, as they have a low to moderate number of variphobic residues and a number of conserved polar residues. In contrast, TM II has the highest variphobic index so, therefore, is more likely to present a hydrophobic face for interaction with lipids. The most prominent ridges are present in TM I, II and VI, and a

number of potential interhelical fits were identified (Table 2).

Discussion

Topology of the *E. coli* AmtB protein

In this study, we have used a minimal set of protein fusions between AmtB and AP (PhoA) to test an *in silico*-derived model and thereby determine the topology of the *E. coli* AmtB protein in the cytoplasmic membrane. This is

Table 2. Variphobicity and ridge/groove analysis for the *E. coli* AmtB protein.

TM helix	Position (length)	Variphobic residues	Conserved polar residues	Variphobicity index	Conserved residues forming potential ridge/groove	Interhelical fits
0	9–27 (19)	NA	NA	NA	NA	NA
I	34–58 (25)	5/25 = 0.20	2/25 = 0.08	0.12	F34, F44, Y54/G44, G56	V, VI, IX
II	70–91 (22)	7/22 = 0.32	1/22 = 0.05	0.27	W81, Y84, Y86, F90/G85	VI, VIII, IX, X
III	123–141 (19)	3/19 = 0.16	1/19 = 0.05	0.11	F128, F132/G140	V, VII, X
IV	150–166 (17)	5/17 = 0.29	1/17 = 0.06	0.23	Y163	X
V	183–205 (23)	4/23 = 0.17	3/23 = 0.13	0.04	F183/G185	I, III, VI, VIII
VI	220–237 (18)	4/18 = 0.22	3/18 = 0.17	0.05	Y231, W234, F237/G226, G233	I, II, V, X
VII	248–270 (23)	5/23 = 0.22	1/23 = 0.04	0.18	None	None
VIII	280–299 (20)	4/20 = 0.20	0/20 = 0.00	0.20	G286, G290	II, V
IX	306–324 (19)	2/19 = 0.11	1/19 = 0.05	0.06	G310	I, II
X	336–360 (25)	3/25 = 0.12	3/25 = 0.12	0.00	H341/G345, G348	II, III, IV, V, VI, XI
XI	380–396 (17)	3/17 = 0.18	0.17 = 0.00	0.18	W387	VII, X

NA, not applicable.

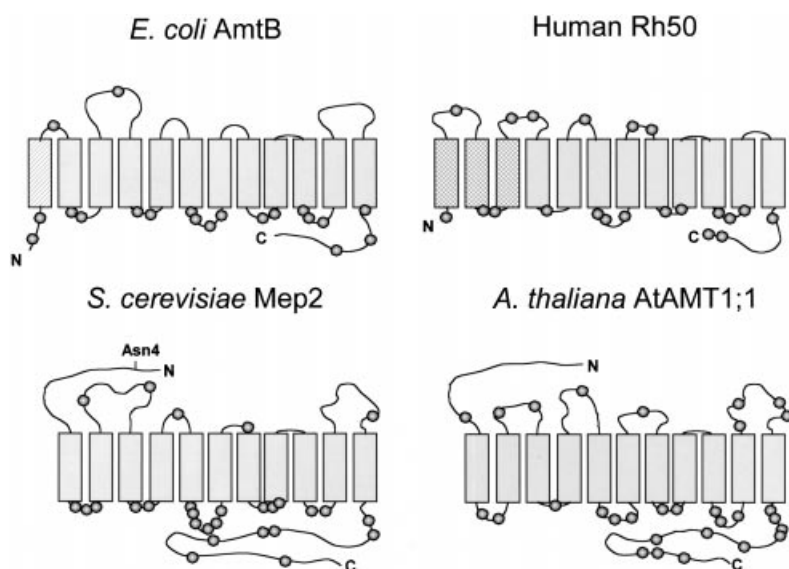


Fig. 7. Schematic models of three members of the Mep/Amt proteins and the human Rh50 proteins. Shaded helices are those in which topology is conserved through the Mep/Amt and Rh families. Hatched boxes represent helices that are not homologous. The balls represent positive charge in these protein sequences.

the first secondary structural information for any member of this new family of ammonium transport proteins. The final model for the *E. coli* protein contains 12 TM helices with an $N_{IN}C_{IN}$ orientation (Fig. 6). The distinction between high- and low-activity AmtB–PhoA fusions was large enough to ascribe the cellular location of each fusion junction unambiguously (Fig. 3). Western blotting demonstrated that each fusion protein was being expressed (Fig. 4), although the total amount of signal was much greater for the periplasmically located fusions. Here, degradation of the fusion proteins releases AP into the periplasm where it is stable (for examples, see Fig. 5) and contributes to the total AP activity measured. We included two PhoA fusions in one cytoplasmic loop, which were positioned near the start and end of the loop (G209 and V223). However, there was no significant difference in the AP activities of the two fusions (Fig. 3), which demonstrated that the region between residues 209 and 223 is not required for efficient insertion of the previous TM helix. This might be because of the presence of two positive charges at residues K206 and R207 (Fig. 3), which would be sufficiently strong topogenic signals to ensure efficient insertion of the preceding TM helix.

Although the data from the AmtB–PhoA fusions was quite conclusive, we constructed a complementary set of fusions to β -galactosidase. However, despite using both *in vivo* and *in vitro* methods with random or engineered sites, the resultant data were not completely unambiguous. This was mainly because of the tendency of the LacZ domain of the fusion protein to be cleaved off, either during synthesis or after insertion of the fusion in the membrane. As LacZ is active in the cytoplasm, this cleavage gives false-positive results that confuse the analysis. These problems have been reported by others

(Hennessey and Broome-Smith, 1993; Geest and Lolke, 2000), and we are led to conclude that this is not a reliable method of topology analysis.

Topology of the Mep/Amt protein family

The hydropathy profile alignment of the Mep/Amt proteins (Fig. 2) illustrates the conservation of a core of 11 TM helices within these proteins. We have concluded that the *E. coli* AmtB protein may have an additional 12th TM helix at its N-terminus and, in the subset of 16 functionally characterized Mep/Amt sequences, up to five other bacterial sequences might have an additional helix at the N-terminus (*A. brasilense* AmtB, *A. caulinodans* AmtB, *A. vinelandii* AmtB and *Synechocystis* PCC6803 Amt1 and Amt2). The balance of charge in these regions of the six proteins also suggests the presence of an additional helix. However, there is no sequence conservation in any of these regions, and most contain a number of prolines. Therefore, they would probably only form short broken helices, which would be unlikely to pack tightly with the other core 11 TM regions and are unlikely to be important for function. In *E. coli*, large integral membrane proteins favour a topology with 12 membrane-spanning regions and an $N_{IN}C_{IN}$ orientation (Jones, 1998; Wallin and von Heijne, 1998). Given that the proposed 'functional core' of the Mep/Amt proteins is 11 TM helices, it is possible that there are evolutionary reasons for the extra helix in *E. coli* AmtB, perhaps related to protein folding during insertion into the cytoplasmic membrane.

As we conclude that the general structure of the Mep/Amt proteins contains 11 TM helices, we propose that the homologous TM helices in different proteins are always numbered in the same way. Therefore, the 11 conserved

TM helices should be designated TM 1–11 and, in cases such as *E. coli*, the additional N-terminal TM should be designated TM 0 (see Fig. 6).

Figure 7 shows the proposed topology models for Mep/Amt proteins from *S. cerevisiae* and *A. thaliana*, the organisms in which these transporters were first discovered, compared with the *E. coli* AmtB protein. In agreement with the proposed topology, a glycosylation site has been identified at residue Asn-4 of the *S. cerevisiae* Mep2 protein, demonstrating that the N-terminus of this protein has an extracytosolic location (A. M. Marini, and B. Andre, personal communication). The distribution of positive charge is also illustrated in Fig. 7, showing that the positive inside rule is maintained in each case. Recently, two homologous proteins with identical hydrophobic regions were shown to have opposite topologies (Sääf *et al.*, 1999), and this was correlated with an inversion of the distribution of positively charged residues in loop regions. Hence, the conservation of positive charge distribution throughout the Mep/Amt family further supports the concept that their topology is conserved.

Structure–function relationships in the Mep/Amt proteins

Multiple sequence alignment of the Mep/Amt proteins, combined with our determined topology, offers some clues to structure–function relationships in this family of transporters. If the determined topology is superimposed over the hydropathy profile (Fig. 2), then it is apparent that there is more variation in the periplasmic loops compared with the cytoplasmic loops, suggesting that the cytoplasmic face of the protein is more conserved. In agreement with this, we find that, of the 11 ‘non-TM’ amino acid residues that are totally conserved within the Mep/Amt proteins, eight of them are located on the cytoplasmic side of the membrane, whereas two of the three that are periplasmic are directly adjacent to TM helices (Fig. 6).

Helices V and VI have low variphobic indices and might form a ridge/groove arrangement with the conserved Trp-234 in helix VI and the conserved Gly-185 in helix V. Helices I and X, which contain multiple conserved residues, also have low variphobic indices and again might not be directly adjacent to lipid. Both these helices might form potential ridge/groove associations with TMs V and VI. As illustrated in Fig. 6, the region of the protein that has the greatest level of sequence conservation is the helix V/VI region. At the periplasmic face of TM V is a conserved Asp-182 residue. As an aspartic acid residue has been implicated in binding ammonium ions in *E. coli* glutamine synthetase (Alibhai and Villafranca, 1994; Liaw *et al.*, 1995), we suggest that this residue is a potential candidate for an initial ammonium binding site within the protein. This might be aided in the formation of a binding pocket by the highly polar conserved asparagine at a

similar position on the periplasmic face of helix VI. The ammonium ion might then pass through a polar corridor formed by the conserved glycines present at different depths in a number of TM helices. Finally, the two adjacent conserved negative charges in the loop IX–X region might associate with the ammonium ion as it passes from the membrane to the cytoplasm. Although the mechanism of these proteins is unknown, the structural features described above all lead us to favour the hypothesis that ammonium is the substrate for this family of transporters. However, it has been proposed recently that the *E. coli* AmtB actually uses the ammonia molecule as the substrate instead of the ammonium cation (Soupene *et al.*, 1998). The precise mechanism of the Mep/Amt proteins clearly still remains to be demonstrated experimentally.

E. coli AmtB has a C-terminal cytoplasmic tail of 32 residues that is present in all Mep/Amt proteins and includes three completely conserved residues (Fig. 6). In other members of the family, especially in eukaryotes, this region is extended by up to another 80 residues. This conservation suggests that the C-terminus has a structural or functional significance and, in agreement with this, we found that perturbation of the C-terminus of AmtB by the addition of other protein domains abolished its activity. Additionally, if the C-terminal tail of the AmtB protein is removed, then activity is significantly reduced, but not abolished (G. Coutts and M. Merrick, unpublished data). Within the cytoplasmic face of *E. coli* AmtB, there is a marked complementarity of charge distribution, in that all but one of the cytoplasmic loops has a net positive charge, whereas the cytoplasmic tail is negatively charged. This may indicate the presence of significant electrostatic interactions between charged residues in the tail and the exposed cytoplasmic face of the protein, a situation that is not unprecedented. Recent data for a metal-tetracycline/H⁺ antiporter (TetA(B)) have demonstrated the presence of an important charged pair that is formed by residues in two cytoplasmic loop regions (Someya *et al.*, 2000).

Implications for Rhesus protein structure and function

The Mep/Amt proteins are homologous to the human Rh (Rhesus) blood group polypeptides, a family of polytopic membrane proteins that display antigens involved in transfusion-incompatible immune reactions (Marini *et al.*, 1997; Avent and Reid, 2000). The different rhesus antigens are displayed on one of a number of Rh30 proteins, which form a tetrameric core complex with the Rh50 glycoprotein. Both the Rh30 and the Rh50 proteins are homologous to the Mep/Amt proteins, although the similarity is stronger for Rh50. The physiological functions of the Rh proteins remain unknown, but the Rh50 proteins

Table 3. Strains and plasmids used in this study.

Strain	Genotype or phenotype	Reference.
ET8000	<i>rbs lacZ::IS1 gyrA hutC_K</i>	Jayakumar <i>et al.</i> (1986)
GT1001	ET8000 Δ <i>amtB</i>	This work
DH5 α	<i>endA hsdR17 supE44 thi-1 λ^- recA1 ϕ80dLacZΔM15</i>	Bethesda Research Laboratories
CC118	<i>araD139 Δ(ara leu)7697 ΔlacX74 <i>phoA</i>Δ20 <i>galE thi rspE rpoB argE recA1</i></i>	Manoil and Beckwith (1985)
Plasmids		
pUC118	<i>lacZ</i> α -peptide fusion vector	Vieira and Messing (1987)
pSK4158	<i>phoA</i> fusion vector	Paulsen <i>et al.</i> (1995)
pNM480	<i>lacZ</i> fusion vector	Minton (1984)
pWSK29		Wang and Kushner (1991)
pWVH141	<i>mdl' glnK⁺ amtB⁺ tesB</i>	van Heeswijk <i>et al.</i> (1996)
pWVH153	Δ <i>glnK1, amtB⁺</i>	Arcondéguy <i>et al.</i> (1999)
pGT14	Δ <i>glnK1, amtB⁺</i>	This work
pA33 _{PHO/LAC/LACα}		This work
pQ69 _{PHO/LAC/LACα}		This work
pH122 _{PHO/LAC/LACα}		This work
pA150 _{PHO/LAC/LACα}		This work
pA180 _{PHO/LAC/LACα}		This work
pG209 _{PHO/LACα}		This work
pV223 _{PHO/LAC/LACα}		This work
pI248 _{PHO/LAC/LACα}		This work
pS279 _{PHO/LAC/LACα}		This work
pG305 _{PHO/LAC/LACα}		This work
pD335 _{PHO/LAC/LACα}		This work
pQ375 _{PHO/LAC/LACα}		This work
pA428 _{PHO/LACα}		This work

have been suggested to be involved in transmembrane solute transport, and Rh_{null} erythrocytes have abnormal cation permeability (Ballas *et al.*, 1984; Anstee and Tanner, 1988). Therefore, given their homology to the Mep/Amt family, one potential function of these proteins is that they transport ammonium.

Hydropathy analysis, together with immunochemical studies, has resulted in topology models for both Rh50 and Rh30 proteins with 12 TM helices (Eyers *et al.*, 1994; Avent *et al.*, 1996). To determine whether the topology of the Rh proteins and the Mep/Amt proteins is identical, we have extended the analysis of Marini *et al.* (1997) and compared the two protein families. A schematic model of our finding is presented in Fig. 7. Both Rh and Mep/Amt have homologous sequence and identical topology over the majority of the polypeptide (helices III to XI in Mep/Amt correspond with helices IV to XII in Rh). However, this alignment breaks down towards the N-terminus of both proteins, and we do not find any significant similarities between helices I and II of the Mep/Amt proteins and helices II and III of the Rh proteins.

In terms of the sequence identity between the two protein families, it is significant that the match is strongest to the TM V/VI region of the Mep/Amt proteins. This includes conservation of residues equivalent to Asp-182 and Asn-238 in *E. coli* AmtB, which we have proposed as being involved in the binding of ammonium. Four of the intramembrane glycines in the protein and one of the aspartates in the cytoplasmic loop IX–X region are also conserved. Although this suggests that the Rh proteins might be able to transport [¹⁴C]-

methylammonium, it is also possible that the absence in the Rh proteins of the conserved residues from TMs I and II of the Mep/Amt proteins (Fig. 6) could change or impair their transport properties.

Experimental procedures

Bacterial strains, plasmids and culture conditions

The bacterial strains and plasmids used in this study are listed in Table 3. Bacteria were routinely grown aerobically on Luria medium [Luria broth or Luria agar (LA)]. For growth under nitrogen limitation, an altered M9 medium was used, in which the ammonium sulphate was replaced by 200 μ g ml⁻¹ glutamine (M9Gln medium). For growth of strain CC118 on M9Gln medium, thiamine was added at 1 μ g ml⁻¹, arginine at 25 μ g ml⁻¹ and leucine at 25 μ g ml⁻¹. Antibiotics were used at the following concentrations: carbenicillin, 100 μ g ml⁻¹ and chloramphenicol, 40 μ g ml⁻¹.

DNA methods

Plasmids were isolated using a plasmid miniprep kit (Qiagen) in accordance with the manufacturer's instructions. The methods of Sambrook *et al.* (1989) were used for restriction enzyme digests and ligations. Automatic sequencing was performed using BigDye polymerase (Applied Biosystems), which was analysed on an ABI sequencer.

Sequences and alignments of Mep/Amt proteins

The sequences of most Mep/Amt proteins can be found in the

Prodom domain PD001450 (<http://www.toulouse.inra.fr/prodom.html>). A number of additional members of the family were identified in searches of uncompleted genome sequences (<http://www.tigr.org/>). Multiple sequence alignments were generated in CLUSTALX (Thompson *et al.*, 1997). The average hydropathy plots were determined using the PROFILE program described by Lolkema and Slotboom (1998). The window size was set at 19 residues, and a subset of 23 sequences with pairwise alignments of between 21% and 64% was used in the alignment.

Enzyme assays

AP activities were assessed on M9Gln agar plates by breakdown of the chromogenic substrate 5-bromo-4-chloro-3-indolyl phosphate (XP), which had been incorporated at a concentration of 40 µg ml⁻¹. Similarly, β-galactosidase activities were assessed on M9Gln agar plates by breakdown of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal), which had been incorporated at a concentration of 40 µg ml⁻¹. For enzyme assays, bacteria were first grown in LB medium, then diluted 50-fold into M9Gln medium for overnight growth. AP activities were quantified after growth in M9Gln medium using the method of Manoil (1991). β-Galactosidase activities were determined using the method of Miller (1972), and activities are expressed in Miller units.

Quantification of the AP activities was initially attempted in the *phoA* mutant strain CC118, which is frequently used in studies of this nature, but activities were surprisingly low. Furthermore, when the AmtB-dependent [¹⁴C]-methylammonium transport activity of CC118 was compared with that of ET8000, a wild-type strain used in our laboratory, strain CC118 had a 100-fold lower activity than strain ET8000 (data not shown). This same difference in activity was seen when plasmid pGT14, which contains *amtB* expressed from its natural promoter, was transformed into both these backgrounds. Expression of the *amtB-phoA* fusions is driven from the natural *glnKamtB* promoter, which is induced by the Ntr system during growth under nitrogen-limiting conditions (van Heeswijk *et al.*, 1996; Atkinson and Ninfa, 1998). We conclude that these low levels of activity were caused by some defect in nitrogen regulation in CC118, but the nature of this defect was not determined. For subsequent quantification of the AmtB-PhoA activities, we used ET8000, which, although *phoA*⁺, is white on indicator plates and gives levels of AP activity of around 10 units, presumably resulting from repression of *phoA* synthesis caused by the high level of phosphate (20 mM) in M9 medium.

Western blotting

Whole-cell extracts from *E. coli* strain ET8000 transformed with each *amtB-phoA* and *amtB-lacZ* fusion construct were prepared from washed stationary phase cells broken by sonication. Periplasmic proteins were released from bacteria by treatment with lysozyme and EDTA according to the method of McEwan *et al.* (1984). The resulting sphaeroplasts were then disrupted by sonication at 10 µm for two 30 s steps. Protein concentrations were determined using the

Bio-Rad protein assay, and 25 µg quantities were separated by SDS-PAGE. After transfer to a nitrocellulose filter (Hybond C; Amersham), the membranes were reacted with either anti-PhoA antibody (5 Prime → 3 Prime) or anti-LacZ antibody (Promega) and detected using the ECL detection system (Amersham). After exposure, membranes were stained with Coomassie blue to confirm the uniform transfer of the proteins to nitrocellulose.

Construction of an *amtB* strain

A strain of *E. coli* containing an unmarked deletion of the *amtB* gene was constructed using the method of Link *et al.* (1997). The following PCR steps used *Taq* DNA polymerase (Gibco BRL) and plasmid pWVH141 (van Heeswijk *et al.*, 1996) as template DNA. Plasmid pWVH141 contains the complete *glnKamtB* region and the ends of the flanking *mdl* and *tesB* genes. The PCR primers AmtBNo (5'-CG CGGATCCGTGCACTGTCATAGTGC GTT-3') and AmtBNI (5'-CCCATCCTCTAGACTTAAACATATCGTTCGC-3') were used to amplify a 400 bp region upstream of *amtB*, which ends with the first six codons of *amtB*. A similar pair of primers AmtBCi (5'-TGTTTAAGTCTAGAGGATGGGGATG TCAACAGCCACGGCGAG-3') and AmtBCo (5'-CGCGG ATCCGCGCAAATGGTAGC-3') were used to amplify a 500 bp region extending from the last seven codons of *amtB* into the downstream *tesB* gene. The two PCR products were purified, mixed and used as templates for a second PCR reaction using primers AmtBNo and AmtBCo. The resulting PCR product was digested with *Bam*HI and cloned into *Bam*HI-cut pKO3 (Link *et al.*, 1997). This unmarked deletion of the *amtB* gene was crossed onto the chromosome of strain ET8000 using the double selection method described by Link *et al.* (1997), creating strain GT1001. Resolution of the recombination was checked extensively by PCR to confirm that no wild-type sequence remained.

Construction of an *amtB* expression plasmid

We constructed a plasmid that contained *amtB* under the control of its natural regulated promoter. The starting point for construction of this plasmid was pWVH153, which contains the *glnKamtB* region with an unmarked deletion of *glnK* (Δ *glnK1*) in the pUC18-Not vector (Arcondéguy *et al.*, 1999). The Δ *glnK1,amtB*⁺ region was removed by digestion of pWVH153 with *Hind*III and *Sma*I, and this was then cloned into the low-copy-number vector pWSK29, creating pGT14. This construct therefore contains full-length *amtB* expressed from its natural NtrC-dependent promoter.

Construction of *amtB-phoA* and *-lacZ* fusions

For each fusion, an upstream primer preceding the *glnK* promoter and a specific primer, which annealed within the coding sequence of *amtB*, were used to amplify DNA fragments from plasmid pGT14 by PCR. These PCR products were purified and then digested by restriction enzymes to facilitate cloning into *phoA* (pSK4158) and *lacZ* (pUC118) reporter plasmids. The *phoA* gene in pSK4158 is truncated so that it expresses a version of the protein without

a signal peptide. To construct the *phoA* fusion series, the PCR products were digested with *HindIII* and *BglII* and cloned into *HindIII*–*BamHI*-cut pSK4158. Similarly, to construct the *lacZ* α fusions, the PCR products were digested with *SmaI* and *BglII* and cloned into *SmaI*–*BamHI*-cut pUC118. A list of the plasmids generated using this method is found in Table 3; they have been named after the last *AmtB* residue in the in frame fusion. Fusions to full-length *lacZ* were constructed using the same PCR products cut with *SmaI* and *BglII* and cloned into *SmaI*–*BamHI*-cut pNM480.

The *in vivo* isolation of mutants was attempted according to the method of Manoil (1991), using pGT14 as the starting construct for infection with Tn*lacZ* (Manoil and Bailey, 1997). Colonies that were blue on indicator plates were purified, and the location of the fusion junction within the plasmid was estimated using PCR with a primer in the *mdl* gene and another at the 5' end of the *lacZ* insert. Fusions that were within the *amtB* gene were sequenced to determine the exact location of the fusion junctions.

Measurement of [¹⁴C]-methylammonium assimilation

Methylammonium uptake and assimilation rates were determined essentially as described by Jack *et al.* (1999). However, the growth medium was M9Gln, and filters were dissolved in Filter Count scintillant (Packard Biosciences). Radioactivity in the samples was then counted with a Wallac 1409 scintillation counter, which had been calibrated with a known set of standards.

Analysis of hydrophathy, variphobicity and potential ridge/groove arrangements

For the variphobicity analysis and the determination of conserved residues within the Mep/Amt proteins, a subset of 16 proteins was used that had been experimentally determined to transport [¹⁴C]-methylamine (identified in Fig. 1). Transmembrane regions were predicted by averaging the outputs of a number of hydrophathy analysis programs, namely TMPRED (Hofmann and Stoffel, 1993), TOPPED 2 (von Heijne, 1992), HMMTOP (Tusnády and Simon, 1998) and TMHMM (Sonnhammer *et al.*, 1998). The predicted transmembrane regions were then aligned to identify variphobic residues according to the method of Donnelly *et al.* (1993). A variphobicity index was calculated by subtracting the proportion of conserved polar residues in a given helix from the proportion of variphobic residues, as shown in Table 2. For this purpose, polar residues included asparagine (N), proline (P), glutamine (Q), serine (S), threonine (T), tryptophan (W) and tyrosine (Y) together with any of the charged residues, namely aspartate (D), glutamate (E), histidine (H), lysine (K) and arginine (R). The hydrophobic residues used in the variphobicity measurements were alanine (A), cysteine (C), isoleucine (I), leucine (L), methionine (M), phenylalanine (F) and valine (V). These classifications were made according to the hydrophaticity scale of Kyte and Doolittle (1982).

Interhelical ridge/groove scores were obtained for *E. coli* *AmtB* using two-dimensional MEM-MAP, a graphic software application (J. Antoniw and J. G. L. Mullins, unpublished). This software can be used to identify potential interhelical

ridge/groove fits, as described by Chothia *et al.* (1981), through detection of compatible amino acids (aromatic residue on one helix, glycine on another), which reside at similar depth (± 1.5 Å) in the membrane. Ridge/groove arrangements between two helices were accepted if there were two or more aromatic/glycine fits predicted between a pair of helices.

Acknowledgements

This study was supported by a grant from CEE (EURATINE:BIOTEC94-2310) and by the BBSRC through a grant-in-aid to the John Innes Centre. We thank Dr Wally van Heeswijk, Dr Melissa Brown and Dr Colin Manoil for strains and plasmids, and are very grateful to Dr Juke Lolkema for help with the hydrophathy analysis. We also thank Dr Melissa Brown and Dr Ian Paulson for helpful advice, and Dr Tania Arcondeguy, Dr Sara Austin, Professor Ray Dixon, Professor Bruno Andre and Dr Gary Sawers for helpful comments on the manuscript.

References

- Alibhai, M., and Villafranca, J.J. (1994) Kinetic and mutagenic studies of the role of the active site residues Asp-50 and Glu-327 of *Escherichia coli* glutamine synthetase. *Biochemistry* **33**: 682–686.
- Anstee, D.J., and Tanner, M.J. (1988) Blood group antigen deficiencies associated with abnormal red cell shape. *Blood Rev* **2**: 115–120.
- Arcondeguy, T., van Heeswijk, W.C., and Merrick, M. (1999) Studies on the roles of GlnK and GlnB in regulating *Klebsiella pneumoniae* NifL-dependent nitrogen control. *FEMS Microbiol Lett* **180**: 263–270.
- Atkinson, M.R., and Ninfa, A.J. (1998) Role of the GlnK signal transduction protein in the regulation of nitrogen assimilation in *Escherichia coli*. *Mol Microbiol* **29**: 431–447.
- Atkinson, M.R., and Ninfa, A.J. (1999) Characterization of the GlnK protein of *Escherichia coli*. *Mol Microbiol* **32**: 301–313.
- Avent, N.D., and Reid, M.E. (2000) The Rh blood group system: a review. *Blood* **95**: 375–387.
- Avent, N.D., Liu, W., Warner, K.M., Mawby, W.J., Jones, J.W., Ridgwell, K., *et al.* (1996) Immunochemical analysis of the human erythrocyte Rh polypeptides. *J Biol Chem* **271**: 14233–14239.
- Ballas, S.K., Clark, M.R., Mohandas, N., Colfer, H.F., Caswell, M.S., Bergren, M.O., *et al.* (1984) Red cell membrane and cation deficiency in Rh null syndrome. *Blood* **63**: 1046–1055.
- Boyd, D., Traxler, B., and Beckwith, J. (1993) Analysis of the topology of a membrane protein by using a minimum number of alkaline phosphatase fusions. *J Bacteriol* **175**: 553–536.
- Chothia, C., Levitt, M., and Richardson, D. (1981) Helix to helix packing in proteins. *J Mol Biol* **145**: 215–250.
- Chou, P.Y., and Fasman, G.D. (1979) Prediction of beta-turns. *Biophys J* **26**: 367–373.
- Dandekar, T., Snel, B., Huynen, M., and Bork, P. (1998)

- Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci* **23**: 324–328.
- Deleage, G., and Roux, B. (1987) An algorithm for protein secondary structure prediction based on class prediction. *Protein Eng* **1**: 289–294.
- van Dommelen, A., Keijers, V., Vanderleyden, J., and de Zamaroczy, M. (1998) (Methyl)ammonium transport in the nitrogen-fixing bacterium *Azospirillum brasilense*. *J Bacteriol* **180**: 2652–2659.
- Donnelly, D., Overington, J.P., Ruffle, S.V., Nugent, J.H., and Blundell, T.L. (1993) Modeling alpha-helical transmembrane domains: the calculation and use of substitution tables for lipid-facing residues. *Protein Sci* **2**: 55–70.
- Eyers, S.A., Ridgwell, K., Mawby, W.J., and Tanner, M.J. (1994) Topology and organization of human Rh (rhesus) blood group-related polypeptides. *J Biol Chem* **269**: 6417–6423.
- Geest, M., and Lolkema, J.S. (2000) Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol Mol Biol Rev* **64**: 13–33.
- van Heeswijk, W.C., Hoving, S., Molenaar, D., Stegeman, B., Kahn, D., and Westerhoff, H.V. (1996) An alternative P_{II} protein in the regulation of glutamine synthetase in *Escherichia coli*. *Mol Microbiol* **21**: 133–146.
- von Heijne, G. (1992) Membrane protein structure prediction, hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225**: 487–494.
- Hennessey, E.S., and Broome-Smith, J.K. (1993) Gene-fusion techniques for determining membrane-protein topology. *Curr Opin Struct Biol* **3**: 524–531.
- Hofmann, J., and Stoffel, W. (1993) TMBASE – a database of membrane spanning proteins segments. *Biol Chem Hoppe-Seyler* **347**: 166–168.
- Jack, R., de Zamaroczy, M., and Merrick, M. (1999) The signal transduction protein GlnK is required for NifL-dependent nitrogen control of *nif* gene expression in *Klebsiella pneumoniae*. *J Bacteriol* **181**: 1156–1162.
- Jayakumar, A., Schulman, D., MacNeil, D., and Barnes, E.M., Jr (1986) Role of the *Escherichia coli* *glnALG* operon in regulation of ammonium transport. *J Bacteriol* **166**: 281–284.
- Jones, D.T. (1998) Do transmembrane protein superfolds exist? *FEBS Lett* **423**: 281–285.
- Jones, D.T., Taylor, W.R., and Thornton, J.M. (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**: 3038–3049.
- Kleiner, D. (1993) Ammonium transport systems – an overview. In *Alkali Cation Transport Systems in Prokaryotes*. Bakker, E.P. (ed.). Florida: CRC Press, pp. 379–396.
- Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**: 105–132.
- Knepper, M.A., Packer, R., and Good, D.W. (1989) Ammonium transport in the kidney. *Physiol Rev* **69**: 179–249.
- Liaw, S.H., Kuo, I., and Eisenberg, D. (1995) Discovery of the ammonium substrate site on glutamine synthetase, a third cation binding site. *Protein Sci* **4**: 2358–2365.
- Link, A.J., Phillips, D.R., and Church, G.M. (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* **179**: 6228–6237.
- Lolkema, J.S., and Slotboom, D.-J. (1998) Estimation of structural similarity of membrane proteins by hydrophathy profile alignment. *Mol Membr Biol* **15**: 33–42.
- McEwan, A.J., Jackson, J.B., and Ferguson, S.J. (1984) Rationalisation of properties of nitrate reductases in *Rhodospseudomonas capsulata*. *Arch Microbiol* **137**: 344–349.
- Manoil, C. (1991) Analysis of membrane protein topology using alkaline phosphatase and β -galactosidase gene fusions. *Methods Cell Biol* **34**: 61–75.
- Manoil, C., and Bailey, J. (1997) A simple screen for permissive sites in proteins: analysis of *Escherichia coli lac* permease. *J Mol Biol* **267**: 250–263.
- Manoil, C., and Beckwith, J. (1985) TnphoA: a transposon probe for protein export signals. *Proc Natl Acad Sci USA* **82**: 8129–8133.
- Marini, A.M., Vissers, S., Urrestarazu, A., and Andre, B. (1994) Cloning and expression of the MEP1 gene encoding an ammonium transporter in *Saccharomyces cerevisiae*. *EMBO J* **13**: 3456–3463.
- Marini, A.M., Urrestarazu, A., Beauwens, R., and Andre, B. (1997) The Rh (Rhesus) blood group polypeptides are related to NH₄⁺ transporters. *Trends Biochem Sci* **22**: 460–461.
- Meletzus, D., Rudnick, P., Doetsch, N., Green, A., and Kennedy, C. (1998) Characterization of the *glnK-amtB* operon of *Azotobacter vinelandii*. *J Bacteriol* **180**: 3260–3264.
- Michel-Reydellet, N., Desnoues, N., de Zamaroczy, M., Elmerich, C., and Kaminski, P.A. (1998) Characterisation of the *glnK-amtB* operon and the involvement of AmtB in methylammonium uptake in *Azorhizobium caulinodans*. *Mol Gen Genet* **258**: 671–677.
- Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbour, NY: Cold Spring Harbour Laboratory Press.
- Minton, N.P. (1984) Improved vectors for the isolation of translational *lac* gene fusions. *Gene* **31**: 269–273.
- Montesinos, M.L., Muro-Pastor, A.M., Herrero, A., and Flores, E. (1998) Ammonium/methylammonium permeases of a cyanobacterium. Identification and analysis of three nitrogen-regulated *amt* genes in *Synechocystis* sp. PCC 6803. *J Biol Chem* **273**: 31463–31470.
- Ninnemann, O., Jauniaux, J.C., and Frommer, W.B. (1994) Identification of a high affinity NH₄⁺ transporter from plants. *EMBO J* **13**: 3464–3471.
- Paulsen, I.T., Brown, M.H., Dunstan, S.J., and Skurray, R.A. (1995) Molecular characterization of the Staphylococcal multidrug resistance export protein QacC. *J Bacteriol* **177**: 2827–2833.
- Persson, B., and Argos, P. (1997) Prediction of membrane protein topology utilizing multiple sequence alignments. *J Protein Chem* **16**: 453–457.
- Sääf, A., Johansson, M., Wallin, E., and von Heijne, G. (1999) Divergent evolution of membrane protein topology: the *Escherichia coli* RnfA and RnfE homologues. *Proc Natl Acad Sci USA* **96**: 8540–8544.
- Saier, M.H., Jr, Eng, B.H., Fard, S., Garg, J., Haggerty, D.A.,

- Hutchinson, W.J., *et al.* (1999) Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochim Biophys Acta* **1422**: 1–56.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Sarsero, J.P., and Pittard, A.J. (1995) Membrane topology analysis of *Escherichia coli* K-12 Mtr permease by alkaline phosphatase and β -galactosidase fusions. *J Bacteriol* **177**: 297–306.
- Servín-González, L., and Bastarrachea, F. (1984) Nitrogen regulation of synthesis of the high affinity methylammonium transport system of *Escherichia coli*. *J Gen Microbiol* **130**: 3071–3077.
- Siewe, R.M., Weil, B., Burkovski, A., Eikmanns, B.J., Eikmanns, M., and Krämer, R. (1996) Functional and genetic characterization of the (methyl)ammonium uptake carrier of *Corynebacterium glutamicum*. *J Biol Chem* **271**: 5398–5403.
- Sonnhammer, E.L.L., von Heijne, G., Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*. Glasgow, J., Littlejohn, T., Major, F., Lathrop, R., Sankoff, D., and Sensen, C. (eds). Menlo Park, CA: AAAI Press, pp. 177–183.
- Someya, Y., Kimura-Someya, T., and Yamaguchi, A. (2000) Role of the charge interaction between Arg(70) and Asp(120) in the *Tn10*-encoded metal-Tetracycline/H⁺ antiporter of *Escherichia coli*. *J Biol Chem* **275**: 210–214.
- Soupe, E., He, L., Yan, D., and Kustu, S. (1998) Ammonia acquisition in enteric bacteria: physiological role of the ammonium/methylammonium transport B (AmtB) protein. *Proc Natl Acad Sci USA* **95**: 7030–7034.
- Taté, R., Riccio, A., Merrick, M., and Patriarca, E.J. (1998) The *Rhizobium etli* *amtB* gene coding for an NH₄⁺ transporter is down-regulated early during bacteroid differentiation. *Mol Plant–Microbe Interact* **11**: 188–198.
- Thomas, G., Coutts, G., and Merrick, M. (2000) The *glnKamtB* operon: a conserved gene pair in prokaryotes. *Trends Genet* **16**: 11–14.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997) The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882.
- Traxler, B., Boyd, D., and Beckwith, J. (1993) The topological analysis of integral cytoplasmic membrane proteins. *J Membr Biol* **132**: 1–11.
- Tusnády, G.E., and Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* **283**: 489–506.
- Vieira, J., and Messing, J. (1987) Production of single-stranded plasmid DNA. *Methods Enzymol* **153**: 3–11.
- Wallin, E., and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7**: 1029–1038.
- Wang, R.F., and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**: 195–199.