A structural genomics approach to membrane transport proteins

Peter J.F. Henderson

Astbury Centre for Structural Molecular Biology, Institute for Membranes and Systems Biology, University of Leeds

1   Introduction

The hydrophobic bilayer membrane that bounds cells is inherently impermeable to the great majority of hydrophilic solutes required for cell nutrition and to many of the waste products and/or toxins that must be excreted. Accordingly, the membrane contains proteins, the sole function of which is to catalyse the translocation of substrates through the membrane. In all types of cells, from microbes to man, there are a large number of such membrane transport proteins comprising 5-15% of the genomic protein portfolio. Generally, each one of these is highly specific for a single substrate.

As the substrates for many membrane processes can be obtained in radioisotope-labelled form, it has been technically feasible to characterise the functions of many of these transport proteins. The 3d structures of the proteins themselves, however, have proved to be difficult to elucidate: they are of low natural abundance in the membrane and so difficult to identify; they are very hydrophobic and refractory to isolation methods in aqueous solutions; and, even when purified, usually in non-denaturing detergents, they are very difficult to crystallize. Consequently, X-ray crystallography, the method of choice for determining structures of proteins, is difficult to implement. So, the numbers of independent structures of membrane proteins is numbered in hundreds (193 on 11 June, 2009 (White, 2009)), whereas the numbers of structures of soluble proteins exceeds ten thousand.
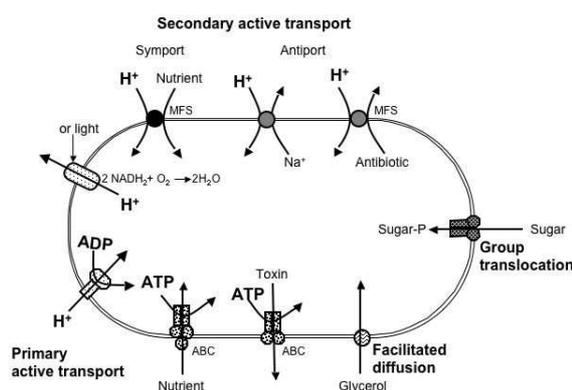

Figure 1: Membrane transport systems in a typical bacterium (see text)

In order to understand the molecular mechanisms of transport proteins, knowledge of their 3d structures is crucial. In this review the key features of transport processes are presented followed by a generic strategy - structural genomics - for elucidation of their structures.

2 Membrane transport processes and bioenergetics - useful concepts

2.1. Passive diffusion is the translocation of a solute across a membrane down its electrochemical gradient without participation of a transport protein. The process follows Fick's law, with the relationship below where velocity has a linear relationship to [solute]

$$v = PAc$$

where v =velocity, P = permeability coefficient for the particular solute, A = area, and c = difference in solute concentration across the cell membrane.

Diffusion has a low temperature coefficient ($v \; \alpha \; {}^o A$) and is non-specific. Typical biologically important compounds following this mechanism are O_2, CO_2, NH_3, HCO_2H, CH_3CO_2H, $CH_2OH.CHOH.CH_2OH$ - small, neutral molecules soluble in lipid membranes.

2.2 Facilitated diffusion is the translocation across a membrane of a solute down its electrochemical gradient catalysed by a transport protein. The Michaelis-Menten relationship often adequately relates the initial rate of transport (v) to initial substrate concentration

$$v = V_{\max}.[S]/(K_m + [S])$$

($V_{\max}$ = maximum velocity, $K_m = [S]$ where v is $V_{\max}/2$). As with enzyme reactions, there is a high temperature coefficient and, usually, strong substrate specificity. Biological substrates that follow this mechanism are typically charged and/or larger than about the size of glycerol, with a very low inherent solubility in biological membranes. Mitchell (1990) classified such transport of a single substrate as 'uniport', and glycerol transport is an example of such facilitated diffusion in *Escherichia coli* (Fig. 1).

2.3 Active transport is a term used to describe the net transport of a solute across a biological membrane from a low to a high electrochemical potential. Active transport shows the following characteristics.

1. Accumulation of solute occurs against a concentration gradient.
2. The solute is not chemically modified during translocation.
3. Saturable steady state kinetics are observed.
4. There is a high temperature coefficient typical of enzyme-catalysed reactions.
5. Substrate specificity is restricted.
6. An input of metabolic energy is required.

Active transport processes embrace a variety of molecular mechanisms, in which energy may be derived from light, oxidoreduction, ATP hydrolysis, or pre-existing solute gradients. It is conceptually helpful to classify them further into "primary" and "secondary" mechanisms. Secondary transport can be subdivided into "symport" or "antiport" (Fig. 1), terms introduced by Mitchell, (1990).

2.4 Primary active transport involves the direct conversion of chemical or photosynthetic energy into an electrochemical potential of solute across the membrane barrier. Thus, translocation of protons driven by oxidation of respiratory substrates (Mitchell, 1990), (Fig. 1), by hydrolysis of ATP, or by light, all fall into this category. Most of these transport one substrate in one direction and so are described as 'uniport' (Mitchell, 1990).

2.5 Secondary active transport involves the conversion of a pre-existing electrochemical gradient, usually of H^+ or Na^+ ions, into a new electrochemical gradient of the transported species. Thus the ultimate energy source for secondary transport systems is a primary chemical or photochemical conversion. In bacteria primary proton ejection by respiration or ATPase powers secondary sugar-H^+ symport (obligatory coupling of H^+ and solute movement in the *same* direction; Fig. 1) or secondary Na^+/H^+ antiport (the obligatory coupling of H^+ and solute movement in the *opposite direction*; Fig. 1). For example, the resulting Na^+-gradient can be further coupled to substrate uptake by a nutrient-Na^+ symport, so that net accumulation is driven by respiration (or ATPase) *via* H^+ and Na^+ gradients.

In *E. coli* the transmembrane H^+-gradient appears to be the 'currency' of many energised transport reactions, and the Na^+-gradient of relatively few. However, in other organisms living in salt environments the Na^+ gradient is the dominant factor maintained by a primary Na^+ pump. This is also true in multicellular eukaryotes.

2.6 Group Translocation All the above mechanisms operate without chemical modification of the solute. Group translocation systems catalyse both the translocation and concomitant chemical modification of the solute. For a range of carbohydrates in many species of bacteria, phosphoenol pyruvate is the donor to produce internal sugar-phosphate from external free sugar (Fig.1).

3 The post-genome age - bioinformatics and statistical analyses of amino acid alignments and emergence of evolutionary families

With the discovery of increasing numbers of transport proteins, it became important to classify the various ways in which transport occurs across biological membranes, particularly in relation to energisation of transport. Despite the enormous number of genetically and biochemically distinct systems, the number of commontypes of energisation could be reduced to five (Fig. 1). Considerable clarification of the classification was achieved when the sequences of the proteins involved became available through the advent of recombinant DNA technology. This yielded statistically robust comparisons of their (dis)similarities and their evolutionary relationships and revealed at least two important insights. Transport proteins related by sequence would usually have the same *modus operandi*, but it might be profoundly different. Thus, sequence-related proteins could operate by facilitated diffusion, substrate-cation symport, or substrate-cation antiport, and therefore it was realised that quite subtle changes in structure might produce profound changes in direction and/or energy linkage of the transport process. Secondly, sequence analyses also confirmed that transport systems could be comprised of more than one protein, each with a different functional role in the overall translocation process, and also that domains of different function could become fused together during evolution so that fewer, or even one, polypeptide chain contained all the necessary functions. One example of this is the protein implicated in cystic fibrosis.

The classification of membrane transport proteins according to the statistical relatedness of their aligned amino acid sequences has been exhaustively explored by Saier (2000), leading to the Transport Classification Database (TCDB). Each time the genomic DNA of an organism is sequenced, the amino acid sequences of all the proteins are deduced and their possible functions annotated. There are sufficient precedents where function has been identified in one organism

for educated guesses to be made for functions of proteins found to be present in all subsequent genomes sequenced. A repository of the conclusions for transport proteins is maintained by Paulsen and co-workers (Ren and Paulsen, 2007).

A fundamental conclusion from these global surveys is that transport proteins tend to fall into two major classes; each of these is defined by a statistically significant thread of similarities throughout evolution from the primordial microorganisms through to higher plants and animals, including man. One is a series of primary active transport systems, known as the 'ATP-binding-cassette', ABC-transporters superfamily (Higgins, 1992), and the other as the 'Major Facilitator Superfamily', MFS-transporters (Pao et al., 1998). For example, in *E. coli* out of 354 transport systems in total 69 are of the ABC family and 70 are of the MFS family; in man out of a total of 1022 53 are ABC and 104 are MFS (Ren and Paulsen, 2007).

4   Structural genomics

In this presentation, an experimental strategy is outlined that enables the amplified expression and purification of bacterial membrane transport proteins, from several species of bacteria, in amounts required for structural studies. Many of these prokaryote transport proteins are homologous to members of the 'Major Facilitator Superfamily' found in a range of higher organisms, such as protozoan parasites, fungi, plants and mammals. It is clear, therefore, that results of structure-activity studies of these bacterial proteins will be relevant to the understanding of the corresponding eukaryotic transporters. In other cases the transport systems are unique to bacteria and, indeed, sometimes critical for growth in cases where these bacteria are pathogenic. The availability of purified active protein, of these key transporters, may then be useful for discovery of novel antibacterials. The transporters chosen to illustrate the strategy, which is illustrated in Fig. 2, belong to the nucleobase-cation-symport, 'NCS-1', family, and the structure of one, the Na^+-hydantoin transport protein Mhp1 from *Microbacterium liquefaciens*, has been determined (Weyand et al., 2008); this has led to the emergence of a novel 'Superfamily' (Weyand et al., 2008).

5   How do transport proteins work? Structures and surprises

Identification of possible transport proteins through statistical analyses of their amino acid sequences is now trivial, but their complete biochemical, functional and structural characterisation in purified systems is only just beginning. Thus the focus of the field is on structural genomics – devising a general strategy to move quickly from identification of a gene of interest to having sufficient quantities of pure protein for functional and structural studies (Szakonyi et al., 2007). The structures of some transport proteins will be illustrated, and a molecular mechanism proposed for Mhp1.

References

Higgins C.F. (1992). ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8**: 67–113.

Mitchell, P. (1990). Osmochemistry of solute translocation. *Res. Microbiol.***141**: 286–289.
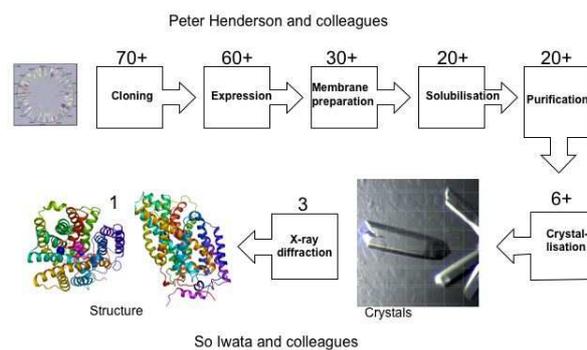
Figure 2: The structural genomics pipeline for bacterial membrane transport proteins

Pao S.S., Paulsen I.T., and Saier M.H. Jr. (1998). Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**: 1–34.

Ren Q. and Paulsen I.T. (2007). Large-scale comparative genomic analyses of cytoplasmic membrane transport systems in prokaryotes. *J Mol Microbiol Biotech* **12**: 165–79.

Saier M.H, Jr. (2000). A functional-phylogenetic classification system for transmembrane solut transporters. *Microbiol Mol Biol Rev* **64**: 354–411.

Szakonyi G., Dong, L., Ma, P., Bettaney K.E., Saidijam M., Ward, A., Zibaei, S., Gardiner, A.T., Cogdell, R.J., Butaye, P., Kolsto, A-B., O'Reilly J., Hope, R.J., Rutherford N.G., Hoyle C.J. and Henderson P.J.F. (2007). A genomic strategy for cloning, expressing and purifying efflux proteins of the major facilitator superfamily. *J. Antimicrob. Chemother*. **59**, 1265–1270.

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J., O'Reilly, J., Ma, P., Saidijam, M., Patching, S.G., Hope, R.J., Norbertczak, H.T., Roach, P.C.J., Iwata, S., Henderson, P.J.F. and Cameron, A.D. (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322**, 709–713.

White, S. (2009). http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html.