Journal of Bioenergetics and Biomembranes, Vol. 22, No. 4, 1990

MINI-REVIEW

Proton-Linked Sugar Transport Systems in Bacteria

Peter J. F. Henderson[1]

Received January 12, 1990

Abstract

The cell membranes of various bacteria contain proton-linked transport systems for D-xylose,[2] L-arabinose, D-galactose, D-glucose, L-rhamnose, L-fucose, lactose, and melibiose. The melibiose transporter of *E. coli* is linked to both Na^+ and H^+ translocation. The substrate and inhibitor specificities of the monosaccharide transporters are described. By locating, cloning, and sequencing the genes encoding the sugar/H^+ transporters in *E. coli*, the primary sequences of the transport proteins have been deduced. Those for xylose/H^+, arabinose/H^+, and galactose/H^+ transport are homologous to each other. Furthermore, they are just as similar to the primary sequences of the following: glucose transport proteins found in a Cyanobacterium, yeast, alga, rat, mouse, and man; proteins for transport of galactose, lactose, or maltose in species of yeast; and to a developmentally regulated protein of Leishmania for which a function is not yet established. Some of these proteins catalyze facilitated diffusion of the sugar without cation transport. From the alignments of the homologous amino acid sequences, predictions of common structural features can be made: there are likely to be twelve membrane-spanning α-helices, possibly in two groups of six; there is a central hydrophilic region, probably comprised largely of α-helix; the highly conserved amino acid residues (40–50 out of 472–522 total) form discrete patterns or motifs throughout the proteins that are presumably critical for substrate recognition and the molecular mechanism of transport. Some of these features are found also in other transport proteins for citrate, tetracycline, lactose, or melibiose, the primary sequences of which are not similar to each other or to the homologous series of transporters. The glucose/Na^+ transporter of rabbit and man is different in primary sequence to all the other sugar transporters characterized, but it is homologous to the proline/Na^+ transporter of *E. coli*, and there is evidence for its structural similarity to glucose/H^+ transporters in Plants. *In vivo* and *in vitro* mutagenesis of the lactose/H^+ and melibiose/Na^+ (H^+) transporters of *E. coli* has identified individual amino acid residues alterations of which affect sugar and/or cation recognition and parameters of transport. Most of the bacterial

[1]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, United Kingdom.

[2]The D- or L-configurations of sugars will be omitted unless ambiguity can arise.

0145-479X/90/0800-0525$06.00/0

transport proteins have been identified and the lactose/H^+ transporter has been purified. The directions of future investigations are discussed.

Key Words: Sugar transport; membrane protein structure; homologous proteins; chemiosmotic therapy; sugar/proton symport.

Introduction

In 1963 Peter Mitchell suggested that the uptake of sugars into microbial cells might be energized by a trans-membrane proton gradient (Mitchell, 1963). The idea required that an individual transport system catalyze the simultaneous translocation of protons with a sugar molecule, "symport," or the experimentally indistinguishable "antiport" of hydroxyl ions, so that energy released by respiration or ATP hydrolysis and "stored" as the electrochemical gradient of protons could drive accumulation of the nutrient (Mitchell, 1961, 1963, 1973). The principle is illustrated in Fig. 1. However, this brilliant prediction remained untested until 1970, when Ian West devised experimental conditions in which the movement of lactose or substrate analogues into cells

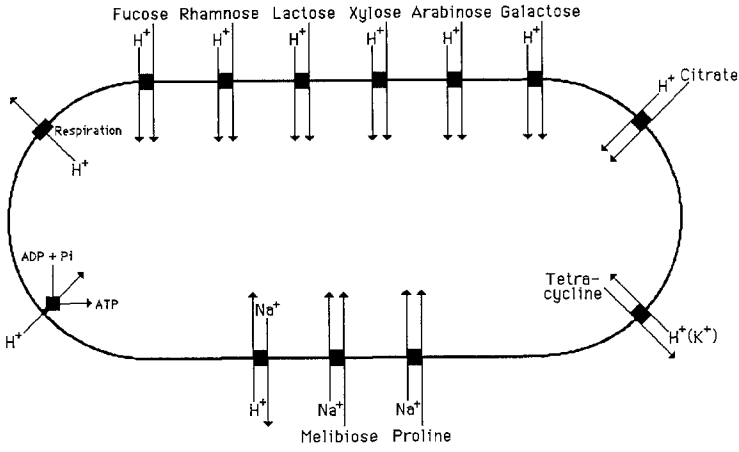

Fig. 1. An illustration of ion-linked transport systems in *Escherichia coli*. The large oval represents the cytoplasmic membrane of the microorganism. A trans-membrane electrochemical gradient of protons is generated by respiration, depicted on the top left. This gradient is used for the energization of ATP synthesis (left) and transport as follows. A series of different sugar/H^+ symport proteins (top) catalyze accumulation of their individual substrate at the expense of the proton gradient. A similar mechanism operates for citrate uptake (top right) and for some other nutrients (not shown). By contrast, a tetracycline/H^+ antiport protein (bottom right) catalyzes efflux of tetracycline at the expense of the proton gradient. Similarly, an Na^+/H^+ antiporter catalyzes efflux of Na^+ ions at the expense of the proton gradient (bottom). The resulting trans-membrane electrochemical gradient of sodium ions is used to energize melibiose uptake or proline uptake via individual substrate/Na^+ symport proteins (bottom). References are given in the text.

of *Escherichia coli* containing the lactose transport protein (Lac Y) evoked an alkaline pH change showing proton movement in the same direction (West, 1970; West and Mitchell, 1972, 1973). Since then the expression of the *lacY* gene has been amplified (Teather *et al.*, 1978), the protein has been solubilized and purified (Newman and Wilson, 1980; Newman *et al.*, 1981), its structure has been explored by immunological and chemical methods (Carrasco *et al.*, 1986; Page and Rosenbusch, 1988; Kaback, 1989), and its molecular mechanism has been investigated by mutagenesis (Kaback, 1987, 1989; Roepe *et al.*, 1990; Brooker, 1990).

The sugar/H^+ symport mechanism was originally conceived to account for the respiration dependence and uncoupler sensitivity of the lactose transport system of *E. coli* (Kennedy, 1970). More recently, other proton-linked transport systems for L-arabinose, D-xylose, D-galactose, L-fucose, and L-rhamnose have been discovered in *E. coli* and several of the Enterobacteriaceae, for D-glucose in Cyanobacteria, and for lactose in Streptococci. Remarkably, some of these proteins are closely related to sugar transport proteins occurring in eukaryotes as diverse as yeasts, protozoa, rat, and man (Maiden *et al.*, 1987; Baldwin and Henderson, 1989; Henderson and Maiden, 1990).

The relationship between prokaryote and eukaryote sugar transport proteins will be the main theme of this article, which will also examine selected properties of the bacterial sugar/H^+ transport proteins—their substrate specificities, the effects of certain inhibitors, their genes, and the amino acid sequences of the proteins. The recurrence of certain motifs in the primary sequences is highlighted, together with the presence or absence of residues that might be expected to interact with the sugars, the proton, or the inhibitors. A similarity in structure is proposed between these proteins and others that transport the apparently unrelated substrates, citrate and tetracycline, and a unifying two-dimensional model is presented for the structure of all the homologous proteins in the membrane.

Additional mechanisms exist for energizing the transport of sugars and other nutrients into *E. coli*, for example the phosphotransferase systems, ATP-dependent binding protein systems, and hexosephosphate/P_i antiports. These are the subjects of other reviews in this volume (see also Henderson, 1986; Furlong, 1987; Lengeler *et al.*, 1990; Kornberg, 1990; Quiocho, 1990; Higgins *et al.*, 1990; Maloney, 1990). Studies on the lactose/H^+ and melibiose/Na^+ transporters are relevant, but they have been well reviewed by Wilson *et al.* (1986), Kaback (1986, 1987, 1989), and Leblanc *et al.* (1989). Accordingly, only selected aspects of their properties will be considered. For the reader who wishes to follow the progress of the sugar molecules into the metabolic pathways after the initial transport event, two reviews by Cooper (1986) and Lin (1987) are highly recommended.

The Number of Proton-Linked Sugar Transport Systems in Bacteria

Experimental Criteria

Measurement of Sugar/H^+ Symport. A proton-linked sugar transport system can be most readily identified in the following type of experiment, in which uphill proton movement is measured in response to an inwardly directed substrate gradient. A concentrated suspension of bacteria is depleted of endogenous substrate and ATP, and resuspended under anaerobic conditions to prevent respiration. A relatively high concentration (1–5 mM) of sugar is added. If the sugar is a substrate for a sugar/H^+ symport system present in the cells (these have usually been grown in the presence of a suitable inducer), it diffuses inward down the concentration gradient. This results in an alkaline pH change in the medium, reflecting the movement of protons with the sugar. The experimental details are described by West (1970) and by Henderson and Macpherson (1986). Quite often the initial alkaline pH change, which is rarely greater than 0.1 pH units, is obscured by subsequent metabolism of the sugar to acidic end-products. Where this is the case, the use of appropriate nonmetabolized sugar analogues (below), or of mutants impaired in a metabolic enzyme, can facilitate detection of the pH change (Henderson and Macpherson, 1986).

Measurement of Respiration-Energized Transport in Vesicles. The second method requires the manufacture of right-side-out subcellular vesicles by the method of Kaback (1972, 1974). The use of vesicles has several advantages: the binding-protein and phosphotransferase types of transport system are not operative; the metabolic enzymes are virtually absent; and vesicles are much more sensitive than intact cells to ionophore and protonophore reagents that modify the electrochemical gradient of protons across the membrane. Transport of radioisotope-labelled sugars is energized by respiration (ascorbate plus tetramethylphenelenediamine is most effective, Kaback, 1974). No net pH change occurs during respiration-energized transport, but the sensitivity to protonophores and ionophores reveals its proton-linked nature. Sugar-promoted pH changes can be observed in such vesicles under de-energized conditions (Patel *et al.*, 1982; Horne and Henderson, 1983), but the amounts of vesicles required and their relative "leakiness" to protons means that this assay is more conveniently practised on intact cells.

Other Assays. This type of transport protein will also catalyze facilitated diffusion of radioisotope-labelled sugar across the membrane, either into or out of the cells/vesicles (net influx or efflux, respectively; Viitanen *et al.*, 1986; Wilson *et al.*, 1986). Various counterflow experiments can be devised, in which the protein catalyzes exchange of the same or different labelled/unlabelled substrates across the membrane, each at the same or

different concentrations (Viitanen *et al.* 1986; Walmsley, 1988). A particularly useful variant of this is the "overshoot" experiment (Wilson *et al.*, 1986; Viitanen *et al.*, 1986; Stein, 1986), because it is very sensitive and therefore valuable in reconstitution and purification assays, where it is not convenient to energize transport and only small amounts of transport protein may be available.

Value of Different Assay Methods. The existence of a repertoire of assay methods is invaluable, because it constitutes a series of "partial" reactions with which to examine the activities of mutants. In the absence of a means of determining the three-dimensional structure of the protein, it is essential to show that this has not been seriously compromised in a mutant by showing that at least one of the assayable activities is unimpaired while characterizing those others that have been changed (see, e.g., Kaback, 1987; Roepe *et al.*, 1990).

Unfortunately, these assays are often performed in a semiquantitative way, expressing activities as percentages of those measured in wild-type controls. Measurements of reproducibility are not always given, and advantage is not taken of computerized methods of calculating best-fit values of kinetic "constants" and their standard deviations (Cleland, 1979; Henderson, 1985). It is always desirable to extrapolate the determination of a K_m and k_{cat} value to conditions where all substrates, i.e., the sugar *and* the cation, are saturating (Cleland, 1979; Fromm, 1979). It is essential to measure k_{cat}, and not just V_{max}, especially in mutants or recombinants, where the level of expression of a protein may well have changed by a factor of 2 or more. Furthermore, the value of an "apparent" K_m for one substrate determined at a less-than-saturating concentration of the other is an arbitrary number, which will vary in an *apparently* unsystematic way when, for example comparing mutants and wild type, if account is not taken of the steady-state mechanism of the overall process, i.e., ordered or random addition and leaving of reactants, with steps at thermodynamic equilibrium or not (see, e.g., Severin *et al.*, 1989). Substrate specificities are best expressed quantitatively as the k_{cat}/K_m ratio (Fersht, 1985).

Different Sugar/H^+ Transport Systems in Escherichia coli

Using the first two assay methods described above, six different sugar/H^+ systems were identified in *E. coli*. Their physiological substrates are lactose (LacY, West, 1970), galactose (GalP, Henderson *et al.*, 1977), xylose (XylE, Lam *et al.*, 1980), arabinose (AraE, Daruwalla *et al.*, 1981), fucose (FucP, Bradley *et al.*, 1987), and rhamnose (RhaT, Muiry, 1989). Since the melibiose/Na^+ transporter can transport H^+ (Wilson *et al.*, 1986; Wilson and Wilson, 1987; Leblanc *et al.*, 1989), it may also be included in this class. Some of their properties are summarized in Table I.

Table I. Properties of Sugar-Cation Symporters in Bacteria

Substrate	Organism	True M_r	Apparent M_r	Amino acids	Gene[a] location (min)	Inhibitors		K_m^{app} (μM)[c]
						NEM	Cyto B[b]	
L-Arabinose	*Escherichia coli*	51683	37000	472	61	Yes	Yes	60–320
D-Xylose	*Escherichia coli*	53603	39000	491	91	Yes	No	70–170
D-Galactose	*Escherichia coli*	–[d]	37000	–	64	Yes	Yes	50–450
Lactose	*Escherichia coli*	46502	30000	417	8	Yes	No	50–900
Melibiose	*Escherichia coli*	52202	39000	469	92	Yes	–	300[e]
L-Fucose	*Escherichia coli*	47773	–	439	60	No	No	18–38
L-Rhamnose	*Salmonella typhimurium*	37390	–	344	88[f]	No	No	16–60
D-Glucose	*Synechocystis* PCC6803	49743	–	468	–	–	–	–
Lactose	*Streptococcus thermophilus*	69454	–	634	–	–	–	–

[a] Location of the gene on the chromosome of *E. coli*.
[b] Cytochalasin B.
[c] The K_m for sugar has not necessarily been determined in the presence of saturating cosubstrate.
[d] A dash implies that the information is not yet known.
[e] The value depends upon the identity of the cation cosubstrate.
[f] The gene from *E. coli* has been located, but not yet sequenced.

Table II. Molecular Recognition by Sugar Transporters

Xylose/H^+ symporter (*E. coli*)[a]	Arabinose/H^+ symporter (*E. coli*)[b]	Galactose/H^+ symporter (*E. coli*)[c]	Glucose transporter (erythrocyte)[d]	Glucose transporter (adipocyte)[e]	Fucose/H^+ symporter (*E. coli*)[f,g]	Rhamnose/H^+ symporter (*E. coli*)[g]
D-Xylose	L-Arabinose	D-Glucose	2-Deoxy-D-glucose	6-F-D-Galactose	L-Fucose	L-Rhamnose
	5-CH_3-L-Arabinose (D-Fucose)	2-Deoxy-D-glucose	D-Glucose	D-Glucose	L-Galactose	L-Mannose
		D-Galactose	6-Deoxy-D-glucose	6-Deoxy-D-glucose	D-Arabinose	L-Lyxose
	5-FCH_2-L-Arabinose (6-F-D-Galactose)	6-F-D-Glucose	D-Mannose	2-Deoxy-D-galactose		
	D-Xylose	6-F-D-Galactose	D-Galactose	D-Galactose		
		6-Deoxy-D-glucose	2-Deoxy-D-galactose	6-Deoxy-D-galactose		
		D-Talose	D-Xylose	D-Talose		
		2-Deoxy-D-galactose	L-Arabinose	3-Deoxy-D-glucose		
		D-Mannose	6-Deoxy-D-galactose	D-Xylose		
		6-Deoxy-D-galactose (D-Fucose)		L-Arabinose		
		D-Xylose				

[a] Davis (1986).
[b] Petro (1988).
[c] Henderson *et al.* (1977); Horne (1980); Henderson and Macpherson (1986).
[d] Lefevre (1961); Barnett *et al.* (1973).
[e] Rees and Holman (1981).
[f] Bradley *et al.* (1987).
[g] Muiry (1989).

Further evidence indicating that each system requires a different protein (or proteins, see below) resulted from the mapping of the genes at different locations on the *E. coli* chromosome (Table I). In addition, each one had a different and characteristic specificity for a range of sugars and their structural analogues (Table II), as determined by the assays described above and discussed below.

Sugar/H^+ Transport Systems in Other Bacteria

The occurrence of proton-linked sugar transport systems in different species of bacteria has not been systematically investigated. Consequently, their distribution *appears* to be rather capricious. The following summarizes the current state of knowledge, but it is expected that many more examples will soon be identified. In view of the occurrence of related proteins in organisms as diverse as the Cyanobacteria, yeasts, and mammals (see below), these proteins are of considerable general interest, and the acquisition of more information on their distribution in bacteria and higher organisms is needed in order to expose their evolutionary relationships.

Enterobacteriaceae. The lactose/H^+ transporter occurs in *E. coli* and in species of *Klebsiella* (Buvinger and Riley, 1985; McMorrow *et al.*, 1988), but not in *Salmonella typhimurium*. Its evolution and distribution may be more related to the need to detoxify environmental galactosides rather than utilize them for growth (Lin, 1987). Rhamnose/H^+ symport activity was found in *E. coli*, *S. typhimurium*, *Klebsiella* species and *Erwinia* species, and both fucose/H^+ and galactose/H^+ activity occurred in the first three of these, whereas xylose/H^+ symport was detected only in *E. coli* (McDonald and Henderson, unpublished observations). DNA homologous at high stringency to the *E. coli araE* gene encoding the arabinose/H^+ symporter was readily detected in species of *S. typhimurium*, *Klebsiella*, and *Enterobacter*, but not in *Erwinia* or *Hafnia* (Charalambous *et al.*, 1989). These authors describe a combination of biochemical and genetical strategies by which sugar/H^+ transport proteins can be detected in any organism.

Lactic Acid Bacteria. Dairy lactic acid bacteria utilize lactose as a carbon source. This is taken up by a phosphotransferase system, or in some species, such as *Streptococcus thermophilus*, *Lactobacillus bulgaricus*, *L. helveticus*, *L. brevis* and *L. buchneri*, by a proton-linked mechanism (Hickey *et al.*, 1986). The gene for one of these, from *Strep. thermophilus*, has recently been cloned and sequenced. It resembles the melibiose/Na^+ transporter of *E. coli* with an additional C-terminal region like the enzyme III of phosphotransferase systems (Poolman *et al.*, 1989), and so is remarkably different from both the *E. coli* and the yeast lactose transporters.

Cyanobacteria. While most Cyanobacteria are photosynthetic, a few species can grow chemo- or photoheterotrophically on sugars as carbon

source (Rippka *et al.*, 1979). Proton-linked glucose transporters were identified in several species, e.g., *Synechocystis*, *Nostoc* Mac and *Plectonema* (Raboy and Padan, 1978; Beauclerk and Smith, 1978; Flores and Schmetterer, 1986; Joset *et al.*, 1988). The gene for one of these, from *Synechocystis*, was recently sequenced to reveal its similarity to the *E. coli* xylose/H^+, arabinose/H^+, and galactose/H^+ proteins and the eukaryote sugar transporters described in the next section (Zhang *et al.*, 1989).

Sugar/H^+ Transport Systems in Eukaryotes

Glucose /H^+ symport activity has been reported in species of the alga, *Chlorella* (Komor and Tanner, 1974), and in the protozoan parasite *Leishmania donovani* (Zilberstein and Dwyer, 1985). The gene of the former was recently cloned and sequenced and found to be a member of the widely dispersed sugar transporter family (Table III). So too was a developmentally regulated gene sequenced from *Leishmania enriettii* that has not yet been shown actually to catalyze transport (Cairns *et al.*, 1989a). A sucrose/H^+ transporter found in the cotyledons of *Ricinus* (Hutchings, 1978a, b; Daie, 1989) may also be a member of the same family. However, the glucose/H^+ symport activities reported in plant cell membranes of *Zea mays* (maize) coleoptiles (Rausch *et al.*, 1989), and in a cell line cultured from *Chenopodium rubrum* ("Fat hen" or "Goosefoot," Gogarten and Bentrup, 1989), are more likely to be related to the glucose/Na^+ transporter of mammalian cells than to the homologous family in Table III, since the former cross-reacts with antibodies raised to the glucose/Na^+ protein and the latter is sensitive to phlorhidzin (Gogarten and Bentrup, 1989; Rausch *et al.*, 1989).

A variety of yeast species have transport systems for glucose, galactose, mannose, maltose, and lactose, some of which are proton-coupled (Eddy, 1982; Romano, 1986). Their relationship to the bacterial sugar/H^+ transporters will be described below.

Inhibitors of Proton-Linked Sugar Transport

Covalent Modifiers, Reversible Inhibitors, and Alternative Substrates

These types of compound have proved of great value in the characterization of transport proteins. For example, the combination of a covalent modifier, the sulfhydryl reagent *N*-ethylmaleimide, and an alternative substrate, "thio-di-galactoside" (TDG), which protected the LacY protein much more effectively against NEM than the normal substrate lactose, enabled the first identification of a transport protein (Fox and Kennedy, 1965; Carter *et al.*, 1968; Jones and Kennedy, 1969; Kennedy, 1970). Subsequently, Kacszorowski

et al. (1980) found that 4-nitrophenyl-α-D-galactoside (NPG) reacted covalently with the LacY protein when the two were exposed to UV light, and conditions were found that achieved reasonable specificity. This convenient labelling method was invaluable for assaying the protein during purification (Newman *et al.*, 1981; Viitanen *et al.*, 1986; Kaback, 1986, 1989). Furthermore, NPG is a tight binding reversible ligand that can be used to quantitate the amount of both the LacY and MelB proteins and to test the integrity of their sugar binding sites in mutants (see, e.g, Damiano-Forano *et al.*, 1986; Overath *et al.*, 1987; Kaback, 1987; Leblanc *et al.*, 1989; Roepe *et al.*, 1990).

Alternative substrates serve first to characterize the specificity of a transport protein, which is especially useful when comparing a homologous series of proteins, like those in Table III, or in comparing wild-type and mutant proteins (see later). They can also be useful to isolate transport from further metabolism; for example, methyl-β-D-thiogalactoside (TMG) is a substrate for LacY and MelB, but not for LacZ and MelA, the enzymes for entry of the disaccharides into metabolism (Beckwith and Zipser, 1970; Kennedy, 1970). Some, like TMG, are gratuitous inducers of transport that are often more effective than the natural substrate (Beckwith and Zipser, 1970; Kennedy, 1970). Some, like TDG, are relatively poor substrates that are nevertheless more effective than the natural one for protecting the protein against covalent modification (Fox and Kennedy, 1965; Carter *et al.*, 1968; Jones and Kennedy, 1969; Kennedy, 1970; Lombardi, 1981).

In the rest of this section, I briefly describe compounds that are likely to be of similar value in future investigations of the relatively novel series of transporters listed in Table III.

N-Ethylmaleimide

When incubated with a sulfhydryl reagent, such as 1 mM *N*-ethylmaleimide for 15 min at 25°C, the *E. coli* proteins for arabinose/H^+, xylose/H^+, and galactose/H^+ transport lose 70–90% of their activity (Macpherson *et al.*, 1981, 1983; Henderson and Macpherson, 1986; Davies, 1986). The appropriate sugar substrates protect these proteins against *N*-ethylmaleimide. In this respect they resemble the lactose/H^+, melibiose/Na^+, and proline/Na^+ symporters (Fox and Kennedy, 1965; Carter *et al.*, 1968; Hanada *et al.*, 1985; Damiano-Forano *et al.* 1986), although the amino acid sequences of their proteins are different (see below). However, the fucose/H^+ or rhamnose/H^+ transport proteins are insensitive to *N*-ethylmaleimide under the same conditions (Henderson, McDonald and Bradley, unpublished data).

The glucose transporter of human erythrocytes is also inhibited by sulfhydryl reagents (Bloch, 1974). However, in this case the presence of a transport substrate *accelerates* the rate of inactivation by, for example,

N-ethylmaleimide (Rampal and Jung, 1987). The implication of all these results is that the environment of a sulhydryl group is affected by the binding of a sugar to the transport protein. It is not easy to identify the residue(s) involved (see below). However, since there are no cysteine residues in the N-terminal half of the arabinose/H^+ transport protein, the region interacting with *N*-ethylmaleimide must be in the C-terminal half of this protein and, by implication, in all the homologous proteins.

Cytochalasin B

Cytochalasin B is a fungal product that inhibits mammalian glucose transport proteins (Jung and Rampal, 1977) and binds covalently to them when irradiated with UV light (Carter-Su *et al.*, 1982; Shanahan, 1982; Baly and Horuk, 1988). It inhibited the transport of arabinose via AraE or galactose via GalP (but not xylose via XylE) into vesicles made from appropriately induced strains of *E. coli* (McDonald and Henderson, unpublished data). Binding of radioactive cytochalasin B to the amplified arabinose transport protein (AraE) could be measured directly, revealing a dissociation constant of 0.7–1.1 μM (Petro, 1988). The cytochalasin B was displaced competitively by substrate, with a dissociation constant for arabinose of 20–30 mM (Petro, 1988). Thus, cytochalasin B bound much more tightly than substrate, reminiscent of the situation for a transition state analogue (Fersht, 1985). Similar results were obtained with GalP, though binding constants could not be estimated so precisely because the expression of GalP protein had not been amplified as much as that of AraE (Cairns *et al.*, 1989b; Cairns, unpublished data).

These results reinforce the conclusion that the prokaryote and eukaryote sugar transport proteins are structurally similar. The corollary is that all transport proteins that interact with cytochalasin B are likely to be related. The glucose transport protein of *Leishmania donovani* (Zilberstein *et al.*, 1986; Baly and Horuk, 1988) is therefore one of this family of transport proteins (Table III). Since the fucose/H^+, rhamnose/H^+, and lactose/H^+ transporters are not susceptible to cytochalasin B, it might be concluded that they are not structurally related to this family. However, this suggestion must be treated with caution, because attempts to inhibit or photolabel the XylE transporter with cytochalasin B have failed, despite its homology to the proteins in Table III.

When vesicles from appropriately induced strains of *E. coli* or *S. typhimurium* are irradiated with UV light in the presence of [^{3}H]-cytochalasin B, the label becomes covalently attached to the AraE or GalP proteins (Petro, 1988; Cairns *et al.*, 1989b; Charalambous *et al.*, 1989; Smith, unpublished data). Substrates and inhibitors protect the proteins against labelling with cytochalasin B.

Table III. Homologous Sugar Transport Proteins in Different Organisms

Sugar	Organism	Reference
D-Glucose	*Synechocystis* PCC6803	Zhang *et al.* (1989)
D-Xylose	*Escherichia coli*	Maiden *et al.* (1987)
L-Arabinose	*Escherichia coli*	Maiden *et al.* (1987)
D-Galactose	*Escherichia coli*	Roberts, unpublished data
D-Galactose	*Saccharomyces cerevisiae*	Szkutnicka *et al.* (1989)
D-Glucose	*Saccharomyces cerevisiae*	Celenza *et al.* (1988)
Maltose	*Saccharomyces carlsbergensis* *Saccharomyces cerevisiae*	Yao *et al.* (1989) Szkutnicka *et al.* (1989)
Lactose	*Kluyveromyces lactis*	Chang and Dickson (1988)
D-Glucose	*Chlorella kessleri*	Sauer and Tanner (1989), in press
D-Glucose	*Leishmania donovanii* *Leishmania enriettii*	Zilberstein *et al.* (1986) Cairns *et al.* (1989a)
D-Glucose (isoform E)	Human hepatoma/erythrocyte Rat brain	Mueckler *et al.* (1985) Birnbaum *et al.* (1986)
D-Glucose (isoform L)	Rat liver Human liver	Thorens *et al.* (1988) Fukumoto *et al.* (1988)
D-Glucose (isoform M, insulin-sensitive)	Rat adipocyte/heart/muscle Mouse adipocyte/heart/muscle Rat muscle and fat endothelia	James *et al.* (1989) Birnbaum (1989) Kaestner *et al.* (1989) Vilaro *et al.* (1989)

Cytochalasin B will be an invaluable tool for explaining the differences in specificity between these sugar transport proteins, for labelling and assaying the proteins in purification experiments, and for the determination of their structures and molecular mechanisms.

6-Deoxy-D-glucose

This sugar analogue (= 5-methyl-D-xylose) was not a substrate for XylE in the sugar/H^+ symport assay, although it did inhibit the transport of xylose, the natural substrate (Davis, 1986; McKeown, unpublished data). 6-Deoxy-D-glucose was also more effective at protecting XylE against reaction with *N*-ethylmaleimide than xylose (Davis and Henderson, unpublished data). Although it was not a substrate for AraE in the sugar/H^+ symport assay, 6-deoxy-D-glucose did protect AraE against cytochalasin B. In contrast, 6-deoxy-D-glucose was a good substrate for GalP (Table II).

These preliminary observations imply that 6-deoxy-D-glucose either binds to the normal (external?) sugar binding site of XylE and AraE, but fails to trigger the evens leading to sugar/H^+ symport, or that it binds to a second

(internal?) site that may also interact with cytochalasin B (Walmsley, 1988). More investigations are required to confirm and expand these initial observations. However, the availability of dead-end inhibitors, of which 6-deoxy-D-glucose and cytochalasin B appear to be two, will be important for exploiting kinetic measurements of the bacterial transporter activities to the level that has already been valuable in mechanistic studies of the mammalian glucose transporter (Deves and Krupka, 1978; Lowe and Walmsley, 1986; Walmsley, 1988).

Other Alternative Substrates

These are listed in Table II in the order of their affinities for each transport protein as reported in the literature. The following observations can be made.

Of the homologous transporters, XylE is most selective for its substrate, followed by AraE and then GalP and the mammalian sugar transporters. One may speculate that, if evolution started with the glucose transporter, then the changes in primary sequence may first have effected discrimination against the 6 OH position (not the 6 $HOCH_2$) to achieve a pentose transporter (AraE) and then enhanced recognition of the precise 4 OH configuration to achieve a xylose transporter (XylE). Analysis of the recognition of each hydroxyl group in the hexose molecule has led to models of the sugar recognition site in the glucose transporters of mammals (Barnett *et al.*, 1973; Rees and Holman, 1981; Walmsley, 1988). These may be equally relevant to the binding of sugars by the bacterial protein, GalP.

Future mutagenesis experiments may succeed in associating changes in just a few amino acid residues in each protein with the changes in sugar recognition. This is an important goal in the elucidation of the relationship of structure to function of these transport proteins.

Individual sugar analogues have proved to be of particular value. 2-Deoxy-D-galactose, D-talose, and D-fucose are all substrates for GalP but not for the first metabolic enzyme, galactokinase, to any significant extent. Thus, their radioisotope-labelled forms are useful for isolating transport from metabolism in transport assays, and the unlabelled sugars give maximal alkaline pH changes in the direct sugar/H^+ symport assays without the complication of acid production due to metabolism (Henderson and Macpherson, 1986). Furthermore, D-fucose is a gratuitous inducer of all the *gal* operons, useful for maintaining high levels of expression in cultures of *E. coli*, and probably other Enterobacteriaceae. D-Fucose is also a valuable nonmetabolizable substrate of the AraE transporter (Henderson and Macpherson, 1986).

L-Rhamnose is 6-deoxy-L-mannose and L-fucose is 6-deoxy-L-galactose, and they can be regarded as either 6-deoxy-hexoses or 5-methyl-pentoses.

The L-rhamnose/H^+ or L-fucose/H^+ transporters of *E. coli* accept the corresponding hexoses or pentoses (Table II), albeit with reduced affinity, so loss of the normal substrate's methyl group or substitution of an OH group in it does not prevent binding (Bradley *et al.*, 1987; Muiry, 1989). However, each transport protein recognizes the correct orientation of the OH at the C-2, C-3, or C-4 position in the pyranose ring, and so a sugar differing in only one of these configurations is not a substrate (Muiry, 1989). Thus, the L-rhamnose transporter accepts L-rhamnose, L-mannose, and L-lyxose, but not L-fucose, L-galactose, and D-arabinose, which are substrates of the L-fucose transporter (Muiry, 1989). Each of the substrate analogues is metabolized relatively slowly, and has consequently proved very useful in optimization of sugar/H^+ symport assays. None of them appear to be gratuitous inducers of the *E. coli* operons. Since the predicted primary sequences of the RhaT and FucP transport proteins are not homologous to each other or to any of the other sugar/H^+ symporters (see below), comparative studies are unlikely to yield insights into the nature of the substrate recognition process by the two L-deoxy-sugar transporters.

The Locations of Genes Encoding Sugar/H^+ Transport Proteins on the Chromosome of *Escherichia coli*

An important step in cloning and sequencing a transport protein of *E. coli* is to establish the location of the corresponding gene on the chromosome. This can be achieved by first isolating a mutant impaired in the relevant transport activity. Random mutagenesis of *E. coli* is conveniently accomplished by insertion of specially constructed transposons or phages into the genome (Jones-Mortimer and Henderson, 1986). These usually contain antibiotic resistance or other markers to facilitate selection of cells into which they have become incorporated. It is then necessary to devise selection and screening procedures, which show that a particular transport system has been mutated. Amongst these are the appearance of sugar-resistance phenotypes in *fda*, *fdp*, or *ppc* strains, the acquisition of resistance to toxic sugar analogues, and the disappearance of sugar/H^+ symport activity (Henderson and Kornberg, 1975; Jones-Mortimer and Henderson, 1986).

One mutant was isolated that contained Mu d(Ap^r) phage inserted into the *xylE* gene, which encodes the xylose/H^+ symport protein; this facilitated the location of the gene, because it was much easier to score the ampicillin-resistance phenotype on solid media than the presence or absence of the transport activity (Davis *et al.*, 1984). Classical genetical techniques of interrupted conjugation and phage P1-mediated transduction were then used to locate the ampicillin-resistance determinant, and consequently the *xylE* gene,

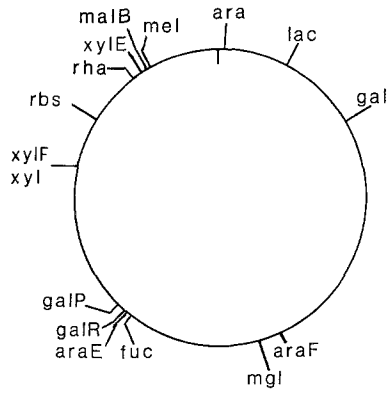

Fig. 2. Locations of some genes encoding enzymes of sugar transport and metabolism on the gene linkage map of *Escherichia coli*. The circle represents the chromosome of *E. coli*, which is calibrated arbitrarily 0–100 min from the index mark on the inside at the top. Details are given by Bachmann (1987), and the nomenclature is as follows; *ara* contains genes for regulation of transcription and for the enzymes of metabolism of arabinose; *lac* contains genes for regulation of transcription, lactose/H^+ transport, and metabolism of lactose; *gal* contains genes for metabolism of galactose; *araF* contains three genes encoding proteins for ATP-energized binding protein-mediated transport of arabinose; *mgl* contains genes (probably three) for ATP-energized binding protein-mediated transport of galactose; *fuc* contains genes for regulation of transcription, fucose/H^+ transport, and metabolism of fucose; *araE* encodes the arabinose/H^+ symport protein; *galR* produces a repressor protein for regulation of transcription of *gal*, *mgl*, and *galP*; *galP* encodes the galactose/H^+ symport protein; *xyl* contains the genes for internal metabolism of xylose and for regulation of transcription; *xylF* contains genes (probably three) for ATP-energized binding protein-mediated transport of xylose; *rbs* contains genes for regulation of transcription, ATP-energized binding protein-mediated transport (three genes), and metabolism of ribose; *rha* contains genes for regulation of transcription, rhamnose/H^+ transport (probably one gene), and metabolism of rhamnose; *xylE* encodes the xylose/H^+ symport protein; *malB* contains five genes involved in ATP-energized binding protein-mediated transport of maltose; and *mel* contains a gene for melibiose/Na^+ transport and one for melibiose metabolism. References are in the text.

at 91.4 min on the *E. coli* chromosome. Another useful mutagenic phage is lambda p *lac* Mu (Bremer *et al.*, 1984; Silhavy *et al.*, 1984), which was exploited to clone the *xylE*, *araE*, and *galP* genes located at 91.4, 61.2, and 63.7 min, respectively (Davis and Henderson, 1987; Maiden *et al.*, 1988; Roberts, Moore, and Henderson, unpublished data).

The positions of a number of relevant sugar transport genes on the *E. coli* chromosome are summarized in Fig. 2. The transport genes can be located more precisely by comparison of the positions of restriction sites in their genes and in flanking DNA with a restriction map of the entire *E. coli* genome (Kohara *et al.*, 1987).

Some transport genes occur in operons where they are adjacent to the genes encoding the enzymes for metabolism of the sugar and to the genetic elements regulating gene expression. The best known case of this type is the

contiguous *lac* operon (Beckwith and Zipser, 1970). Similarly, the gene encoding the fucose/H^+ transporter is part of a contiguous fucose operon at 60 min (Fig. 2; Lu and Lin, 1989), and the gene encoding the rhamnose/H^+ transporter is probably part of a contiguous rhamnose operon at 88 min (Fig. 2; Badia *et al.*, 1989; Muiry and Tate, unpublished data). Both of these operons also contain the genes for the metabolic enzymes and regulatory proteins (Tobin and Schleif, 1987; Lu and Lin, 1989; Badia *et al.*, 1989; Muiry and Tate, unpublished data).

In other cases the transport gene is entirely separate. For example, the *araE* gene encoding the arabinose/H^+ transporter is at 61 min, distant from the genes for the enzymes of arabinose metabolism at 0 min, from the associated regulatory region, also at 0 min, and from the genes for the binding protein transport system at 45 min (Fig. 2; Bachmann, 1987). Similarly, the *xylE* gene for the xylose/H^+ transporter is at 91 min (Davis *et al.*, 1984), separate from the other genes encoding the enzymes of xylose metabolism, the regulatory region, and the binding protein xylose transport system. These are clustered together at about 80 min (Bachmann, 1987; Sumiya, 1989). The *galP* gene for galactose/H^+ transport is located at about 64 min (Riordan and Kornberg, 1977; Roberts, Moore, and Henderson, unpublished data), and is again separate from the genes encoding the metabolic enzymes at 17 min, the *galR* regulatory region at 61 min, and the genes for the galactose binding protein transport system at 45 min (Fig. 2; Bachmann, 1987).

However discontinuous the operon, the expression of all genes involved in the transport and metabolism of a particular sugar is highly coordinated. The presence of the sugar in the environment derepresses expression of all the genes by factors of 10 to several hundred (Buttin, 1968), though this may be attenuated by catabolite repression (Saier, 1985; Postma and Lengeler, 1985; Lengeler *et al.*, 1990). To date, only the expression of the *ara* operons has been studied in any detail (Kolodrubetz and Schleif, 1981; Kosiba and Schleif, 1982; Stoner and Schleif, 1983), and it is still unclear precisely how the transcription of the *araE* gene for transport is coordinated with the expression of the other *ara* operons. The physiological and/or evolutionary significance of this complex organization of the genetic elements involved in sugar transport, metabolism, and the regulation of gene expression is, as yet, unclear.

The gene for each of the transport proteins in Table I has been cloned and sequenced. In the case of the *E. coli* lactose/H^+ transporter, it has been shown unequivocally that only a single gene product is necessary for activity (Kaback, 1989). For the other sugar/H^+ symporters, no evidence for the involvement of more than one protein has been found, though the possibility has not been eliminated.

The Primary Sequences of Sugar/H^+ Transport Proteins

The Arabinose/H^+ and Xylose/H^+ Transport Proteins of Escherichia coli

The sequences of the amino acids in the arabinose/H^+ and xylose/H^+ transport proteins of *E. coli* were established from the DNA sequences of their genes (Davis and Henderson, 1987; Maiden *et al.*, 1987, 1988). The AraE and XylE proteins are homologous, with 141 identical residues out of 472 and 491, respectively, in the aligned sequences (Fig. 3; Maiden *et al.*, 1987). There are additional conservative substitutions throughout the proteins, so that nearly 40% of the residues can be regarded as homologous (Maiden *et al.*, 1987; Baldwin and Henderson, 1989). In contrast, there is little homology in the DNA sequences of the genes (Maiden, 1987).

The Galactose(glucose)/H^+ Transport Protein of Escherichia coli

Only 270 residues of the *E. coli* galactose/H^+ transporter (GalP) have been sequenced so far, but 176 are identical to those in the aligned arabinose/H^+ (AraE) transporter (Roberts, Moore, and Henderson, unpublished data). Since GalP and AraE are therefore more closely related to each other than they are to XylE, which is reflected also in the similarity of the DNA sequences of the *galP* and *araE* genes, they perhaps resulted from a more recent gene duplication and subsequent divergence (Doolittle, 1981; von Heijne, 1987).

The substrate specificity of GalP is more similar to that of the mammalian glucose transporter than to those of AraE or XylE (Table II). Therefore, if the high degree of identity persists throughout their sequences, AraE and GalP will be two proteins of relatively similar amino acid sequence with different substrate specificities, and GalP and the glucose transporter will be of relatively different sequence and similar specificities. It seems likely that, when the sequence of GalP is completed, comparisons will identify regions of all the proteins, and even individual amino acids, that must contribute to molecular recognition of the substrate.

The Glucose/H^+ Transport Protein of Synechocystis

Zhang *et al.* (1989) selected a clone containing the glucose transport protein of *Synechocystis* PCC6803 by complementation of a fructose-resistant mutant (cf. Flores and Schmetterer, 1986). Of the two reading frames in the sequenced DNA in this clone, one encoded a protein predicted to contain 468 amino acids. This could be readily aligned with the other proteins listed in Table III, showing 25–39% identity and 46–60% homology (Zhang *et al.*, 1989). It was most similar to the *E. coli* xylose/H^+ symporter (XylE).

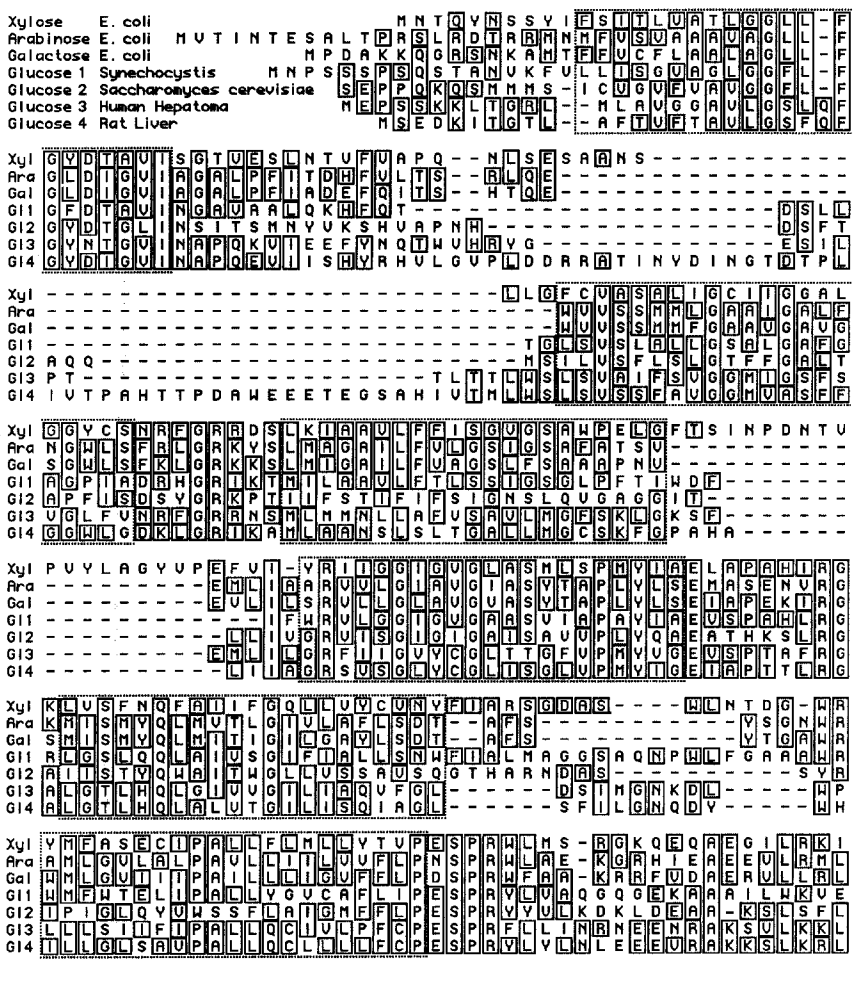

Fig. 3. The aligned amino acid sequences of homologous sugar transport proteins from bacteria and eukaryotes. The single letter code for amino acids is employed for the following transport proteins: Xyl, xylose/H^+ symporter of *E. coli*; Ara, arabinose/H^+ symporter of *E. coli*; Gal, galactose/H^+ symporter of *E. coli*; G11, glucose/H^+ symporter of *Synechocystis* PCC6803 (a cyanobacterium); G12, glucose transporter of *Saccharomyces cerevisiae* (a yeast); G13, glucose transporter of a human hepatoma cell line; G14, glucose transporter of rat liver. The alignments have been modified slightly from the individual published ones (references in Table I and the text). The dotted rectangles enclose hydrophobic regions.

The Primary Sequences of Related Sugar Transport Proteins in Eukaryotes

Yeast Transport Proteins for Glucose, Galactose, Lactose, and Maltose

The existence of proton-linked and facilitated diffusion transport systems for sugars in various species of yeast is well established (Eddy, 1982; Kotyk,

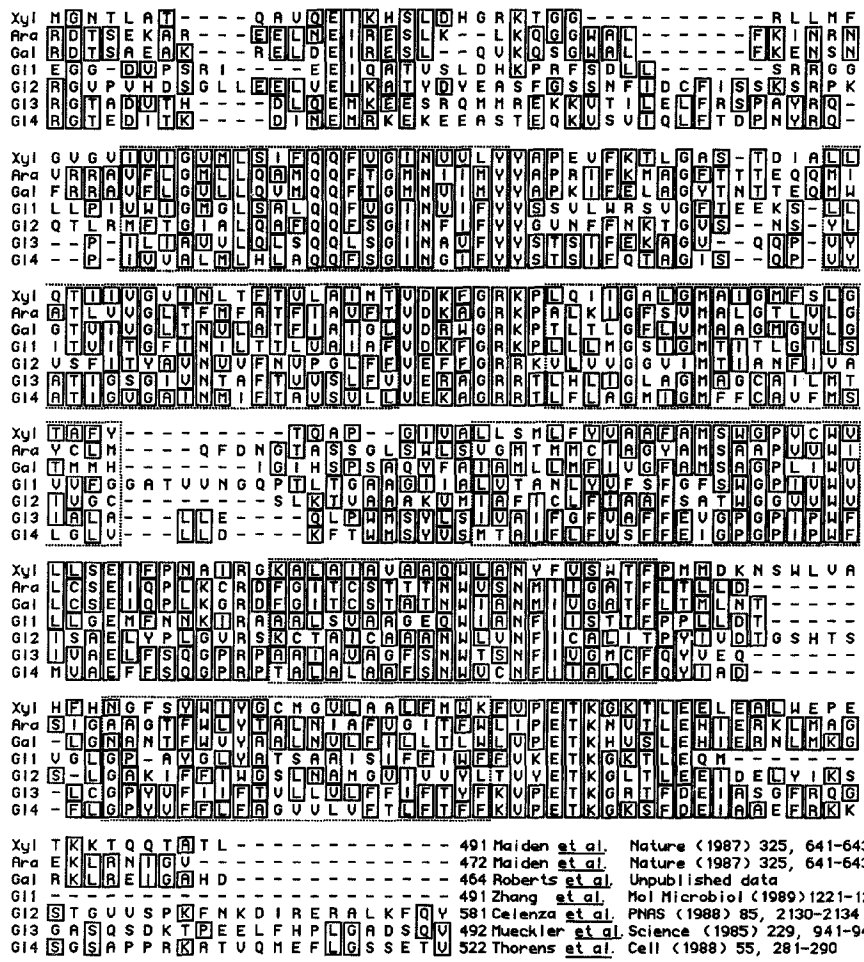

Fig. 3. Continued.

1983; van Dijken and Scheffers, 1986). The primary sequence of a glucose facilitated diffusion transport protein from *Saccharomyces cerevisiae* was recently deduced from the DNA sequence of the *SNF3* gene (Celenza *et al.*, 1988). Out of 884 residues in the amino acid sequence, those from 86 to 581 were 28–31% identical with the bacterial transporters AraE and XylE and with the mammalian D-glucose transporters discussed below; the C-terminal 87 residues were not necessary for transport function (Celenza *et al.*, 1988).

The primary sequence of a lactose transport protein, probably proton-linked, from *Kluyveromyces lactis* was predicted from the DNA sequence of the *LAC12* gene (Chang and Dickson, 1988). Out of 587 residues, those from 117 to 530 revealed homology with residues 60 to 480 of the sequence of

AraE, XylE, and the mammalian glucose transporters, even though the number of identities between the monosaccharide transporters was greater than between any one of them and the yeast disaccharide transporter. The yeast lactose transporter was not similar to the LacY protein of *E. coli* (Chang and Dickson, 1988). Thus, two proteins of very different primary sequence have evolved to transport the same sugar. It remains to be determined whether the two proteins have similar three-dimensional structures.

The primary sequence of a putative galactose transport protein from *Saccharomyces cerevisiae* was recently predicted from the DNA sequence of the *GAL2* gene (Szkutnicka *et al.*, 1989). Out of 574 amino acids, those from about 70 to about 530 revealed clear homologies with the family in Table III. In the same publication a homologous maltose transport protein is described, quoting unpublished work of Michels and Cheng. The sequence of a maltose transport protein deduced from the DNA sequence of the *MAL6T* locus of *Saccharomyces carlsbergensis* is extremely similar to this, although the author's assignment of hydrophobic regions is different (Yao *et al.*, 1989).

The presence of much longer N-terminal and C-terminal segments in the yeast proteins is perhaps related to the need to target proteins to different membranous organelles in these organisms (Szkutnicka *et al.*, 1989).

Mammalian Glucose Transport Proteins

Mueckler *et al.* (1985) cloned and sequenced *c*DNA encoding a facilitated diffusion glucose transport protein from the human hepatoma HepG2 cell line. Of its 492 residues, 28–31% were identical to those in the aligned AraE and XylE pentose/H^+ transporters from *E. coli* (Maiden *et al.*, 1987; Fig. 3), and, taking into account conservative substitutions, each bacterial protein was as similar to the human protein as they were to each other. Immunological and peptide sequence data indicated that the hepatoma protein was very similar, if not identical, to the intensively studied glucose transporter of human erythrocytes (Wheeler and Hinkle, 1985; Baly and Horuk, 1988; Baldwin and Henderson, 1989). The glucose transporter of rat brain was virtually identical to the human hepatoma/erythrocyte ones, with only seven different residues out of 492 (Birnbaum *et al.*, 1986). There is probably another related glucose transporter in human brain (Kayano *et al.*, 1988).

A rat liver glucose transport protein, cloned, sequenced, and expressed in *E. coli*, was substantially homologous to the hepatoma/brain/erythrocyte transporter (the E isoform), with about 55% identical residues in the aligned sequences (Fukumoto *et al.*, 1989; Thorens *et al.*, 1988). These structural differences must account for the different kinetic properties of the liver protein (Ciaraldi *et al.*, 1986; Baly and Horuk, 1988), which are presumably allied to the special physiological role of glucose transport in liver tissue

(Nordlie, 1985; Baly and Horuk, 1988). Messenger RNA hybridization experiments showed that the same protein (the L isoform) also occurred in intestine, kidney, and β-pancreatic islet cells (Thorens *et al.*, 1988).

James *et al.* (1989) and Birnbaum (1989) sequenced *c*DNA of the insulin-regulatable glucose transporter from rat adipose tissue. Its predicted amino acid sequence was 65% identical to that of the HepG2/brain transporter. The differences between the three types of glucose transporters (Table III) appeared to be most significant in the third predicted membrane-spanning segment (James *et al.*, 1989, see below), presumably reflecting different properties necessary for the special physiological role of glucose transport in each tissue (Baly and Horuk, 1988). DNA/RNA hybridization and immunological experiments indicated that the insulin-regulatable glucose transporter (the M isoform) was present in brown and white adipose tissue, heart and red and white skeletal muscle, but not in brain, liver, or HepG2 cells (James *et al.*, 1989; Vilaro *et al.*, 1989). Significantly, the highest level of its expression occurred in the epithelial cells of insulin-responsive tissues (Vilaro *et al.*, 1989).

It seems, therefore, that mammals have at least three different facilitated diffusion glucose transport proteins expressed to different levels in each tissue, each with a particular role in the body's mechanisms for capturing, distributing, or utilizing glucose (Nordlie, 1985; Kayano *et al.*, 1988; Baly and Horuk, 1988; Vilaro *et al.*, 1989). These proteins are homologous to each other (55–65% identity) and to the sugar transport proteins found in single-celled bacteria, yeasts, algae, and protozoa (25–33% identity, Table III). It seems likely that they also occur in the great diversity of intervening multicellular life forms.

Features of the Aligned Primary Sequences of the Homologous Sugar Transport Proteins

There Are Probably Twelve Membrane-Spanning α-Helices

When the hydropathic profile of each of the homologous sugar transport proteins is analyzed independently, there appear to be twelve hydrophobic regions, predicted to be membrane-spanning α-helices (references in Table III). In some cases uncertainty over the predicted position of these regions is overcome when the sequences are aligned according to the amino acid identities. One example of this is loop 11 of the AraE protein (Baldwin and Henderson, 1989; Fig. 4). Some faint evidence of duplication in the two halves of each protein (see below) suggests that the membrane-spanning regions should be regarded as two groups of six (Maiden *et al.*, 1987).

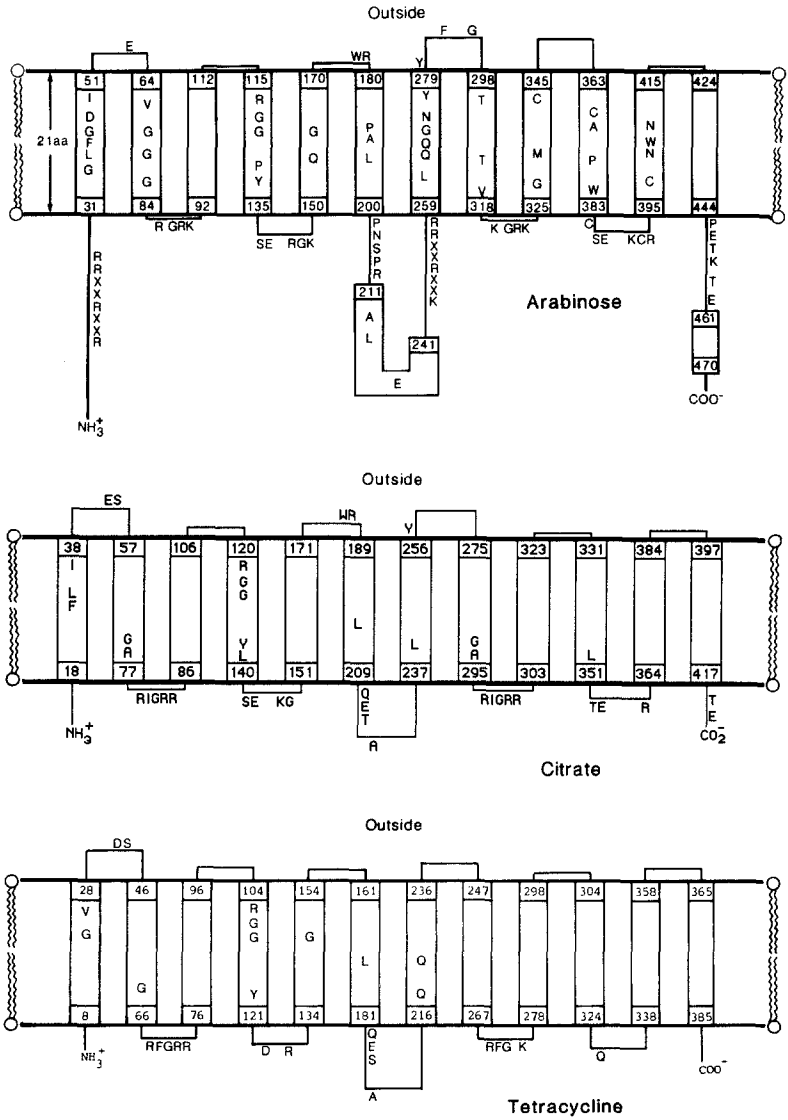

Fig. 4. Proposed orientations of the arabinose/H^+ symporter, the citrate/H^+ symporter, and the tetracyline/H^+ antiporter in the cytoplasmic membrane of bacterial cells. The models are based on predictions of membrane-spanning regions and α-helix content and on the alignment with the glucose transporter (Eisenberg *et al.*, 1984; Mueckler *et al.*, 1985; von Heijne, 1987). The 12 putative membrane-spanning segments, usually of 21 residues, are depicted as rectangles with their end residues numbered from the N-terminus. The rectangles outside the membrane represent predicted α-helices. The single residues indicated are usually conserved in all the aligned sugar transporters (Fig. 3), and some recur at equivalent positions in the two halves of the molecule as discussed in the text.

The predicted occurrence of twelve membrane-spanning α-helices is not restricted to the homologous sugar transport proteins listed in Table III. It is also the case for the lactose/H^+ (Foster *et al.* 1983), glucose-6-phosphate/P_i (Friedrich and Kadner, 1987), glycerol-3-phosphate/P_i (Eiglmeier *et al.*, 1987), proline/Na^+ (Nakao *et al.*, 1987), and possibly the melibiose/Na^+ (Leblanc *et al.*, 1989) transporters of *E. coli*, the plasmid- or transposon-encoded citrate and tetracycline transporters (see below) of bacteria, and the uracil transporter of yeast (Jund *et al.*, 1988). Perhaps this is a general property of many other transport proteins (cf. Kartner and Ling, 1989; Riordan *et al.*, 1989; Higgins, 1989). However, there are exceptions, for example the rhamnose/H^+ transporter of *S. typhimurium* (Muiry, 1989, see below).

There Is a Central Hydrophilic Region

In all the homologous sugar transport proteins there is a central hydrophilic region of 55–70 residues strongly predicted to comprise a high proportion of α-helix. It contains several conserved regions (Fig. 3), and so presumably plays an essential role in the maintenance of structure and function. In the erythrocyte glucose transporter it may be identified with a protease-sensitive region that is cleaved by trypsin only when it is applied to the cytoplasmic face of the membrane (Baldwin *et al.*, 1980; Fig. 2 in Baldwin and Henderson, 1989). This may imply that the region is cytoplasmic in all the transporters.

Significance of the Conserved Amino Acid Residues

When six different sugar transporters for arabinose (*E. coli*) or xylose (*E. coli*) or glucose (*S. cerevisiae*, human hepatoma, human liver, rat adipocyte) are aligned, there are 45 absolutely conserved residues and about 150 conservative substitutions. These residues may be important features of the common structure and mechanism of these proteins. An analysis (Maiden, 1987) of the four transporters for arabinose, xylose, glucose, and citrate showed that the most commonly conserved residues were glycine, proline, arginine, and glutamate. The significance of the first two probably reflects the unique spatial properties of their side chains (von Heijne, 1987). The second two are probably involved in structurally important charge interactions. The conserved glutamate residues are also of interest because of their potential interaction with protons in the sugar/H^+ symporters. However, so far there is no obvious relationship between the presence or absence of a particular glutamate (or aspartate) and the capacity for cation symport. This highlights the need to be certain whether a particular protein transports cations or not (see below).

The bacterial transporters, unlike the mammalian transporters, are energized by the cotransport of a proton with the sugar; hence, the proton-translocating residues might be conserved only in the bacterial proteins. Likely candidates for such residues are histidine, glutamate, and aspartate. However, despite the proposed involvement of histidine in proton translocation by the functionally similar lactose/H^+ transporter of *E. coli* (Kaback, 1989, 1990; but see also King and Wilson, 1989), no histidine residues are conserved in the bacterial monosaccharide transport proteins (Fig. 3; Baldwin and Henderson, 1989). Furthermore, directed mutagenesis of all seven histidine residues in the melibiose/Na^+ transport protein failed to associate any of them with proton translocation (Leblanc, personal communication), even though alterations in His 94, but not any of the other histidine residues, did profoundly affect transport competence.

Aspartate, asparagine, glutamate, and glutamine residues are involved in the H-bonded interaction of soluble proteins with sugar molecules in their binding sites (Quiocho, 1986, 1990). It is therefore of interest that these residues are highly conserved in some locations predicted to be within membrane-spanning helices (loops, 1, 6, 7, 8, 10; see Figs. 3, 4). These are therefore candidates for site directed mutagenesis experiments to evaluate their possible roles in sugar binding.

In view of their inhibition by sulfhydryl reagents (above), it is interesting that no cysteine residues are conserved in the aligned transporters (Maiden *et al.*, 1987). Such inhibition may stem from steric hindrance of substrate binding or translocation, rather than from an essential mechanistic role of SH groups. It happens that there are no cysteine residues in the N-terminal half of the AraE protein, so the substrate-protected residue(s) must be in the C-terminal part of AraE and, by extension, of the other proteins. Walmsley (1988) has suggested that a reagent-sensitive cysteine residue in the glucose transporter may be at position 347, or possibly 417, in the C-terminal half of the protein. Directed mutagenesis could confirm the location of such residues, as it has in the case of the lactose/H^+ transporter—both peptide labelling/mapping and site-directed mutagenesis experiments identified Cys148 as the residue susceptible to *N*-ethylmaleimide in Lac Y (Beyreuther *et al.*, 1981; Menick *et al.*, 1987). However, this is predicted to occur in the N-terminal half of the protein (Kaback, 1986, 1989).

The RXGRR Motif

Between the putative membrane-spanning helices 2 and 3 and 8 and 9 is a sequence motif RXGRR (Fig. 5). The first and last R may be replaced by K, X is usually an amino acid with a large hydrophobic side chain, and the motif is often preceded by D or N. It is predicted to form a β-turn, linking

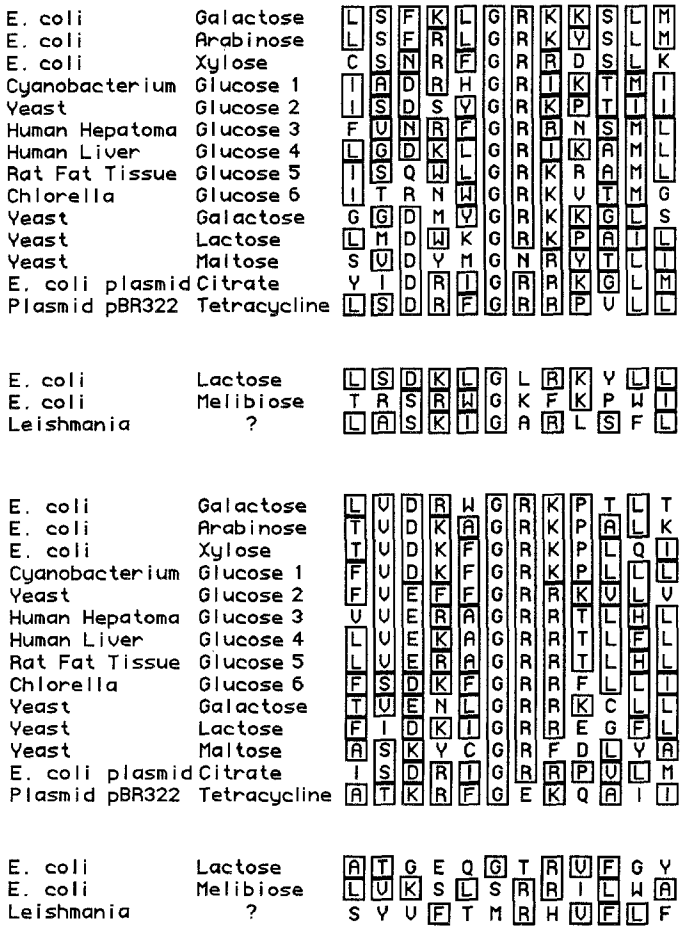

Fig. 5. Conserved sequence motifs between putative helices 2/3 and 7/8 (see, e.g., Fig. 4) of transport proteins.

the adjacent helices. The positively charged side chains may interact with the head groups of lipids. The finding of this motif and the occurrence of twelve hydrophobic regions in a novel sequence are initial clues that a protein belongs to this homologous group (see below). A quite similar motif is found in the corresponding regions of the Lac Y and MelB proteins (Maiden *et al.*, 1987), despite their otherwise differences in primary sequence.

The PESPR and PETK Motifs

These occur after the ends of the putative membrane-spanning helices numbers 6 and 12, respectively, in most of the monosaccharide transporters

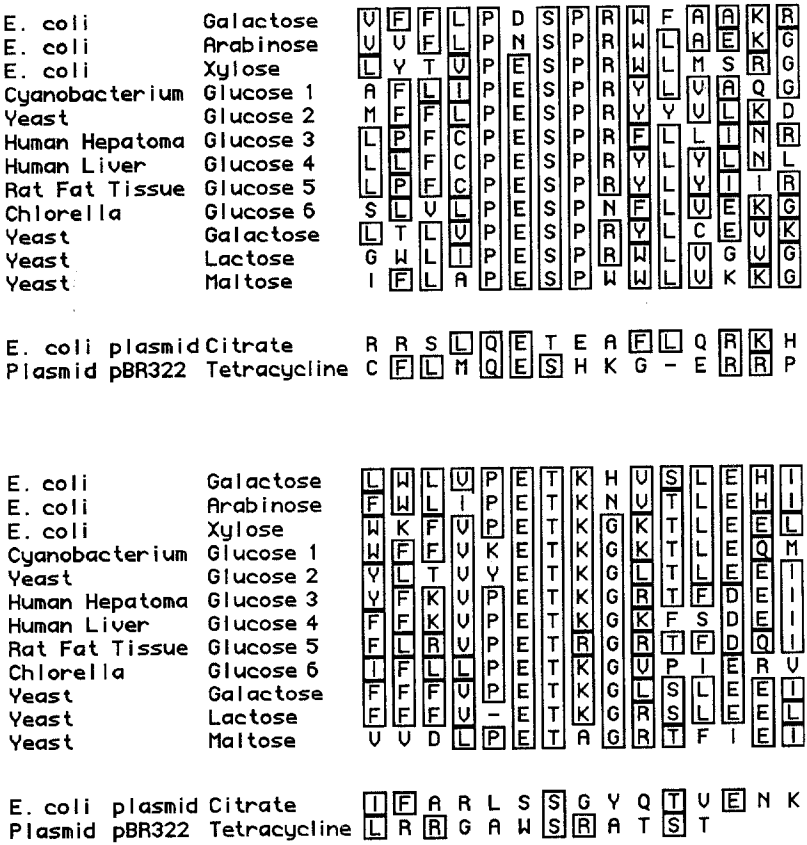

Fig. 6. Conserved sequence motifs at the ends of putative helices 6 and 12 (see, e.g., Fig. 4) of transport proteins.

(Fig. 6). They are also found in the lactose transporter of *K. lactis* (Chang and Dickson, 1988), but not in the citrate or tetracycline transporters (see below). In the AraE and GalP proteins the E in the first motif is changed to N or D, and in some of the glucose transporters the P in the second motif is not conserved. Again, site-directed mutagenesis experiments need to be undertaken to evaluate the possible significance of these differences in relation to substrate specificity, cation recognition, and other parameters of the molecular mechanism.

A Diffused Motif Through Helices 4 and 5

The following pattern is conserved in the region including putative helices 4 and 5 of the aligned monosaccharide transporters (Figs. 3, 4):

R---G---G------P-Y--E------RG------Q-----G

No gaps are inserted to achieve this matching. This motif presumably has structural and/or functional significance, and it is conserved, albeit imperfectly, in the transport proteins for lactose (Chang and Dickson, 1988), citrate, and tetracycline (Henderson and Maiden, 1990; Figs. 3, 4; see below).

Residues That May be Implicated in Sugar Binding from Comparisons with the Lactose Transporter

Many of the homologous sugar transport proteins have a conserved motif P A L L in the predicted membrane-spanning region six (Fig. 3). Since mutation of the residue, alanine 177, in the equivalent region of the lactose/H^+ transporter has been associated with changes in sugar specificity (Brooker and Wilson, 1985; see below) it may be speculated that the P A L L motif is also associated with sugar binding. By a similar, and admittedly tentative, argument the following conserved residues in the homologous proteins might be involved in sugar binding: a tyrosine residue at the end of helix 7; a threonine residue in helix 8; and an isoleucine residue in helix 9. These are all altered in mutants of the Lac Y protein with altered binding of sugars (Markgraf *et al.*, 1985; Brooker and Wilson, 1985; Brooker, 1990).

A Diffused Motif Through Helix 7

Most of the aligned sugar transport proteins of Table I contain the following motif in putative helix 7:

L - - - Q Q - - G I N - - F Y Y

Elements of this can also be recognized in the corresponding regions of the transport proteins for lactose, melibiose, citrate, and tetracycline (see below). The presence of the glutamine and asparagine residues generates the suspicion that the region may contribute to recognition of the sugar, but it may equally be involved in the translocation mechanism. Perhaps the presence of the I N pair here is a repetition of its presence at the end of putative helix 1 in both the aligned glucose transporters (Fig. 3) and the Lac Y protein (Kaback, 1989).

A Diffused Motif Through Helices 10 and 11

In this region there is a pattern of conserved residues as follows

F - - - - G - - - W - - - - E - - - - - - R - - - - - - - - - - N W - - N F

The position and spacing of the glutamate and arginine reflect their occurrence in the corresponding region between helices 4 and 5 in the N-terminal half of the protein. Photolabelling of a sugar transporter with cytochalasin B probably results from photoactivation of an aromatic amino acid residue in the

protein rather than from activation of the inhibitor (Deziel *et al.*, 1984; see above). Peptide mapping experiments indicated that the covalently bound cytochalasin B was located between Phe_{389} and Trp_{412} in the mammalian D-glucose transporter (Cairns *et al.*, 1984; Karim *et al.*, 1987; Holman and Rees, 1987). The only aromatic residue conserved in this region of the bacterial transporters is the second tryptophan in helix 11 of the above motif. Experiments are in progress to determine whether this is the actual site of labelling, and to discover why the XylE protein does not bind cytochalasin B. It may be relevant that this region of the MelB protein is suggested to be part of a sugar recognition site (Botfield and Wilson, 1988; Botfield *et al.*, 1990).

Similarities Between the Sugar, Citrate, and Tetracycline Transport Proteins

The Citrate/H^+ Transporter

A plasmid-encoded citrate transport activity is thought to be proton-linked (Reynolds and Silver, 1983). The transport protein was sequenced, via its gene, by two groups independently (Sasatsu *et al.*, 1985; Ishiguro and Sato, 1985). At the primary sequence level it appeared to have only a slight similarity to the sugar transporters (Maiden *et al.*, 1987), but, when analyzed by the algorithm of Eisenberg *et al.* (1984), the citrate transporter, like the sugar transporters, was predicted to have two groups of six membrane-spanning α-helices separated by a central hydrophilic domain (Maiden, 1987). Furthermore, between the predicted helices 2/3 and 8/9 were the DRXGRR motifs, and through the predicted helices 4 and 5 was the diffused motif described above (Fig. 4). There were other similarities between the citrate transporter and individual sugar transport proteins (Fig. 4), but only a few more identities that occurred in every protein (Fig. 4). The most noticeable differences were the relative shortness of the central (24 residues) and the C-terminal (14 residues) hydrophilic regions in the citrate transporter. These largely account for its shorter length (431 residues) compared to the sugar transporters (472–522 residues).

The Tetracycline/H^+(K^+) Transporter

One class of tetracycline genes encodes proteins that effect the active efflux of the antibiotic from the cell (McMurry *et al.*, 1980; Levy, 1984, 1988). The mechanism may be an obligatory coupling to the movement of protons (presumably antiport) or potassium ions (Dosch *et al.*, 1984). The sequences of several such tetracycline transporters have been determined (Peden, 1983; Waters *et al.*, 1983; Hillen and Schollmeier, 1983). They are extensively

homologous to each other (Hillen and Schollmeier, 1983), but not obviously so to the sugar transporters. Nevertheless, each one is predicted, by the same algorithm used to analyze the sugar transporters, to have twelve membrane-spanning α-helices (Maiden, 1987; Eckert and Beck, 1989; Fig. 4). Between the putative helices 2/3 there is a perfect DRXGRR motif, and between the putative helices 8/9 is RXGEK, similar to RXGRR (Fig. 4). Through the putative helices 4/5 is most of the diffused motif (Fig. 4). There are other similarities in primary sequence between the tetracycline, citrate, and individual sugar transporters (Fig. 4), but few of them are conserved in all of the transporters. Superficially, the structures of tetracycline and cytochalasin B appear similar (Henderson and Maiden, 1990), or at least they resemble each other more than they resemble a monosaccharide (but see Griffin *et al.*, 1982). Perhaps the cytochalasin binding site on a sugar transport protein resembles a tetracycline binding site?

These similarities permit the speculation that the tetracycline transporters, which are probably representatives of a wider class of antibiotic resistance factors, have three-dimensional structures in the membrane that are fundamentally similar to those of the series of homologous sugar transporters.

Structural Models of the Transport Proteins

Many membrane proteins are believed to be folded so that hydrophobic α-helices inserted through the membrane are joined by hydrophilic regions located in the head groups of the lipid and/or the aqueous environment (Kyte and Doolittle, 1982; von Heijne, 1988; Lodish, 1988). There are very few proteins where this model is supported by detailed structural information determined by electron diffraction or X-ray crystallographic measurements, the exceptions being bacteriorhodopsin from *Halobacterium halobium* and components of the photosynthetic reaction center of *Rhodopseudomonas viridis* (Henderson and Unwin, 1975; Henderson and Schertler, 1990; Diesenhofer *et al.*, 1985). Nevertheless, the burgeoning mass of primary sequence information shows that numerous other membrane proteins, including those involved in transport, also contain a series of hydrophobic regions, each sufficiently long to span the membrane as an α-helix, separated by hydrophilic regions (Kyte and Doolittle, 1982; von Heijne, 1988). Computerized algorithms have been derived to predict the occurrence of secondary structural features from primary sequence information (Garnier *et al.*, 1978; Eisenberg *et al.*, 1984; Eisenberg, 1985), and incorporated into models of the lactose (Foster *et al.*, 1983) and glucose (Mueckler *et al.*, 1985) transporters. Models of membrane protein structure based on such sequence data alone should be

viewed with caution for several reasons (Lodish, 1988). However, they do generate testable hypotheses of structural features—for example that the N-terminus and the C-terminus are on the same side of the membrane—and they set criteria for comparisons between different proteins, as illustrated above. When evidence is obtained that, for example, the C-terminal end of the mammalian glucose transporter is inside the cytoplasmic membrane (Davies *et al.*, 1987; Haspel *et al.*, 1988), or that the central hydrophilic region is also inside the membrane (Cairns *et al.*, 1987), the model is confirmed and the extrapolation to the related transporters becomes more tenable (see also Baldwin and Henderson, 1989).

Three examples of such models are given in Fig. 4, the arabinose/H^+ transporter of *E. coli*, the citrate/H^+ transporter, and the tetracycline/H^+ (K^+) transporter, the last two of which are encoded on plasmids (see also Eckert and Beck, 1989; Henderson and Maiden, 1990). Importantly, they include the positions of residues conserved in all the monosaccharide transporters and some sequence motifs (above) aparently repeated in similar positions in both halves of the citrate and tetracycline transporters. For the latter, experiments using labelling reagents and proteolytic cleavage indicated that the N-terminus and C-terminus are located inside the cytoplasmic membrane (Eckert and Beck, 1989), which is consistent with the results obtained with the glucose transporter and with the predictions in the models.

Sugar/H^+ Transport Proteins with Dissimilar Primary Sequences

The Lactose/H^+ Transport Protein of Escherichia coli

Primary Sequence of the Lac Y Protein. There is no obvious homology between the primary sequence of the lactose/H^+ transport protein of *E. coli* (Buchel *et al.*, 1980) and the sequences of any of the other sugar/cation transporters except the equivalent protein in *Klebsiella pneumoniae* (Buvinger and Riley, 1985; McMorrow *et al.*, 1988). Some possible similarities have been mentioned already, such as the predicted occurrence of twelve membrane-spanning segments (Foster *et al.*, 1983; Kaback, 1989), and the equivalent location of sequences similar to the R X G R R motif (Maiden *et al.*, 1987). Nevertheless, it is possible that the three-dimensional shape of the Lac Y protein will turn out to be similar to that of other sugar transport proteins, since there are several precedents for similarity of tertiary structure between proteins that are not homologous at the primary sequence level (Fersht, 1985; Quiocho, 1986, 1990). Since studies of the lactose/H^+ transport protein do lead the field, but have been extensively reviewed elsewhere (Kaback, 1986, 1987, 1989), it is appropriate here only to highlight a few selected areas, which may be relevant to future experimentation on the other sugar transport systems.

Specific Amino Acids within Putative Membrane-Spanning Helices 9 and 10 Are Essential for Sugar/H^+ Symport. Directed mutagenesis experiments imply that Arg 302, His 322, and Glu 325 in the Lac Y protein are intimately involved in coupling lactose and proton transport (Kaback, 1987, 1989). Until recently this was interpreted as their involvement only in H^+ movement. However, since mutants with His 322 changed to Tyr or Phe retain the ability to cotransport H^+ with sugar (King and Wilson, 1989a, b; Brooker, 1990), and since His 322 and Arg 302 mutants exhibit drastically lowered binding affinities for at least one sugar analogue, nitrophenyl-α-D-galactoside (Roepe *et al.*, 1990), it seems that an altered binding of the sugar molecule may equally explain the altered transport phenotype (Roepe *et al.*, 1990; Brooker, 1990). Neither Arg, His, nor Glu residues appear to be located in the equivalent regions of the homologous sugar transporters (Table III; Fig. 3). The possible contribution of other individual amino acid residues to the binding and translocation of H^+ ions by the Lac Y protein are discussed by Roepe *et al.* (1990) and Brooker (1990).

Glycine 24 and Proline 28 Are Necessary for Translocation, but Not Sugar Binding. Overath *et al.* (1987) obtained mutants of *E. coli* in which Gly 24 was changed to Glu or Arg and Pro 28 was changed to Ser. These were severely impaired in the ability to transport galactosides, but were able to bind them with wild-type, or greater, affinities. The conclusions from these experiments were well supported by measurements of the amount of transport protein in the membrane and the appropriate kinetic constants. Again this leads to the speculation that one or both of the highly conserved glycine residues in the first putative membrane-spanning loop of the aligned monosaccharide transporters (Fig. 3, 4) may also be involved in translocation, but not sugar binding.

Specific Amino Acids Are Associated with Sugar Binding. By some elegant *in vivo* mutagenesis experiments, individual amino acid residues have been associated with alternations in the specificity of sugar binding to Lac Y. Brooker (1990) has summarized these data and listed the residues as follows: Ala 177 (VI), Tyr 236 (immediately after VII), Thr 266 (VII), Ile 303 (IX), Ser 306 (IX), Lys 319 (X), His 322 (X), and Ala 389 (XII). the roman numerals in brackets refer to the predicted membrane-spanning α-helix proposed by Kaback and co-workers (Foster *et al.*, 1983; Kaback, 1986, 1987, 1989; Roepe and Kaback, 1989). It is possible that highly conserved residues, of identical nature and similar predicted location in the aligned sequences (Fig. 3), and proposed models (Fig. 4) of the homologous sugar transporters (Table III) are also involved in sugar binding.

Availability of the Purified Active Protein. Amplified expression of the *lacY* gene (Teather *et al.*, 1978) enabled subsequent purification of the apparently homogeneous Lac Y protein with high specific activity (Newman

et al., 1981; Kaback, 1989; cf. Wright *et al.*, 1983; Page *et al.*, 1988). Further amplification in an improved vector system led to the remarkable observation that the protein could be produced in a soluble form that retained sugar binding activity and could be reconstituted into lipid membranes with retention of transport activity (Roepe and Kaback, 1989). Purification of active protein is particularly important because biophysical techniques can then be exploited for the determination of the three-dimensional structure of the transport protein (Kaback, 1989).

The Melibiose/Cation Transport Protein of Escherichia coli

Primary Sequence of the MelB Protein. Since the sequence of the melibiose/cation transport protein of *E. coli* (Yazyu *et al.*, 1984) has similarities to the lactose transport protein of *Strep. thermophilus* (Poolman *et al.*, 1989), it is likely to be the first representative of a dispersed family of sugar transport proteins. There is no obvious homology to any other sequenced transport protein. However, recent reviews favor a structural model that also predicts twelve membrane-spanning helices (Leblanc *et al.*, 1989; Botfield *et al.*, 1990; but see Yazyu *et al.*, 1984, who suggest 10–11 helices), and between helices 2/3 and 8/9 are sequences similar to the R X G R R motif found in the transporters of Table III (discussed above). All these proteins are of similar size and display similar, anomalous, behaviour in SDS/polyacrylamide gels (below). Despite the differences in primary sequence, one can speculate that the melibiose transporter may therefore have a similar three-dimensional structure to many other transport proteins that are not thought to be related at the moment.

Interdependence of Sugar and Cation Specificity. The nature of the sugar determines which cation is the preferred cosubstrate of the MelB protein (Wilson and Wilson, 1987). For melibiose transport Na^+ or H^+ is utilized, whereas Na^+ or Li^+ is utilized for β-galactoside transport and Na^+, H^+, or Li^+ for α-galactoside transport. Also, the kinetic constants for the transport of melibiose depend upon the chemical identity of the coupling cation (Bassilana *et al.*, 1985; Pourcher *et al.*, 1990). Leblanc *et al.* (1989) concluded that, when the protein catalyzes melibiose/H^+ cotransport, the trans-membrane proton-motive force essentially affects the transport constant K_t, whereas when the same carrier catalyzes melibiose/Na^+ cotransport the trans-membrane sodium motive force selectively modifies the transport V_{max}. These observations are interpreted in terms of a structural interaction between the sugar and cation binding sites in the protein (Leblanc *et al.*, 1989; Botfield *et al.*, 1990), but it is important also to take kinetic considerations into account (see below).

Mutations That Alter Sugar and/or Cation Specificity. Yazyu *et al.* (1985) isolated a Li^+-resistant mutant of *E. coli* in which a point mutation

altered Pro 122 to a serine residue. The carrier lost the ability to cotransport H^+ with melibiose and showed an absolute requirement for Na^+ or Li^+. Botfield and Wilson (1988) isolated 23 independent TMG-resistant mutants altered in their sugar specificity. Most also displayed altered recognition for cations. The sites of mutation were clustered in four regions of the protein (Botfield and Wilson, 1988), Asp 15 through Ile 18 (I), Tyr 116 through Pro 122 (IV), Val 342 through Ile 348 (X), and Ala 364 through Gly 374 (hydrophilic region before XI). The roman numerals in brackets indicate the putative helix in a 12-helix model (Botfield *et al.*, 1990). The last two regions are of particular interest because their positions correspond to regions of Lac Y and the homologous sugar transporters that mutagenesis or sequence alignments implicate in the mechanism of sugar binding and/or transport (see above).

The L-Fucose/H^+ Transport Protein of Escherichia coli

Lu and Lin (1989) briefly reported the DNA sequence of the fucose operon of *E. coli*. By comparison with the complementation map of the genes they identified one of the open reading frames in the translated sequence (439 amino acids) as that of the fucose transport protein. This has not yet been confirmed by expression of this DNA, and it may not actually encode the fucose/H^+ transport activity described by Bradley *et al.* (1987). Nevertheless, the predicted amino acid sequence is very hydrophobic, characteristic of a membrane protein, but it does not exhibit homology to any of the known transport proteins, nor is it easy to predict the number of membrane-spanning helices. Its further characterization is in progress.

The L-Rhamnose/H^+ Transport Protein of Salmonella typhimurium

Muiry (1989) reported the predicted amino acid sequence of a protein of *S. typhimurium*, expression of which produced rhamnose/H^+ symport activity. This protein is very hydrophobic and contains ten clearly identifiable hydrophobic regions. It is not homologous to any of the other transport proteins. Investigations of its structure and activity are in progress.

The Glucose/Na^+ Transport Protein of Mammals

Hediger *et al.* (1987) used a novel cloning method of expressing rabbit mRNA in Xenopus oocytes to select and sequence the gene encoding the phlorizin-sensitive intestinal glucose/Na^+ symport activity. The predicted amino acid sequence of the protein is not homologous to any other characterized sugar transport protein. The rabbit DNA was then used as a probe to isolate the gene for the human intestinal glucose/Na^+ transporter, the

predicted amino acid sequence of which is 84% identical to the rabbit protein (Hediger *et al.*, 1989). At this stage it was found that the glucose/Na^+ transporter (664 residues) can be aligned with the *E. coli* proline/Na^+ transporter (502 residues, Nakao *et al.*, 1987) to reveal 28% identity between them, and conservative substitutions at a further 25% of residues (Hediger *et al.*, 1989). Although the proline/Na^+ transporter is predicted to have twelve membrane-spanning segments (Nakao *et al.*, 1987), Hediger *et al.* (1987, 1989) prefer a model for the glucose/Na^+ transporter with eleven. There is no homology between the mammalian glucose/Na^+ transporters and the *E. coli* melibiose/Na^+ transporter (see above).

As already mentioned, the sensitivity of one plant glucose/H^+ transporter to phlorizin (Gogarten and Bentrup, 1989) and cross-reactivity of another with antibodies to the mammalian glucose/Na^+ transporter (Rausch *et al.*, 1989), together with the above homologies, indicate the existence of yet another widely dispersed family of homologous substrate/cation transport proteins.

Identification of the Transport Proteins

The biochemical identification of a transport protein is an essential step in the characterization of any transport system. However, in bacteria such proteins usually constitute only 0.1–0.5% of the membrane proteins and are impossible to identify by conventional means. By combining genetical and biochemical techniques, several strategies have evolved to overcome this problem, and they are illustrated briefly here. The specific covalent labelling of a transport protein is also a most important aid in its purification (see, e.g., Newman *et al.*, 1981).

Amplification of Gene Expression

In bacteria, gene expression can be increased by factors of 2–5 by isolating mutants in which accumulation of the natural or gratuitous inducers is enhanced. Examples include galactokinase-negative, arabinose isomerase-negative, and xylose isomerase-negative strains of *E. coli*, in all of which expression of the corresponding sugar transporter was increased (Henderson and Macpherson, 1986). Some mutants may express the transport (and other metabolism genes) constitutively, and this level is usually higher than is achieved in "fully-induced" wild-type strains. Even these modest increases in expression depress the efficiency of cell growth for reasons that are not understood. More substantial levels of amplification require transfer of the transport gene to an appropriate vector, where its expression is suppressed during growth of the cells, and then activated.

Teather *et al.* (1978) were the first to achieve amplified expression of a bacterial sugar transport protein, Lac Y, on a multi-copy plasmid to the level where it could be identified as a Coomassie blue-stained protein of M_r about 30 000 by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The genes *araE* and *xylE* were recently transferred to vectors in which their amplified expression was achieved under the control of the efficient lambda $O_L P_L$ promotor (Maiden *et al.*, 1988; McKeown, 1988). Each gene product, AraE or XylE, was then identified as a Coomassie blue-stained protein of M_r 35–38,000 (Table I) constituting 15–20% and 5–7% of the membrane protein, respectively. Since the N-terminal sequences of the proteins were essentially as predicted from their DNA sequences, it was unlikely that they had undergone post-translational modification. The melibiose transporter MelB has proved more difficult to overexpress, but exploitation of the T7 bacteriophage RNA polymerase expression system developed by Tabor and Richardson (1985) and protein labelling with [^{35}S]-methionine revealed a single protein of M_r 39 000 (Leblanc *et al.*, 1989).

Expression in Maxi-Cells or Mini-Cells

Even when the amplification of expression is insufficiently successful, the availability of the cloned DNA on a vector means that the gene can be preferentially expressed in these special cells when the host cell protein synthesis is incapacitated (Stoker *et al.*, 1984). The cloned gene product, albeit present at low absolute levels, may then be specifically labelled with [^{35}S]-methionine.

Labelling with N-Ethylmaleimide

The proton-linked transporters for arabinose, xylose, or galactose from *E. coli* were all inhibited by *N*-ethylmaleimide (Table I). The protection by substrates was exploited to label specifically the susceptible transport proteins, AraE, XylE, and GalP, with radioactive *N*-ethylmaleimide. They were then identified as proteins that migrated with an apparent M_r of 35,000–40,000 in SDS-PAGE (Macpherson *et al.*, 1981, 1983; Henderson *et al.*, 1983b; Henderson and Macpherson, 1986; Davis, 1986; Maiden *et al.*, 1988), entirely consistent with the values observed in the amplified expression systems (above). By similar procedures the lactose/H^+ transport protein was the first to be identified using *N*-ethylmaleimide, with an apparent M_r of about 30 000 (Jones and Kennedy, 1969). Similarly, the proline/Na^+ transporter was identified as a protein of apparent M_r about 35,000 (Hanada *et al.*, 1985). Such labelling experiments have been very important for the identification of hydrophobic transport proteins that are of low abundance (0.1–0.5% of the membrane proteins), though they are being superseded by the expression systems.

Labelling with Photoactivatable Substrate Analogues and Cytochalasin B

It has already been mentioned how Kaczorowski *et al.* (1980) found that 4-nitrophenyl-α-D-galactoside (NPG) reacted covalently, and specifically, with the Lac Y protein when the two were exposed to UV light. Similarly, when activated by UV light the AraE and GalP transport proteins react covalently with radioactive cytochalasin B, and they can again be identified by SDS/PAGE as labelled proteins of apparent M_r 36,000–38,000. The presence of appropriate sugar substrates in the incubation mixture protected the transporters against this labelling (Smith, unpublished data; Cairns *et al.*, 1989b). Cytochalasin B will be a particularly convenient tool in purification studies on these proteins.

Anomalous Migration in SDS/Polyacrylamide Gel Electrophoresis

An important general point, which has been made elsewhere (Leblanc *et al.*, 1989), is that all these sugar transport proteins, and the citrate, tetracycline, and proline transporters (Reynolds and Silver, 1983; Eckert and Beck, 1989; Hanada *et al.*, 1985), migrate in SDS/PAGE with an apparent M_r less than that predicted from the DNA sequence. The values for the sugar transporters are summarized in Table I. In the case of several of them there is evidence that no post-translational modification has occurred, which might have accounted for the phenomenon. It may be due to the binding of higher proportions of SDS than occurs with less hydrophobic proteins or to the retention of some secondary/tertiary structure.

Reconstitution

The important general point is that membrane transport proteins of a variety of primary sequences can be solubilized under very similar conditions (usually with added lipid in 50 mM phosphate buffer, pH 7.5, 1 mM dithiothreitol) in the same detergent, octyglucoside (1.25% concentration, dodecylmaltoside is often a good substitute), and then successfully reconstituted by the convenient detergent dilution technique into proteoliposomes without much loss of activity. The pertinent examples are the mammalian glucose transporter (Kasahara and Hinkle, 1977) and the *E. coli* transporters for lactose/H^+ (Newman and Wilson, 1980), melibiose/Na^+ (Tsuchiya *et al.*, 1982), galactose/H^+ (Henderson *et al.*, 1983a), arabinose/H^+ (Henderson and Macpherson, 1986), and proline/Na^+ (Hanada *et al.*, 19988a, b).

The availability of a reconstitution assay is particularly important for purification of membrane transport proteins. Purification has been successful for the mammalian glucose transporter (Kasahara and Hinkle, 1977;

Baldwin *et al.*, 1982), the *E. coli* lactose/H^+ transporters (Newman *et al.*, 1981; Wright *et al.*, 1983; Viitanen *et al.*, 1986; Page *et al.*, 1988), and the *E. coli* proline/Na^+ transporter (Hanada *et al.*, 1988b). It seems reasonable to expect that methods based on those used for these proteins will be successfully applied to the other sugar transport proteins described in this article when their expression is amplified sufficiently.

Conclusions

Evolutionary and Structural Relationships

The family of sugar transport proteins listed in Table III and described in this article has been found in such diverse organisms so far that it seems likely to occur in most forms of life. This implies that investigations of the molecular mechanism of their transport activities in microorganisms, with the attendant conveniences of experimental manipulation, will uncover features equally valid for their operation in higher organisms, including man. It seems that insight into the mechanism of transport of substrates other than sugars will also accrue, if the apparent similarities to the citrate and tetracycline (and other?) transporters described here are real. Also, the homologies described imply a fundamental similarity between proteins that catalyze facilitated diffusion ("passive transport") and substrate/H^+ symport ("secondary active transport")—a perceived difference in mechanism seems not to require a profound difference in structure.

The purpose of cataloguing the substrate and inhibitor specificities of the transport proteins (Table II) is to gain insight into the nature of the ligand binding sites and to define the relationship between the proteins. The availability of a variety of sugars and their structural analogues, and the known sequence homologies of the arabinose, xylose, galactose, and glucose transporters (Table III), make this a particularly promising series of membrane proteins in which to establish the characteristics of sugar recognition and their evolution. If the relatively unspecific glucose/galactose hexose transporters arose first, one can speculate that they then evolved to discriminate against the 6-OH residue and recognize the precise orientation of the 4-OH residue in order to become more specific for either of the pentoses, arabinose or xylose. When the substrate specificities of the homologous lactose and maltose transporters of yeast (Chang and Dickson, 1988; Yao *et al.*, 1989; Sczkutnicka *et al.*, 1989) are known in more detail, then it may be possible to learn about molecular discrimination between mono- and disaccharides.

There also exist transport proteins with the same function of substrate/cation symport, but without any overall similarity in primary sequence, i.e., the lactose/H^+, melibiose/Na^+, fucose/H^+, rhamnose/H^+, and proline/Na^+

transporters of Enterobacteriaceae and the glucose/Na^+ transporter of mammalian intestine. Perhaps these arose by convergent evolution, and their tertiary structures and molecular mechanisms are actually similar to those of the homologous series. Amongst soluble proteins there are several examples of those with dissimilar primary sequences that nevertheless have similar tertiary structures and molecular mechanisms, e.g., the periplasmic binding proteins (Quiocho, 1986, 1990) and the serine proteases (Fersht, 1985).

The Roles of Individual Amino Acids in the Transport Proteins

The structural and functional roles of individual amino acids may be evaluated by modification with appropriate reagents (e.g., *N*-ethylmaleimide for cysteine and diethylpyrocarbonate for histidine residues) and *in vivo* and *in vitro* mutagenesis (Botfield and Wilson, 1988; Kaback, 1987; Leblanc *et al.*, 1989; Kaback, 1990; Brooker, 1990; Botfield *et al.*, 1990). The initial difficulty of identifying the role of an *individual* amino acid is best overcome by exploiting the *in vivo* selection of mutants impaired in an identifiable function as exemplified by experiments on the lactose/H^+ and melibiose/Na^+ symporters described above. The comparisons of aligned amino acid sequences (Figs. 3, 4) serves to identify how essential *each and every one* of the residues is, but without determining their function. Directed mutagenesis studies, which are expensive and time-consuming, are probably best undertaken after other techniques have indicated the significance of a particular residue or sequence motif. Obviously, these approaches will be combined in future.

Such genetical strategies are the most practicable for relating structure and function of membrane proteins at the present time, but they are unlikely to yield a molecular model of sufficient detail to understand entirely the different substrate specificities, the process of cation recognition, the mechanism of inhibitor action, the translocation mechanism, and the evolution of the transport proteins. This will also require knowledge of their three-dimensional structures.

Kinetic Studies

Since we can now manipulate the structure of transport proteins, it is profoundly important to characterize properly the resultant changes in individual partial reactions, substrate specificity, inhibitor specificity, transport parameters, reaction with covalent modifiers, etc. Kinetically rigorous analyses and statistically valid comparisons between activities in mutants and wild-type proteins will enable real insights into molecular mechanism to be made.

Steady-state kinetic studies of the transport activities described in this article are generally based on models involving the ordered combination of the transport protein with first the cation and then the sugar substrate

followed by the translocation step and an ordered dissociation of substrates on the other side of the membrane (see, e.g., Leblanc *et al.*, 1989; Kaback, 1989). For the lactose/H^+ symporter the unloaded protein is perceived to carry a net negative charge and the loaded carrier is uncharged (Kaback, 1989), while for the melibiose/Na^+ carrier the unloaded protein is perceived to be uncharged and the loaded carrier has a net positive charge (Leblanc *et al.*, 1989). In both cases the stoichiometry appears to be one mole of substrate transported with one mole of cation (West and Mitchell, 1973; Page *et al.*, 1988; Pourcher *et al.*, 1989). More general treatments of kinetic mechanisms for transport proteins may be accessed through articles by Stein (1986), Walmsley (1988), Eddy (1989), and Severin *et al.* (1989).

It is very important that the kinetic parameters of novel mutations in transport proteins are determined rigorously and compared with features of the defined kinetic mechanism in the wild-type protein. Otherwise conclusions about the effect of the mutation on sugar specificity, or cation recognition, or translocation steps, or energization mechanism may be at the least naive and at the worst incorrect. For example, in a steady-state ordered mechanism for a two-substrate reaction without any steps at thermodynamic equilibrium the apparent K_m measured for one substrate depends upon the concentration of the other (Cleland, 1963; Fromm, 1979). Only when one substrate is at saturating concentrations (a situation that is very difficult to achieve in practice, but very easily extrapolated from measurements in a practicable range of substrate concentrations) is the K_m determined for the second substrate the true value (Cleland, 1963; Fromm, 1979). If a mutation alters the true (and apparent!) K_m for the sugar substrate, then the apparent K_m for the cation will almost always also be changed for the steady-state mechanisms outlined above. This does not necessarily imply that there is a profound molecular link between the binding of the two ligands. Indeed, the proper measurement of the true K_m for the cation may well find that the value is unchanged. Furthermore, it is dangerous to compare only K_m values, when a much more rigorous criterion for the affinity of a protein for its substrates is the ratio of k_{cat}/K_m (Fersht, 1985).

Such basic considerations have usually been applied in kinetic studies on wild-type bacterial transport proteins, but the ability to generate interesting mutants seems sometimes to outpace the resources or will to characterize them quantitatively, with a consequent dimunition in the validity of the conclusions drawn about them. The following rules should perhaps be implemented in future studies:

1. The steady-state mechanism of the wild-type transport protein should be defined.
2. True K_m and k_{cat} values should be used in comparative studies, not apparent K_m and V_{max} values.

3. The numerical values of kinetic parameters should be determined by statistically rigorous methods quoting standard deviations (Cleland, 1979; Cornish-Bowden, 1979; Henderson, 1985).
4. Rigorous analytical methods should also be applied in determining dissociation constants and concentrations of binding sites in equilibrium binding measurements.

Determination of the Three-Dimensional Structures of the Proteins

A real understanding of the mechanism of transport is critically dependent upon determination of the three-dimensional structure of transport proteins at the level of atomic resolution. Tentative models of transport protein structure such as those in Fig. 4 are the basis of present and future experiments in which "topological" protein reagents and antibodies, gene fusions (e.g, Cairns *et al.*, 1984; Carrasco *et al.*, 1986; Manoil and Beckwith, 1986; Broome-Smith and Spratt, 1986; Page and Rosenbusch, 1988; Eckert and Beck, 1989), and other techniques are used to explore the actual folding of the protein in the membrane. However, the determination of a more detailed three-dimensional structure probably requires crystallization and X-ray diffraction analysis. The first step is the purification of sufficient stable, undenatured protein, which has been achieved only for the lactose/H^+ transporter (Page *et al.*, 1988; Kaback, 1989) and the erythrocyte glucose transporter (Walmsley, 1988). Crystallization may be difficult to accomplish for all these membrane transport proteins, but perhaps no more so than the apparent impossibility of establishing their primary sequences that we faced a decade ago.

Acknowledgments

Research in the author's laboratory is supported by the SERC, the MRC, The Smith Kline Foundation, and The Wellcome Trust. I am indebted to Dr. S. A. Baldwin, Dr. M. T. Cairns, Dr. B. M. Charalambous, Dr. E. O. Davis, Prof. H. R. Kaback, Dr. G. Leblanc, Dr. V. A. Lucas, Mr. T. MacDonald, Mr. B. J. McKeown, Dr. D. C. M. Moore, Ms. J. A. R. Muiry, Ms. K. Petro, Mr. P. E. Roberts, Dr. G. Schmetterer, Dr. K. P. Shatwell, Dr. G. Smith, Dr. C. G. Tate, Prof. W. Tanner, and Prof. T. H. Wilson for unpublished information and helpful discussions. Drs. C. G. Tate and K. P. Shatwell provided invaluable help in the preparation of the manuscript.

References

Bachmann, B. J. (1987). In *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed.), ASM, Washington, pp. 807–876.

Badia, J., Baldoma, L., Aguilar, J., and Boronat, A. (1989). *FEMS Microbiol. Lett.* **65**, 253–258.
Baldwin, S. A., and Henderson, P. J. F. (1989). *Annu. Rev. Physiol.* **51**, 459–471.
Baldwin, J. M., Lienhard, G. E., and Baldwin, S. A. (1980). *Biochim. Biophys. Acta* **599**, 699–714.
Baldwin, S. A., Baldwin, J. M., and Lienhard, G. E. (1982). *Biochemistry* **21**, 3836–3842.
Baly, D. L., and Horuk, R. (1988). *Biochim. Biophys. Acta* **947**, 571–590.
Barnett, J. E. G., Holman, G. D., and Munday, K. A. (1973). *Biochem. J.* **131**, 211–231.
Bassilana, M., Damiano-Ferano, E., and Leblanc, G. (1985). *Biochem. Biophys. Res. Commun.* **129**, 626–631.
Beauclerk, A. D. D., and Smith, A. J. (1978). *Eur. J. Biochem.* **82**, 187–197.
Beckwith, J. R., and Zipser, D. (1970). *The Lactose Operon*, Cold Spring Harbor Laboratory, New York.
Beyreuther, K., Bieseler, B., Ehring, R., and Muller-Hill, B. (1981). In *Methods in Protein Sequence Analysis* (Elzina, M., ed.), Humana Press, Clifton, New Jersey, pp. 139–148.
Birnbaum, M. J. (1989). *Cell* **57**, 305–315.
Birnbaum, M. J., Haspel, H. C., and Rosen, O. M. (1986). *Proc. Natl. Acad Sci. U.S.A.* **83**, 5784–5788.
Bloch, R. (1974). *J. Biol. Chem.* **249**, 1814–1822.
Botfield, M. C., and Wilson, T. H. (1988). *J. Biol. Chem.* **263**, 12909–12915.
Botfield, M. C., Wilson, D. M., and Wilson, T. H. (1990). *Res. Microbiol.* **141**, 328–331.
Bradley, S. A., Tinsley, C. R., Muiry, J. A. R., and Henderson, P. J. F. (1987). *Biochem. J.* **248**, 495–500.
Bremer, E., Silhavy, T. J., Weisemann, J. M., and Weinstock, G. M. (1984). *J. Bacteriol.* **158**, 1084–1093.
Brooker, R. J. (1990). *Res. Microbiol.* **141**, 309–315.
Brooker, R. J., and Wilson, T. H. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 3959–3963.
Broome-Smith, J. K., and Spratt, B. G. (1986). *Gene* **49**, 341–349.
Buchel, D. E., Gronenborn, B., and Muller-Hill, B. (1980). *Nature* (*London*) **283**, 541–545.
Buttin (1968). *Adv. Enzymol. Relat. Areas Mol. Biol.* **30**, 81–137.
Buvinger, W. E., and Riley, M. (1985). *J. Bacteriol.* **163**, 850–857.
Cairns, M. T., Elliot, D. A., Scudder, P. R., and Baldwin, S. A. (1984). *Biochem. J.* **221**, 179–188.
Cairns, M. T., Alvarez, J., Panico, M., Gibbs, A. F., Morris, H. R., Chapman, D., and Baldwin, S. A. (1987). *Biochim. Biophys. Acta* **905**, 295–310.
Cairns, B. R., Collard, M. W., and Landfear, S. M. (1989a). *Proc. Natl. Acad. Sci. USA* **86**, 7682–7686.
Cairns, M. T., Smith, G., Henderson, P. J. F., and Baldwin, S. A. (1989b). *Biochem. Soc. Trans.* **17**, 552–553.
Carrasco, N., Herzlinger, D., Danho, W., and Kaback, H. R. (1986). *Methods Enzymol.* **125**, 453–467.
Carter, J. R., Fox, C. F., and Kennedy, E. P. (1968). *Proc. Natl. Acad. Sci. USA* **60**, 725–732.
Carter-Su, C., Pessin, J. E., Mora, R., Gitomer, W., and Czech, M. P. (1982). *J. Biol. Chem.* **257**, 5419–5425.
Celenza, J. L., Marshall-Carlson, L., and Carlson, M. (1988). *Proc. Natl. Acad. Sci. USA* **85**, 2130–2134.
Chang, Y.-D., and Dickson, R. C. (1988). *J. Biol. Chem.* **263**, 16696–16703.
Charalambous, B. M., Maiden, M. C. J., McDonald, T. P., Cunningham, I. J., and Henderson, P. J. F. (1989). *Biochem. Soc. Trans.* **17**, 441–444.
Ciaraldi, T. P., Horuk, R., and Matthaei, S. (1986). *Biochem. J.* **240**, 115–123.
Cleland, W. W. (1963). *Biochim. Biophys. Acta* **67**, 104–137.
Cleland, W. W. (1979). *Methods Enzymol.* **63**, 103–138.
Cooper, R. A. (1986). In *Carbohydrate Metabolism in Cultured Cells* (Morgan, M. J., ed.), Plenum Press, London, pp. 461–491.
Cornish-Bowden, A. (1979). *Fundamentals of Enzyme Kinetics*, Butterworths, London.
Daie, J. (1989). *Plant Mol. Biol. Rep.* **7**, 106–115.
Damiano-Forano, E., Bassilana, M., and Leblanc, G. (1986). *J. Biol. Chem.* **261**, 6893–6899.
Daruwalla, K. R., Paxton, A. T., and Henderson, P. J. F. (1981). *Biochem. J.* **200**, 611–627.
Davies, A., Meeran, K., Cairns, M. T., and Baldwin, S. A. (1987). *J. Biol. Chem.* **262**, 9347–9352.

Davis, E. O. (1986). Ph.D. thesis, University of Cambridge.
Davis, E. O., and Henderson, P. J. F. (1987). *J. Biol. Chem.* **262**, 13928–13932.
Davis, E. O., Jones-Mortimer, M. C., and Henderson, P. J. F. (1984). *J. Biol. Chem.* **259**, 1520–1525.
Deves, R., and Krupka, R. M. (1978). *Biochim. Biophys. Acta* **510**, 339–348.
Deziel, M., Pegg, W., Mack, E., Rothstein, A., and Klip, A. (1984). *Biochim. Biophys. Acta* **772**, 403–406.
Diesenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985). *Nature (London)* **318**, 618–624.
Doolittle, R. F. (1981). *Science* **214**, 149–159.
Dosch, D., Salvacion, F., and Epstein, W. (1984). *J. Bacteriol.* **160**, 1188–1190.
Eckert, B., and Beck, C. F. (1989). *J. Biol. Chem.* **264**, 11663–11670.
Eddy, A. A. (1982). *Adv. Microbial. Physiol.* **23**, 1–78.
Eddy, A. A. (1989). In press.
Eiglmeier, K., Boos, W., and Cole, S. T. (1987). *Mol. Microbiol.* **1**, 251–258.
Eisenberg, D. (1985). *Annu. Rev. Biochem.* **53**, 595–623.
Eisenberg, D., Schwartz, E., Komaromy, M., and Wall, R. (1984). *J. Mol. Biol.* **179**, 125–142.
Fersht, A. R. (1985). *Enzyme Structure and Mechanism.* Wiley, New York, pp. 317–331.
Flores, E., and Schmetterer, G. (1986). *J. Bacteriol.* **166**, 693–696.
Foster, D. L., Boublik, M., and Kaback, H. R. (1983). *J. Biol. Chem.* **258**, 31–34.
Fox, C. F., and Kennedy, E. P. (1965). *Proc. Natl. Acad. Sci. USA* **54**, 891–899.
Friedrich, M. J., and Kadner, R. J. (1987). *J. Bacteriol.* **169**, 3556–3563.
Fromm, H. J. (1979). *Methods Enzymol.* **63**, 42–53.
Fukumoto, H., Kayano, T., Buse, J. B., Edwards, Y., Pilch, P. F., Bell, G. I., and Seino, S. (1989). *J. Biol. Chem.* **264**, 7776–7779.
Fukumoto, H., Seino, S., Imura, H., Seino, Y., Eddy, R. L., Fukushima, Y., Byers, M. G., Shows, T. B., and Bell, G. I. (1988). *Proc. Natl. Acad. Sci. USA* **85**, 5434–5438.
Furlong, C. E. (1987). In *Escherichia coli and Salmonella typhimurium*' (Neidhardt, F. C., ed.), ASM, Washington, pp. 768–796.
Garnier, J., Osguthorpe, D. J., and Rosen, B. (1978). *J. Mol. Biol.* **120**, 97–120.
Gogarten, P. G., and Bentrup, F.-W. (1989). *Planta* **178**, 52–60.
Griffin, J. F., Rampal, A. L., and Jung, C. Y. (1982). *Proc. Natl. Acad. Sci. USA* **79**, 3759–3763.
Hanada, K., Yamato, I., and Anraku, Y. (1985). *FEBS Lett.* **191**, 278–282.
Hanada, K., Yamato, I., and Anraku, Y. (1988a). *Biochim. Biophys. Acta* **939**, 282–288.
Hanada, K., Yamato, I., and Anraku, Y. (1988b). *J. Biol. Chem.* **263**, 7181–7185.
Haspel, H. C., Rosenfeld, M. G., and Rosen, O. M. (1988). *J. Biol. Chem.* **263**, 398–403.
Hediger, M. A., Coady, M. J., Ikeda, T. S., and Wright, E. M. (1987). *Nature (London)* **330**, 379–381.
Hediger, M. A., Turk, E., and Wright, E. M. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 5748–5752.
Henderson, R., and Unwin, P. N. T. (1975). *Nature (London)* **257**, 28–32.
Henderson, P. J. F. (1985). In *Techniques in Protein and Enzyme Biochemistry*, Part II Supplement (Tipton, K. F., ed.), Elsevier, Dublin, pp. 1–48.
Henderson, P. J. F. (1986). In *Carbohydrate Metabolism in Cultured Cells* (Morgan, M. J., ed.), Plenum Press, London, pp. 409–460.
Henderson, P. J. F., and Kornberg, H. L. (1975). In *Energy Transformation in Biological systems*, Ciba Foundation Symposium Vol. 31, Elsevier, Amsterdam, pp. 243–269.
Henderson, P. J. F., and Macpherson, A. J. S. (1986). *Methods Enzymol.* **125**, 387–429.
Henderson, P. J. F., and Maiden, M. C. J. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 391–410.
Henderson, R., and Schertler, G. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 379–390.
Henderson, P. J. F., Giddens, R. A., and Jones-Mortimer, M. C. (1977). *Biochem. J.* **162**, 309–320.
Henderson, P. J. F., Hirata, H., and Kagawa, Y. (1983a). *Biochim. Biophys. Acta* **732**, 204–209.
Henderson, P. J. F., Hirata, H., Horne, P., Jones-Mortimer, M. C., Kaethner, T. E., and Macpherson, A. J. S. (1983b). In *Biochemistry of Metabolic Processes* (Lennon, D. L. F., Stratman, F. W., and Zahlten, R. N., eds.), Elsevier, New York, pp. 339–352.

Hickey, M. W., Hiller, A. J., and Jago, G. R. (1986). *Appl. Environ. Microbiol.* **51**, 825–831.
Higgins, C. J. (1989). *Nature (London)* **341**, 103.
Higgins, C. F., Gallagher, M. P., Hyde, S. C., Mimmack, M. L., and Pearce, S. R. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 353–366.
Hillen, W., and Schollmeier, K. (1983). *Nucleic Acids Res.* **11**, 525–539.
Holman, G. D., and Rees, W. D. (1987). *Biochim. Biophys. Acta* **897**, 395–405.
Horne, P. (1980). Ph.D. Thesis, University of Cambridge.
Horne, P., and Henderson, P. J. F. (1983). *Biochem. J.* **210**, 699–705.
Hutchings, V. (1978a). *Planta* **138**, 229–235.
Hutchings, V. (1978b). *Planta* **138**, 237–241.
Ishiguro, N., and Sato, G. (1985). *J. Bacteriol.* **164**, 977–982.
James, D. E., Strube, M., and Mueckler, M. (1989). *Nature (London)* **338**, 83–87.
Jones, T. D. H., and Kennedy, E. P. (1969). *J. Biol. Chem.* **244**, 5981–5987.
Jones-Mortimer, M. C., and Henderson, P. J. F. (1986). *Methods Enzymol.* **125**, 157–180.
Joset, F., Buchou, T., Zhang, C.-C., and Jeanjean, R. (1988). *Arch. Microbiol.* **149**, 417–421.
Jund, R., Weber, E., and Chevallier, M.-R. (1988). *Eur. J. Biochem.* **171**, 417–424.
Jung, C. Y., and Rampal, A. L. (1977). *J. Biol. Chem.* **252**, 5456–5463.
Kaback, H. R. (1972). *Biochim. Biophys. Acta* **265**, 367–416.
Kaback, H. R. (1974). *Science* **186**, 882–892.
Kaback, H. R. (1986). *Ann. N.Y. Acad. Sci.* **456**, 291–304.
Kaback, H. R. (1987). *Biochemistry* **26**, 2071–2076.
Kaback, H. R. (1989). *Harvey Lect.* **83**, 77–105.
Kaback, H. R. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 425–436.
Kaczorowski, G. J., Leblanc, G., and Kaback, H. R. (1980). *Proc. Natl. Acad. Sci. USA* **77**, 6319–6323.
Kaestner, K. M., Christy, R. J., McLenithan, J. C., Braiterman, L. T., Cornelius, P., Pekala, P. M., and Lane, M. D. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 3150–3154.
Karim, A. R., Rees, W. D., and Holman, G. D. (1987). *Biochim. Biophys. Acta* **902**, 402–405.
Kartner, N. and Ling, V. (1989). *Sci. Am.*, March, 26–33.
Kasahara, M., and Hinkle, P. (1977). *J. Biol. Chem.* **252**, 7384–7390.
Kayano, T., Fukumoto, H., Eddy, R. L., Fan, Y. S., Byers, M. G., Shows, T. B., and Bell, G. I. (1988). *J. Biol. Chem.* **263**, 15245–15248.
Kennedy, E. P. (1970). In *The Lactose Operon* (Beckwith, J. R., and Zipser, D., eds.), Cold Spring Harbor, New York, pp. 49–92.
King, S. C., and Wilson, T. H. (1989a). *J. Biol. Chem.* **264**, 7390–7394.
King, S. C., and Wilson, T. H. (1989b). *Biochim. Biophys. Acta* **982**, 253–264.
Kohara, Y., Akiyama, K., and Isono, K. (1987). *Cell* **50**, 495–508.
Kolodrubetz, D., and Schleif, R. F. (1981). *J. Mol. Biol.* **151**, 215–227.
Komor, E., and Tanner, W. (1974). *Eur. J. Biochem.* **44**, 219–223.
Kosiba, B. E., and Schleif, R. F. (1982). *J. Mol. Biol.* **156**, 53–56.
Kornberg, H. L. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 505–514.
Kotyk, A. (1983). *J. Bioenerg. Biomembr.* **15**, 307–319.
Kyte, J., and Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.
Lam, V. M. S., Daruwalla, K. D., Henderson, P. J. F., and Jones-Mortimer, M. C. J. (1980). *J. Bacteriol.* **143**, 396–402.
Leblanc, G., Pourcher, T., and Bassilana, M. (1989). *Biochimie* **71**, 969–979.
Lefevre, P. G. (1961). *Pharmacol. Rev.* **13**, 39–70.
Lengeler, J. W., Titgemeyer, F., Vogler, A. P., and Wohrl, B. M. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 489–504.
Levy, S. (1984). In *Antimicrobial Drug Resistance* (Brian, L. E., ed.), Academic Press, New York, pp. 191–240.
Levy, S. (1988). *ASM News* **54**, 418–421.
Lin, E. C. C. (1987). In *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., ed.), ASM, Washington, pp. 244–284.
Lodish, H. F. (1988). *Trends Biochem. Sci.* **13**, 332–334.
Lombardi, F. J. (1981). *Biochim. Biophys. Acta* **649**, 661–679.

Lowe, A. G., and Walmsley, A. R. (1986). *Biochim. Biophys. Acta* **857**, 146–154.
Lu, Z., and Lin, E. C. C. (1989). *Nucleic Acids Res.* **17**, 4883–4884.
Macpherson, A. J. S., Jones-Mortimer, M. C., and Henderson, P. J. F. (1981). *Biochem. J.* **196**, 269–283.
Macpherson, A. J. S., Jones-Mortimer, M. C., Horne, P., and Henderson, P. J. F. (1983). *J. Biol. Chem.* **258**, 4390–4396.
Maiden, M. C. J. (1987). Ph.D. thesis, University of Cambridge.
Maiden, M. C. J., Davis, E. O., Baldwin, S. A., Moore, D. C. M., and Henderson, P. J. F. (1987). *Nature (London)* **325**, 641–643.
Maiden, M. C. J., Jones-Mortimer, M. C., and Henderson, P. J. F. (1988). *J. Biol. Chem.* **263**, 8003–8010.
Maloney, P. C. (1990). *Philos. Trans. R. Soc. London Ser. B* **326**, 437–454.
Manoil, C., and Beckwith, J. (1986). *Science* **233**, 1403–1408.
Markgraf, M., Bocklage, H., and Muller-Hill, B. (1985). *Mol. Gen. Genet.* **198**, 473–475.
McKeown, B. J. (1988). CPGS thesis, University of Cambridge.
McMorrow, I., Chin, D. T., Fiebig, K., Pierce, J. L., Wilson, D. M., Reeve, E. C. R., and Wilson, T. H. (1988). *Biochim. Biophys. Acta* **945**, 315–323.
McMurry, L. M., Petrucci, R. R., and Levy, S. B. (1980). *Proc. Natl. Acad. Sci. USA* **77**, 3974–3977.
Menick, D., Lee, J. A., Brooker, R. J., Wilson, T. H., and Kaback, H. R. (1987). *Biochemistry* **26**, 1132–1136.
Mitchell, P. (1961). *Nature (London)* **191**, 144–148.
Mitchell, P. (1963). *Biochem. Soc. Symp.* **22**, 142–169.
Mitchell, P. (1973). *J. Bioenerg.* **4**, 63–91.
Mueckler, M., Caruso, C., Baldwin, S. A., Panico, M., Blench, I., Morris, H. R., Allard, W. J., Leinhard, G. E., and Lodish, H. (1985). *Science* **229**, 941–945.
Muiry, J. A. R. (1989). Ph.D. Thesis, University of Cambridge.
Nakao, T., Yamato, I., and Anraku, Y. (1987). *Mol. Gen. Genet.* **208**, 70–75.
Newman, M. J., and Wilson, T. H. (1980). *J. Biol. Chem.* **255**, 10583–10586.
Newman, M. J., Foster, D. L., Wilson, T. H., and Kaback, H. R. (1981). *J. Biol. Chem.* **256**, 11804–11808.
Nordlie, R. C. (1985). In *Metabolic Regulation* (Ochs, R. S., Hanson, R. W., and Hall, J., eds.), Elsevier, Amsterdam, pp. 60–69.
Overath, P., Weigel, U., Neuhaus, J.-M., Soppa, J., Seckler, R., Riede, I., Bocklage, H., Muller-Hill, B., Aichelle, G., and Wright, J. K. (1987). *Proc. Natl. Acad. Sci. USA* **84**, 5535–5539.
Page, M. G., and Rosenbusch, J. P. (1988). *J. Biol. Chem.* **263**, 15906–15914.
Page, M. G., Rosenbusch, J. P., and Yamato, I. (1988). *J. Biol. Chem.* **263**, 15897–15905.
Patel, L., Garcia, M. L., and Kaback, H. R. (1982). *Biochemistry* **21**, 5805–5810.
Peden, K. W. (1983). *Gene* **22**, 277–280.
Petro (1988). M. Philos. Thesis, University of Cambridge.
Poolman, B., Royer, T. J., Mainzer, S. E., and Schmidt, B. F. (1989). *J. Bacteriol.* **171**, 244–253.
Postma, P. W., and Lengeler, J. W. (1985). *Microbiol. Rev.* **49**, 232–269.
Pourcher, T., Bassilana, M., Sarkar, H. K., Kaback, H. R., and Leblanc, G. (1990). *Philos. Trans. R. Soc. London Ser. B.* **326**, 411–424.
Quiocho, F. A. (1986). *Annu. Rev. Biochem.* **55**, 287–315.
Quiocho, F. A. (1990). *Philos. Trans. R. Soc. London. Ser. B* **326**, 341–352.
Raboy, B., and Padan, E. (1978). *J. Biol. Chem.* **253**, 3287–3291.
Rampal, A. L., and Jung, C. Y. (1987). *Biochim. Biophys. Acta* **896**, 287–294.
Rausch, T., Raszeja-Specht, A., and Koepsell, H. (1989). *Biochim. Biophys. Acta* **985**, 133–138.
Rees, W. D., and Holman, G. D. (1981). *Biochim. Biophys. Acta* **646**, 251–260.
Reynolds, C. H., and Silver, S. (1983). *J. Bacteriol.* **156**, 1019–1024.
Riordan, C., and Kornberg, H. L. (1977). *Proc. R. Soc. London Ser. B* **198**, 401–410.
Riordan *et al.* (1989). *Science* **245**, 1066–1073.
Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., and Stanier, R. Y. (1979). *J. Gen. Microbiol.* **111**, 1–61.

Roepe, P. D., and Kaback, H. R. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 6087–6091.
Roepe, P. D., Consler, T. J., Menezes, M. E., and Kaback, H. R. (1990). *Res. Microbiol.* **141**, 290–308.
Romano, A. H. (1986). In *Carbohydrate Metabolism in Cultured Cells* (Morgan, M. J., ed.), Plenum Press, London, pp. 225–244.
Saier, M. H. (1985). *Mechanisms and Regulation of Carbohydrate Transport in Bacteria*, Academic Press, New York.
Sasatsu, M., Misra, T. K., Chu, L., Ladagu, R., and Silver, S. (1985). *J. Bacteriol.* **164**, 983–993.
Sauer, N., and Tanner, W. (1989). *FEBS Lett.* **259**, 43–46.
Severin, J., Langel, P., and Hofer, M. (1989). *J. Bioenerg. Biomembr.* **21**, 321–334.
Shanahan, M. F. (1982). *J. Biol. Chem.* **257**, 7290–7293.
Silhavy, T. J., Berman, M. L., and Enquist, L. W. (1984). *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, New York.
Stein, W. (1986). *Transport and Diffusion across Cell Membranes*, Academic Press, Orlando, Florida.
Stoker, N. G., Pratt, J. M., and Holland, I. B. (1984). In *Transcription and Translation—A Practical Approach* (Hames, B. D., and Higgins, S. J., eds.), IRL Press, Oxford, pp. 153–177.
Stoner, C., and Schleif, R. F. (1983). *J. Mol. Biol.* **171**, 369–381.
Sumiya, M. (1989). Ph.D. thesis, University of Cambridge.
Szkutnicka, K., Tschopp, J. F., Andrews, L., and Cirillo, V. P. (1989). *J. Bacteriol.* **171**, 4486–4493.
Tabor, S., and Richardson, C. C. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 1074–1079.
Teather, R. M., Muller-Hill, B., Abrutsch, U., Aichele, G., and Overath, P. (1978). *Mol. Gen. Genet.* **159**, 239–248.
Thorens, B., Sarkar, H. K., Kaback, H. R., and Lodish, H. F. (1988). *Cell* **55**, 281–291.
Tobin, J. F., and Schleif, R. F. (1987). *J. Mol. Biol.* **196**, 789–799.
Tsuchiya, T., Ottina, K., Moriyama, Y., Newman, M. J., and Wilson, T. H. (1982). *J. Biol. Chem.* **257**, 5125–5128.
van Dijken, J. P., and Scheffers, W. A. (1986). *FEMS Microbiol. Rev.* **32**, 199–224.
Viitanen, P., Newman, M. J., Foster, D. L., Wilson, T. H., and Kaback, H. R. (1986). *Methods Enzymol.* **125**, 429–452.
Vilaro, S., Palacin, M., Pilch, P. F., Testar, X., and Zorzano, A. (1989). *Nature (London)* **342**, 798–800.
von Heijne, G. (1987). *Sequence Analysis in Molecular Biology—Treasure Trove or Trivial Pursuit*. Academic Press, London.
Von Heijne, G. (1988). *Biochim. Biophys. Acta* **947**, 307–333.
Walmsley, A. R. (1988). *Trends Biochem. Sci.* **13**, 226–231.
Waters, S. H., Rogowsky, J., Grinstead, J., Altenbuchner, J., and Schmitt, R. (1983). *Nucleic Acids Res.* **11**, 6089–6140.
West, I. C., (1970). *Biochem. Biophys. Res. Commun.* **41**, 655–661.
West, I. C., and Mitchell, P. (1972). *J. Bioenerg.* **3**, 445–462.
West, I. C., and Mitchell, P. (1973). *Biochem. J.* **132**, 587–592.
Wheeler, T. J., and Hinkle, P. C. (1985). *Annu. Rev. Physiol.* **47**, 503–517.
Wilson, D. M., and Wilson, T. H. (1987). *Biochim. Biophys. Acta* **904**, 191–200.
Wilson, D. M., Tsuchiya, T., and Wilson, T. H. (1986). *Methods Enzymol.* **125**, 377–387.
Wright, J. K., Teather, R. M., and Overath, P. (1983). *Methods Enzymol.* **97**, 158–175.
Yao, B., Sollitti, P., and Marmur, J. (1989). *Gene* **79**, 189–197.
Yazyu, H., Shiota-Niija, S., Shimamoto, T., Kanazawa, H., Futai, M., and Tsuchiya, T. (1984). *J. Biol. Chem.* **259**, 4320–4326.
Yazyu, H., Shiota-Niija, S., Futai, M., and Tsuchiya, T. (1985). *J. Bacteriol.* **162**, 933–937.
Zhang, C.-C., Durand, M.-C., Jeanjean, R., and Joset, F. (1989). *Mol. Microbiol.* **3**, 1221–1229.
Zilberstein, D., and Dwyer, D. M. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 1716–1720.
Zilberstein, D., Dwyer, D. M., Matthei, S., and Horuk, R. (1986). *J. Biol. Chem.* **261**, 15053–15057.