MicroReview

Cation and sugar selectivity determinants in a novel family of transport proteins

Bert Poolman,1*, Jan Knol,1 Chris van der Does,1 Peter J. F. Henderson,2 Wei-Jun Liang,2 Gérard Leblanc,3 Thierry Pourcher3 and Isabelle Mus-Veteau3
1Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands.
2Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, UK.
3Laboratoire J. Maetz, Département de Biologie Cellulaire et Moléculaire du Commissariat à l’Energie Atomique, 06230 Villefranche-sur-mer, France.

Summary

A new family of homologous membrane proteins that transport galactosides–pentoses–hexuronides (GPH) is described. By analysing the aligned amino acid sequences of the GPH family, and by exploiting their different specificities for cations and sugars, we have designed mutations that yield novel insights into the nature of ligand binding sites in membrane proteins. Mutants have been isolated/constructed in the melibiose transport proteins of Escherichia coli, Klebsiella pneumoniae and Salmonella typhimurium, and the lactose transport protein of Streptococcus thermophilus which facilitate uncoupled transport or have an altered cation and/or substrate specificity. Most of the mutations map in the amino-terminal region, in or near amphipathic α-helices II and IV, or in interhelix-loop 10–11 of the transport proteins. On the basis of the kinetic properties of these mutants, and the primary and secondary structure analyses presented here, we speculate on the cation binding pocket of this family of transporters. The regulation of the transporters through interaction with, or phosphorylation by, components of the phosphoenolpyruvate:sugar phosphotransferase system is also discussed.

Introduction

In the absence of a three-dimensional structure, elucidation of the structure–activity relationship of a membrane protein relies upon combining a variety of strategies, such as comparisons of homologous sequences, mutagenesis, gene fusions, kinetic analyses, measurements of cation and substrate binding, spectroscopy, etc. (Henderson, 1990; 1991; 1993; Griffith et al., 1992). Our discovery of a novel family of cation–sugar transport proteins provided a new opportunity for application of these techniques, which has proved particularly fruitful in determining the molecular nature of cation and substrate recognition. In this article we illustrate this approach. It is generally applicable for exploiting the ligand binding sites of a variety of membrane proteins whose amino-acid sequences are known, but which are refractory towards crystallographic methods for determining 3D structure.

In secondary transport, a solute is translocated across the membrane in uniport or symport, or antiport with another solute. Most of the sensible secondary coupling mechanisms that one can imagine in terms of direction of transport, reaction stoichiometry, electrogenic/electrophoretic nature, and type of coupling solute have been discovered in the past two decades (Poolman and Konings, 1993). As more and more primary sequences of these proteins have become available it has become apparent that a single family of homologous transporters can be composed of uniporters, symporters and antiporters, while among the symporters and antiporters the reaction stoichiometries, type of coupling solute etc. can be different (Henderson, 1990; 1991). In fact, it has so far been impossible to infer from the primary sequences, information regarding the coupling mechanisms of the transporters. However, on the basis of structure predictions and various types of topological studies, the secondary structure and folding of the proteins within the membrane is probably very similar, even among members of different families. Although tertiary structures are not yet available, it is plausible that most of the secondary transporters are composed of twelve hydrophobic segments in α-helical configuration which span the membrane (Griffith et al., 1992; Henderson, 1993).

In order to relate a catalytic or kinetic property of a transporter to a specific region of the protein or even a particu-
lar amino acid, chimeric and single/multiple amino-acid-substitution proteins have been constructed. The largest number of mutants is available for the lactose-transport protein (LacY) of Escherichia coli (Kaback, 1990; 1992; Poolman and Konings, 1993), and the second best studied bacterial transporter in this respect is the melibiose-carrier protein (MelB) of E. coli (Leblanc et al., 1991; Poolman and Konings, 1993). While LacY obligatorily couples the transport of galactosides to protons (Kaback, 1990), the MelB protein can use either Na+, Li+ or H+ as coupling ions (Tsuchiya and Wilson, 1978; Wilson and Wilson, 1987). We have discovered that homologues of the E. coli MelB protein include transporters that facilitate the uptake of galactosides, pentoses or hexuronides (GPH family, see below), and use Na+, Li+, H+ and Li+ as coupling ions (Table 1). These different cation selectivities of the transporters and the varying degrees of primary sequence similarity in the aligned amino-acid sequences of this family, ranging from 18 to 85% identity between pairs of proteins, offer an enormous advantage in identifying locations/residues with a particular function. The study of mutants has resulted in the specification of at least some of the important ligand-selectivity determinants in the transporters, particularly in relation to cation coupling. A subgroup of the GPH family of transport proteins has a carboxy-terminal hydrophilic extension of approx. 160 amino acids, which forms the target for a novel type of regulation by the phosphoenolpyruvate:sugar phosphotransferase system (PTS). The unique features of this family of solute transporters are presented in this review, and the most recent findings pertinent to understanding their structure–activity relationships are explored.

### Sequence comparisons

The members of the galactoside–pentose–hexuronide transporter family (GPH) share sequence identity with each other over their entire length. The binary comparison scores expressed in standard deviations are $\geq 9$ ($P \leq 10^{-15}$), which is sufficient to establish that the proteins

---

**Table 1. Cation selectivity of members of the GPH family.**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>MelBEC</th>
<th>MelBEC</th>
<th>MelBKP</th>
<th>LacSST</th>
<th>GusBEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-Galactosides</td>
<td>H+, Na+, Li+</td>
<td>Na+, Li+</td>
<td>H+</td>
<td>H+</td>
<td>NT</td>
</tr>
<tr>
<td>β-Galactosides</td>
<td>Na+, Li+, (H+)</td>
<td>Na+, Li+</td>
<td>H+, Li+</td>
<td>H+</td>
<td>NT</td>
</tr>
<tr>
<td>Lactose</td>
<td>Na+, Li+</td>
<td>nd</td>
<td>Li+</td>
<td>H+</td>
<td>NT</td>
</tr>
<tr>
<td>Monosaccharides</td>
<td>H+, Na+, (Li+)</td>
<td>nd</td>
<td>nd</td>
<td>H+</td>
<td>NT</td>
</tr>
<tr>
<td>β-Glucuronides</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>H+</td>
<td></td>
</tr>
</tbody>
</table>

Melibiose, 6-O-α-galactopyranosyl-β-glucose; α-NPG, p-nitrophenyl-α-β-galactopyranoside; Methyl-α-Gal, methyl-α-β-galactopyranoside; TMG, methyl-β-α-thiogalactopyranoside; Lactose, 4-O-β-galactopyranosyl-α-glucose; nd, not determined; NT, not transported. Parentheses refer to poor coupling with the corresponding cation. Data were taken from Wilson and Wilson, 1987; Leblanc et al., 1991 (MelBEC); Tsuchiya and Wilson, 1978 (MelBEC); Hama and Wilson, 1992 (MelBKP); Foucaud and Poolman, 1992 (LacSST); W.-J. Liang, unpublished (GusBEC). Subscripts are defined in the legend to Fig. 2.
are derived from the same ancestor. The binary comparison scores for some of the members of the GPH family are provided by Reizer et al. (1994). Within the GPH family, subfamilies can be recognized (Fig. 1), i.e. the three MelB proteins share between 78 and 85% identical residues (‘MelB’ subfamily), the four proteins within the ‘LacS’ subfamily are 35–68% identical, whereas the identity between members of the ‘MelB’ and ‘LacS’ subfamily varies between 19 and 27%. The percentage identity between the members of the ‘LacS’ subfamily and the subfamily represented by GusB varies between 18–22%, whereas the ‘MelB’ and ‘GusB’ subfamilies share 24–28% identity. Portions of the multiple alignment of the sequences are shown in Fig. 2; these regions have high similarity scores and, in addition, mutant analysis has revealed their importance in translocation catalysis (see below).

The multiple sequence alignment was made by the CLUSTAL program in the PCGENE package (release 6.26, Genofit). The computer-generated alignments were optimized by minimizing the number of gaps using typical sequence motifs as a guide for the adjustments, and taking into account the charge-bias of the cytoplasmic loop regions as well as the propensities of amino acids to reside in transmembrane or loop regions (Von Heijne, 1986). In particular, gaps in the putative transmembrane segments were eliminated, which greatly improved the correspondence of the whole sequence alignment. The proposed positions of the transmembrane segments were also adjusted further in the light of any experimentally obtained topological information.

Secondary structure predictions and membrane topology

The membrane topology of the MelB protein has been analysed by fusing various portions of the protein to the compartment-specific reporter protein alkaline phosphatase (Botfield et al., 1992; Pourcher et al., in preparation). For most of the putative transmembrane segments, fusions have been constructed that were spaced at 3–4 amino-acid intervals. By assuming that fusions in the
outer half of transmembrane segments give rise to high alkaline phosphatase activities, whereas those in the inner half result in low activities (Calamia and Manoil, 1990), the membrane topology was fine-tuned and the middle of each transmembrane segment was defined (Pourcher et al., in preparation). In general, the hydropathy/alignment predictions corresponded well with the MelB–PhoA analysis; the adjustments needed resulted in a shift of 4–5 residues in transmembrane segments I and II, and pushed 8–9 residues at each end of interhelix loop 6–7 into the membrane. The cytoplasmic location of the carboxy-terminus of MelB is supported by vectorial labelling with antibodies (Botfield and Wilson, 1989). The secondary structure model of MelB, which is based on the multiple sequence alignment and the topological studies, is shown in Fig. 3. The residues (red) and charges (purple) that are conserved in all the sequences are indicated; the borders of the transmembrane segments are numbered according to the sequence of MelB. Edman degradation of the purified MelB protein has shown that the amino-terminus of MelB corresponds to SISMTTKL (Pourcher et al., 1995) rather than MTTKL as was assumed from the nucleotide sequence of melB (Yazyu et al., 1984). The amino-terminal methionine has probably been removed by post-translational cleavage. The amino-acid numbering of the E. coli MelB protein used throughout the manuscript is based on the newly proposed translation initiation site (the numbering is n + 4 relative to n in earlier papers concerning E. coli MelB). The translated sequence MSIS is also present at the amino-terminus of the Klebsiella pneumoniae and Salmonella typhimurium carrier proteins.

Using the topology information schematically represented for MelB in Fig. 3, the transmembrane segments of the members of the GPH family were examined for their patterns of hydrophobicity and residue substitution (Eisenberg et al., 1984; Donnelly et al., 1993). The amphipathy analysis detects periodicities in the hydrophobicities of amino acid sequences, whereas the substitution analysis discriminates between buried and lipid-facing residues; the latter method is based on the assumption that lipid-accessible residues are less conserved and have substitution patterns that are different from the inaccessible residues in water-soluble proteins (Donnelly et al., 1993). The alignments of the transmembrane regions, such as those shown for helix I, II, X and XI in Fig. 2, were used for the analysis. The periodicities in the hydrophobicity and substitution patterns were calculated for a window of 15, using a Fourier transform method; the substitution tables are derived from lipid-accessible residues (see Donnelly et al., 1993). Distinct patterns of amphipathy that are compatible with an α-helix, i.e. a period of 3.6 residues, are observed in the transmembrane segments II, IV, VII, VIII, IX, X, XI and XII, whereas similar residue-substitution periodicities are found in the transmembrane segments II, IV and XI (Fig. 4); the cut off for assigning significance to a particular periodicity is given by an α-index of 2.5. Heli-cal wheel projections of the transmembrane segments together with the arrows corresponding with the hydrophobic moments (red) and the predicted internal face of the helix (based on substitution periodicity; blue arrows) are shown in Fig. 4. The residues indicated are those of MelB protein, but the analysis is based on the complete alignment. For putative α-helices II and IV the hydrophobic

Fig. 3. Secondary structure model of MelB showing the positions of various critical residues. The colours are defined in Fig. 2. The residues that are essential for regulation of MelB by IIA^met^ are indicated in black.
Fig. 4. Helical wheel representations of the periodicity in hydrophobicity (red arrows, vector length indicates hydrophobic moment; scale from 0 to 3), and residue substitution pattern (blue arrows, internal face residues; scale from 0 to 1.2) in putative transmembrane α-helices of the members of the GPH family. The vectors are generated using the multiple sequence alignment, the predicted transmembrane segments (varying from 19–23 residues, see Fig. 3) and a window size of 15 residues. The 12 helices are arranged in an antiparallel fashion with the first residue of each helix starting at 0° in the top/right quadrant; the residues corresponding with the MelB sequence are depicted clockwise and counter-clockwise for the odd and even numbered helices, respectively. Circled residues refer to positions that have been implicated in substrate and/or cation specificity (see Fig. 2).
moments match with the face of the helix that is predicted to be in contact with the lipids, i.e. the orientations of the vectors for hydrophobicity and substitution pattern are opposite. Apparently, the packing of these helices is well conserved in each of the members of the GPH family so that the charged residues are excluded from the energetically unfavourable environment of the membrane. Although charged residues in regions of low dielectric constants such as the membrane could be stabilized by ion pairing (Honig and Hubbell, 1984), formed between residues in adjacent helices or within the same helix, the charged residues could also protrude into an hydrophilic ‘channel’ in which ligand binding and/or translocation catalysis takes place. Substitutions made for each of the charged residues in helix II (and IV) of both E. coli MelB and Streptococcus thermophilus LacS provide support for the latter suggestion as the results suggest that the acidic residues may contribute to the co-ordination of the coupling ion (see Intramembrane charged and polar residues).

Chimeras

Despite high identity in their primary sequences, the transporters from the ‘MelB’ subfamily exhibit significant differences in ionic selectivity during transport. For example, the MelB carrier from E. coli (MelBEc) uses Na⁺ as a coupling ion for melibiose (or methyl-β-D-thiogalactopyranoside, TMG) symport whereas the MelB carrier from K. pneumoniae (MelBkp) cannot use sodium (Table 1). This difference has been exploited to identify locations involved in cation recognition. In the first instance, Hama and Wilson (1993) constructed hybrid MelB carriers by swopping various portions of MelBEc and MelBkp and showed that the amino-terminal 81 residues of the E. coli protein are essential for sodium coupling to melibiose transport (Table 2). The amino-terminal 81 residues of the E. coli MelB protein alone, however, may not be sufficient to make up a complete binding/translocation site. Indeed, chimeras composed of the amino-terminal 302 or 353 residues of MelBEc (and the remaining carboxy-terminal portions of MelBkp) are ineffective in coupling Na⁺ to transport (Hama and Wilson, 1993). This suggests that additional residues from the carboxy-terminal end of MelBEc are also required for Na⁺ recognition. Finally, it should be recalled that the ionic selectivity profile of the symporters is determined to some extent by the configuration of the transported sugar, implying that cation and sugar binding are co-operative. It may thus be anticipated that alteration of sugar binding to the enzyme may indirectly modify the ionic selectivity or ligand affinity of the transporter. This would explain why Na⁺ activation of melibiose transport with some of the MelBEc→MelBkp chimeras is not or only poorly observed with TMG or lactose as substrate.

In contrast to the three members of the ‘MelB’ subfamily, neither the S. thermophilus galactoside-transporting protein LacS nor the E. coli glucuronide-transporting protein GusB uses Na⁺ or Li⁺ (alkali metal ions) as coupling species (Foucaud and Poolman, 1992; Poolman et al., 1995b; W.-J. Liang, unpublished observations). Most of the MelBEc→LacSst chimeras, constructed with the purpose of changing the ion selectivity profile of the LacS protein, were inactive or poorly expressed (J. Knol, B. Poolman, T. Poucher and G. Leblanc, unpublished results). It should, however, be mentioned that one of the constructs with the amino-terminal 78 residues from MelB (helix I and II) and the remaining residues from LacS (helix III–XII) was unable to catalyse uphill (coupled) transport but facilitated downhill uptake and equilibrium exchange of lactose with near wild-type rates (Table 2). As only 23% of the residues in MelB and LacS are identical, this result suggests that packing of these helices is quite flexible, allowing a large number of ‘mismatches’ without compromising galactoside translocation per se.

**Table 2.** Cation selectivity of MelB and LacS mutants.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Allele</th>
<th>H⁺-coupling</th>
<th>Na⁺-coupling</th>
</tr>
</thead>
<tbody>
<tr>
<td>MelBEc</td>
<td>Wild-type</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>D55C</td>
<td>reduced</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>D55E</td>
<td>+</td>
<td>reduced</td>
</tr>
<tr>
<td></td>
<td>D59N/C</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>D59E</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>D124N/C</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>D124E</td>
<td>reduced</td>
<td>+</td>
</tr>
<tr>
<td>LacSst</td>
<td>Wild-type</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>E67Q</td>
<td>conditional</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td>D71N</td>
<td>—</td>
<td>nd</td>
</tr>
<tr>
<td></td>
<td>D71E</td>
<td>+</td>
<td>nd</td>
</tr>
<tr>
<td>MelBkp</td>
<td>Wild-type</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>A58N</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>N₂MelBEc/C₅₋MelBkp⁺</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>N₂MelBEc/C₅₋LacSst⁺</td>
<td>—</td>
<td>nd</td>
<td></td>
</tr>
</tbody>
</table>

**a.** Amino-terminal 81 residues of MelBEc fused to the carboxy-terminal 389 residues of MelBkp.

**b.** Amino-terminal 78 residues of MelBEc fused to the carboxy-terminal 444 residues of LacSst.

+ , coupled transport (comparable to wild-type); — , no coupling with particular ion; reduced, coupling is observed but less than in wild-type; conditional, coupling is highly dependent on temperature; nd, not determined. Data were taken from Hama and Wilson, 1993; 1994; Pourcher et al., 1993; Zani et al., 1993; 1994; Wilson and Wilson, 1994; J. Knol and B. Poolman, unpublished results.

Dissection of individual amino acid residues implicated in cation and/or sugar binding and transport catalysis

**Intramembrane charged and polar residues**

Transporters of the GPH family have several negatively and positively charged residues in different amino-terminal
helices, some of which are highly conserved, particularly in helix II (Fig. 2). The catalytic role of several of these residues in this region has been examined by making various substitutions in both the *E. coli* MelB and the *S. thermophilus* LacS carrier proteins (Table 2). The most important results can be summarized in the light of the hypothetical structural model of the amino-terminal region of MelB (Fig. 5). Neutral residues at MelB positions 19 (helix I), 55 and 59 (helix II) and 124 (helix IV) lead to (i) loss of Na\(^+\)-linked TMG transport, and (ii) retention of binding of \(\beta\)- and \(\alpha\)-galactosides, but now independent of sodium ions (Pourcher et al., 1993; Zani et al., 1994; our unpublished results). The corresponding Asp to Glu substitutions impair transport but sodium activation is still observed, albeit with lower efficiency (Zani et al., 1994). Comparatively, neutral substitutions of Asp-35 (near the periplasmic extremity of helix I) or Glu-100 (interhelix loop 3–4) do not seriously affect the ion selectivity properties of the transporter. In keeping with the observations on the MelB chimeras, these results suggest that the amino-terminal region of MelB is critical for recognition of the coupling ions. As mutagenesis of Asp-19, Asp-55, Asp-59 and Asp-124 has strikingly similar effects, and these residues are located in the cytoplasmic halves of the (amphipathic) helices I, II and IV, we speculate that the side-chains of these amino acids form part of a network involved in the co-ordination of Na\(^+\), Li\(^+\) or H\(_3\)O\(^+\). It is important to recall that the ionic radii of Na\(^+\) and Li\(^+\) (0.95 and 0.65 Å, respectively; Glusker, 1991; Hille, 1990) are similar to the bond length of O-H in H\(_2\)O\(^+\) (i.e. approx. 0.9 Å; Behr et al., 1982), which is consistent with the binding of Na\(^+\), Li\(^+\) and H\(_3\)O\(^+\) to the same site (Boyer, 1988). Depending on the protonation state of the carboxylates and the number of oxygens binding to the cation, the ion is probably surrounded by more than one negative charge. Therefore, it is possible that additional cationic residues interact with the carboxylates as well (see below).

Several properties of the wild-type and mutant MelB transporters are consistent with the hypothesis presented above. First, earlier work has indicated that H\(^+\) and Na\(^+\) (or Li\(^+\)) compete for the same cationic binding site on MelB (Damiano-Forano et al., 1986). It is probable that competition between H\(^+\) and the alkali metal ions occurs, at least partly, at the level of co-ordinating ligands to H\(_3\)O\(^+,\) Na\(^+\) or Li\(^+.\) Support for this idea comes from studies of crown-ethers (Behr et al., 1982; Glusker, 1991), and of the alkali-metal ion-activated dialkylglycine decarboxylase (Toney et al., 1993; Hohenester et al., 1994), which indicate that the co-ordination number for the different ions is variable. (The co-ordination number is the number of groups or atoms that surround a metal ion; these groups are chemically bonded to the metal by donating an electron pair to the bond.) If the same suggestion is valid for the binding of cations by MelB, one can expect differences in ion-selectivity for the various Asp mutants (Zani et al., 1994; Wilson and Wilson, 1994). Indeed, neutral substitutions at MelB positions 19, 59 or 124 affect both Na\(^+\) and H\(^+\)-dependent transport, whereas similar substitutions at position 55 do affect Na\(^+\) coupling but the mutant enzymes retain the capacity to transport the sugar with H\(^+\). Second, mutagenesis of individual highly conserved polar residues located in the vicinity of Asp-19, 55, 59 and 124 has less drastic effects on cation binding or coupling efficiency than substitution of the acidic residues. At most, the transport activity and the binding affinity for Na\(^+\) are lowered in these mutants (Zani et al., 1994). Incidentally, the defects are most pronounced for mutated residues located at a depth close to that of the crucial acidic residues, which supports the view that the site involved in cation binding is buried in the inner half of the molecule.

Following the analysis of the MelB<sub>EC</sub>/MelB<sub>KP</sub> chimera, individual residues in helix I and II of MelB<sub>KP</sub> were replaced by residues corresponding to the MelB<sub>EC</sub> sequence (Hama and Wilson, 1994). These studies showed that replacing Ala-58 by Asn in helix II of MelB<sub>KP</sub> is sufficient to convert the Na\(^+\)-independent transporter into a Na\(^+\)-dependent one. It seems very significant, therefore, that only those transporters of the GPH family that have this Asn residue at an equivalent position (MelB<sub>EC</sub> and MelB<sub>SY</sub>, Fig. 2) use Na\(^+\) as coupling ion (Table 1). The requirement for an Asn residue at position 58 in *K. pneumoniae* MelB, in order to use Na\(^+\) as coupling ion, could be a result of a proper positioning of Asp-59. Alternatively, the side-
chain of the Asn residue itself could form one of the co-
ordinates in the binding of Na⁺, an interaction which is not essential when Li⁺ or H⁺ (H₂O⁺) is used as the coupling ion. The role of the Asp₅₅-Asp₅₉ pair in MelB is reminiscent of the X-Glu₆₇-Ser₇₁-Y motif in subunit c of the E. coli Fₒ-F₁-ATPase, which enables the enzyme to bind Li⁺ (Zhang and Fillingame, 1995). It is suggested that the polar residues in the X-Glu-Ser-Y motif of the mutant enzyme donate the major ligand groups necessary for Li⁺ binding, i.e. Li⁺ is co-ordinated by the two glutamate carboxylate oxygens, the glutamate-backbone carbonyl and the serine side-chain oxygen. We speculate that a similar cation-binding pocket may be present in the members of the GPH family and that Asn-58 and Asp-59, in addition to Asp-19, 55 and 124, are crucial for the co-ordination of Na⁺, whereas the binding of protons (or hydro-
dium ions) allows more flexibility of the binding pocket (see also Tolner et al., 1995).

On the basis of topological constraints imposed by the proposed structural model (Fig. 5), the helices II and IV should be in close proximity. Recent analysis of a second-site revertant supports this view (Wilson et al., 1995). The mutant phenotype of MelB D₅₅S can be restored by a second mutation, i.e. Gly-117 to Asp, which allows the enzyme to again use Na⁺ and Li⁺ as coupling ions. If one looks at the position of these residues in the membrane it is apparent that Gly-117 is close to the middle of helix IV and at a similar depth to the position of Arg-52 in helix II (Fig. 3). The latter residue is on the same face of the helix as Asp-55 (Fig. 4). In view of the proposed role of Asp-55 in cation binding, one could envisage that the carboxylate of Asp-117 substitutes for the missing ligand in the D₅₅S mutant. Alternatively, if one assumes that the negative charge of Asp-55 is compensated by the positive charge of Arg-52, e.g. by forming a salt-bridge (Fig. 5), it is possible that a similar interaction is formed between Arg-52 and Asp-117 in the double mutant D₅₅S/G117D. Removing one of these charges leaves the opposing charge unpaired and is likely to be detrimental for cation binding and/or translocation activity, even though the residues may not be obligatory for these functions per se. Indeed, replacing Arg-52 by Ala decreases the transport activity of E. coli MelB by 95% without any effect on the expression level. The residual transport of melibiose, however, is still affected by Na⁺ and Li⁺ (Zani et al., 1994). Interestingly, the equivalent of Gly-117 in the MelB pro-
teins is an aspartic acid in the members of the 'LacS' sub-
family (Asp-133 in LacS), which would allow the corresponding Arg (Arg-64 in LacS) to form a salt-bridge with either Glu-67 or Asp-133.

It should be stressed that the acidic residues in helix II are highly conserved in all members of the GPH family. The equivalents of Asp-55 and 59 in the LacS protein of S. thermophilus (Glu-67 and Asp-71) have been replaced with neutral and acidic residues. Because LacS couples the transport of galactosides exclusively to protons (Fou-
caud and Poolman, 1992), sodium activation of transport is observed in neither the wild-type nor the mutant pro-
teins. Strikingly, however, substitution of Glu-67 by a neutral amino acid (E67Q) results in a conditionally uncoupled phenotype of LacS, i.e. galactoside accumulation is re-
duced to approx. 10% at 37°C but is normal at 25°C, whereas equilibrium exchange is comparable to that of the wild-type at both temperatures (J. Knol, P. Fekkes and B. Poolman, manuscript in preparation). This indi-
cates that Glu-67 ensures a high coupling efficiency but, similar to MelB(D55C), the residue is not essential for the sugar-H⁺ symport reaction. The LacS(D71N) mutant is devoid of any galactoside-H⁺ symport activity, although the mutant protein is normally expressed and facilitates equilibrium exchange with near wild-type activity. These observations reinforce the suggestions made for MelB that the conserved acidic residues in helix II have different roles in the cation coupling to sugar transport. It seems that the carboxylate moiety of Asp-59 in MelB (and of Asp-71 in LacS) is obligatory for coupled transport, possi-
bly by donating one (or more) of the ligands for cation bind-
ing and, at the same time, by ensuring tight coupling of the transport reaction, whereas the carboxylate of Asp-55 in MelB (and of Glu-67 in LacS) only serves for ion selectivity and/or structural organization of the site.

In conclusion, these findings support the contention that the residues of the hydrophilic faces and cytoplasmic halves of helices I, II and IV of MelB, and probably also of the homologous proteins of the GPH family, form at least part of the cation-binding site. Whereas some resi-
dues may contribute directly to the catalytic activity, others may have a role in the proper positioning of the catalytic residues.

Interhelix loop 10–11

Both the amphipathic nature of helix II, and the high sequence conservation in interhelix loop 10–11 (and helix XI) are features of members of the GPH family. Transmem-
brane segment XI that flanks this loop has, similar to puta-
tive α-helices II and IV, a strong periodicity in terms of lipid/ non-lipid facing residues (Donnelly et al., 1993) which suggests that the 'non-lipid site' of the helix forms part of the core of the protein that interacts with other hydrophilic regions and/or the carrier ligands.

Although the lactose transport protein (LacY) of E. coli is not homologous to members of the GPH family, part of LacY bears some resemblance to a stretch of approx. 20 amino acids in interhelix loop 10–11 of the LacS members of the GPH family (sequence motif Lys-x-x-His-x-x-Glu; Poolman et al., 1992; 1995b). Among the conserved resi-
dues are a histidine (322 in LacY, 376 in LacS) and a glu-

© 1996 Blackwell Science Ltd, Molecular Microbiology, 15, 911–922
tamic acid residue (325 in LacY, 379 in LacS), which have similar properties in both of the galactoside transport proteins (Kaback, 1992; Poolman and Konings, 1993). For instance, glutamic acid-325 in LacY is believed to participate directly in galactoside-coupled proton translocation because sugar-H+ symport is completely defective in neutral substitution mutants, whereas uncoupled sugar transport (uniport) can still be observed (Carrasco et al., 1986; 1989). Similar observations have been made for the E379A and E379Q mutants of LacS, although some coupled transport can be observed at low pH (Poolman et al., 1995b). In fact, in both proteins, the substitution of the acidic residue by a neutral residue results in a leak pathway in which the binary enzyme–substrate complex has become mobile (ES leak) relative to the ternary enzyme–substrate–proton (ESH) complex. The properties of the Glu-379 mutants can be simulated by assuming that the pK for proton binding on the outside is lowered from 10 to 6, and by introducing an ES leak pathway (Poolman et al., 1995b). The lowering of the pK\text{OUT} suggests that the wild-type enzyme, the pK of the proton translocating residue(s) is modulated by Glu-379. In terms of the proton pathway, this would imply that Glu-379 raises the pK of an essential protonatable residue when the carrier is in the conformation with the binding sites facing outwards. The properties of the acidic residues in helices I, II and IV are consistent with site(s) that are on the pathway of proton (hydronium) transport and are possibly modulated by Glu-379. The lowering of pK in the mutants is consistent with the nature of the substitutions made (Glut-379 → Gln and Glut-379 → Ala), i.e. a negative charge raises the pK of a nearby group by stabilizing the protonated form. By removing the negative charge by mutation, or by a conformational change that is associated with the exposure of the binding sites to the inside, the pK is expected to be lowered.

Substitutions at Glu-365 in MelB, which is the equivalent of LacY(E325) and LacS(E379), also impair galactoside uptake but H+, Na+ (or Li+) -coupled transport does occur albeit with \( k_{\text{CAT}} \) values which are more than 10-fold reduced (Leblanc et al., 1991). Some of the mutations at this position render MelB highly temperature-sensitive but, at present, it is not clear whether Glu-365 in MelB serves a role in modulating the pK of an essential residue in the cation binding site or in the translocation pathway.

Each of the histidines of E. coli MelB and S. thermophilus LacS has been substituted by glutamine (and arginine) in order to establish whether these residues fulfill a catalytic role in the transport reaction (Pourcher et al., 1992; Poolman et al., 1992). Whereas none of the histidines in MelB is important for the coupled transport of sugars, the H376Q substitution in interhelix loop 10–11 of LacS results in a lowering of the coupling efficiency of the transporter. The ability of LacS(H376Q) to accumulate sugars against a concentration gradient is completely lost with TMG as substrate, reduced by more than 80% with melibiose but only moderately affected with lactose as substrate. Two other histidines that are present in the IIA domain of LacS (see below) are crucial for the regulation of transport activity by the PTS (Poolman et al., 1995a), but these residues are not needed for galactoside transport per se (Poolman et al., 1992).

Although the analysis of the alkaline-phosphatase fusions of MelB suggests that interhelix loop 10–11 is on the inner surface of the membrane (Botfield et al., 1992; Pourcher et al., 1995a), the properties of the LacS position 376 and 379 mutants are more consistent with a location in the interior of the protein (Poolman et al., 1995b). The proposed location of the equivalent residue of LacY is in the middle of putative \( \alpha \)-helix X (Calamia and Manoil, 1990). If one assumes that interhelix loop 10–11 does not protrude from the membrane (as indicated in Fig. 3), and is possibly in contact with residues of the flanking helix XI (and others), it is likely that alkaline-phosphatase fusions locally disrupt the structure giving rise to an inside location of PhoA at these junction points. The reporter-protein-based topology analysis is clearly inappropriate for detecting subtleties in protein structure such as the location of protein segments that do not span the membrane entirely.

**Selection of mutants**

Various mutants have been isolated using random mutagenesis techniques in combination with selection for specific phenotypes. Strategies for isolating MelB mutants defective in cationic coupling took advantage of the fact that melibiose transport is inhibited by high concentrations of Li+, which results in inhibition of growth when melibiose is the sole source of carbon and energy (Tsuchiya and Wilson, 1978). Li+-resistant mutants, which have lost the ability to transport with H+ but retain the ability to couple galactoside uptake to Na+ (Kawakami et al., 1988), have the following single-site mutations: P146S (interhelix loop 4–5), L236F (helix VII), A240T or A240V (helix VII) (Figs 2 and 3; residues indicated in blue). Isolated Li+-dependent mutants exhibit Na+- and Li+-melibiose co-transport but cannot accept H+ as a coupling ion (Niiya et al., 1982; Tsuchiya et al., 1983; Shiota et al., 1984). In five independently isolated Li+-dependent MelB mutants, Pro-126 (interhelix loop 4–5) was replaced with Ser (Yazyu et al., 1985).

A total of 70 melibiose carrier mutants have been isolated upon growth on melibiose in the presence of an excess of the non-metabolizable galactoside TMG (Botfield and Wilson, 1988). Except for I65V, these mutants are impaired in TMG-recognition properties and resistant to Li+. The amino acid substitutions of the 70 selected ‘TMG-resistant’ mutants occurred at 18 unique positions
within the protein and are clustered in three distinct regions, i.e. the amino-terminal end of \( \alpha \)-helix I, \( \alpha \)-helix IV, and the majority occurring in interhelix loop 10–11; other substitutions are located at position 65 (I65V, interhelix loop 2–3) and 240 (A240T and A240V in helix VII) (see Figs 2 and 3). Although none of the selected mutants has been studied in great detail, the observation that more than 90% of the mutants are in or near the four amino-terminal \( \alpha \)-helices and in interhelix loop 10–11 further strengthens the suggestion that these regions are in close proximity in the tertiary structures of the proteins and may constitute the cation and substrate binding sites. Finally, the majority of the mutants isolated on the basis of TMG-resistance, Li\(^+\)-resistance or Li\(^+\)-dependence, but also the site-directed mutants Y28F (helix I), D59E (helix II), N87A (helix III), Y120F and D124E (helix IV), exhibit simultaneous alterations in cation and substrate recognition (Shiota et al., 1984; Kawakami et al., 1988; Zani et al., 1994).

Regulation of transport

Transport of melibiose in enteric bacteria is regulated not only at the level of gene expression (catabolite repression) but also at the level of enzyme activity (catabolite inhibition or inducer exclusion) (Saier, 1989; Postma et al., 1993). The dual regulation allows an instantaneous response of the organism to the presence or absence of melibiose (inducer exclusion) and a slow response, which involves switching on/off the expression of certain genes (catabolite repression). The phosphoenolpyruvate:sugar transferase system is involved in both processes, and the regulation is mediated by the phosphorylation state of the phosphorl transfer protein (IIA\(_\text{Glc}\), previously enzyme III or II\(_\text{Glc}\)). The inducer-exclusion mechanism involves the (stoichiometric) binding of IIA to the MelB carrier protein, which results in an inhibition of transport; a number of other transporters/enzymes are affected in a similar fashion (Saier, 1989; Roseman and Meadow, 1990; Postma et al., 1993). Sequence analysis of inducer-exclusion-resistant mutants suggests that IIA binds to the carboxy-terminus (residues 438, 441 and 445 of the S. typhimurium enzyme), and, interestingly, to interhelix loop 10–11 (residue 368) of MelB (Kuroda et al., 1992; Postma et al., 1994).

The LacS protein of S. thermophilus is also affected by the PTS but differently from the inducer exclusion mechanism that affects MelB. The lactose transport proteins (LacS) of S. thermophilus, Lactobacillus bulgaricus and Leuconostoc lactis, and RapF of Pediococcus pentosaceus (‘LacS subfamily’, Fig. 1) differ from the other secondary transport proteins by having an extra domain, attached to the carboxy-terminal end of the carrier domain, that is homologous to IIA of various PTS (Poolman et al., 1989). The IIA domain has several structural features in common with the corresponding PTS proteins and can be phosphorylated in the presence of PEP, enzyme I and heat-stable protein HPr (Poolman et al., 1992). Phosphorylation of the S. thermophilus LacS IIA domain results in an inhibition of transport (Poolman et al., 1995a). Assuming that the IIA domains in the other secondary transporters within the ‘LacS subfamily’ are regulated similarly, it seems that most members of the GPH family are affected by the PTS, either by interaction with unphosphorylated IIA or phosphorylation of an endogenous IIA domain. Whether glucuronide transport in E. coli and xylose transport in Lactobacillus pentosus is regulated by the PTS is not yet known.

Conclusions and future directions

Before the era of gene sequencing it was perhaps expected that functional differences among secondary transporters, such as direction of transport (symport versus antiport), substrate and coupling-ion specificity, etc., would ultimately be recognized in the primary sequences of the proteins. At present it is clear that such mechanistic differences can be the result of only ‘minor’, subtle differences between the primary structures of the proteins that are simply not detected by sequence comparisons. From our concerted study of homologous transporters of the GPH family, as well as mutant analysis, it is clear that discrete regions/residues participate directly in cation and/or substrate recognition. For kinetic reasons it is difficult to determine, a priori, whether it is cation recognition or substrate recognition that is primarily affected unless it is technically possible to measure the binding of both as well as translocation (King and Wilson, 1990). The sequence alignment of the homologous transporters, together with topological information, shows that these regions form the most conserved parts of the proteins, but do not constitute a single linear sequence (simple sequence motif).

It has been shown that the cation-selectivity of the secondary transporters can be changed by only a single amino-acid substitution, e.g. Ala-58 to Asn in MelB\(_\text{KP}\) and Asp-55 to Cys in MelB\(_\text{EC}\), and that a solute-H\(^+\) symporter can be converted by a single amino acid change into a uniporter, e.g. Glu-379 to Gln or Ala in the S. thermophilus LacS. Questions that remain unanswered relate to: (i) the minimal number and kind of mutations that are needed to alter the direction of transport (e.g. the ability to convert a symporter into an antiporter; Konings et al., 1992); (ii) the high selectivity of some transporters for a particular cation whereas others such as MelB and SGLT1, the Na\(^+\)-glucose co-transporter of animal cells (Hirayama et al., 1994), are promiscuous with regard to cation coupling; (iii) the residues that interact with the car-
bohydrate and the type of atomic interactions, e.g., in the binding proteins of enteric bacteria the interactions are predominantly formed by hydrogen bondings between hydroxyl groups of the sugar ligand and the side-chains of polar amino acids (Quiocio, 1990). Finally, it will be essential for a more detailed understanding of the catalytic mechanism of secondary transporters that the three-dimensional structures of the proteins are solved. We have recently succeeded in overexpressing three members of the GPH family, i.e. MelB and GusB of E. coli and LacS of S. thermophilus, and purified each of these proteins to > 98% purity in milligram quantities using simple chromatography protocols (Pourcher et al., 1995; our unpublished results). Spectroscopic and/or crystallographic methods are presently used to determine the structure of these transporters and also to obtain additional insight into the structure–function relationships. Illustration of this future line of research is given by a recent fluorescence spectroscopy study of purified MelBEC, which has shown that sugar and cation binding is co-operative and induces conformational change(s) in the transporter (Mus-Veteau et al., 1995).

Acknowledgements

We wish to thank W. N. Konings for critical reading of the manuscript, and W. M. de Vos, E. Vaughan, R. Leer and S. Hastrup for providing sequence information prior to publication.

References