Abstract We present an overview of the architecture of ATP-binding cassette (ABC) transporters and dissect the systems in core and accessory domains. The ABC transporter core is formed by the transmembrane domains (TMDs) and the nucleotide binding domains (NBDs) that constitute the actual translocator. The accessory domains include the substrate-binding proteins, that function as high affinity receptors in ABC type uptake systems, and regulatory or catalytic domains that can be fused to either the TMDs or NBDs. The regulatory domains add unique functions to the transporters allowing the systems to act as channel conductance regulators, osmosensors/regulators, and assemble into macromolecular complexes with specific properties.

© 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: ABC transporter; Accessory domain; Architecture; ATP-binding cassette; Regulation of transport

1. Introduction

ATP-binding cassette (ABC) transporters are vital to any living system and are involved in the translocation of a wide variety of substrates ranging from ions, sugars, amino acids, vitamins, lipids, antibiotics and drugs to larger molecules such as oligosaccharides, oligopeptides and even high molecular weight proteins [1]. Owing to their function ABC transporters are involved in many crucial processes, including nutrient uptake, lipid trafficking, drug and antibiotic excretion, secretion of macromolecules, antigen presentation to cytotoxic T cells, and cell volume regulation, and they are often major players in complex pathways affecting gene expression (e.g., sporulation, competence and virulence development). In humans a still growing number of diseases is related to (dys)function of ABC transporters, such as cancer (e.g., breast and ovarian cancer), in which they are involved in nutrient uptake and cell volume regulation, and they are often major players in complex pathways affecting gene expression. In many cases, two different

2. Architecture of ABC uptake systems

The first identified components of ABC transporters were the substrate-binding proteins (SBPs) of ABC uptake systems that reside in the periplasmic space of Gram-negative bacteria, from which they could be released by cold osmotic shock [2–5]. These water soluble proteins bind the substrate in a cleft between two lobes and deliver it to the translocator in the cytoplasmic membrane. In Gram-positive bacteria and archaea, i.e., microorganisms lacking an outer membrane, SBPs are exposed on the cell surface and attached to the cytoplasmic membrane via a lipid-anchor or a transmembrane-peptide (to date, the latter has been only observed in archaea), or they can be fused to the TMDs resulting in two substrate-binding domains (SBDs) per functional complex (Fig. 1B; OpuA). In some cases, two SBDs fused in tandem are linked to the TMDs and these systems have a total of four extracytoplasmic substrate-binding sites (e.g., GlnPQ). Systems with SBDs fused to the TMDs can be also found in Gram-negative bacteria but less frequently than in Gram-positives [6].

In terms of number of different subunits, the most complex ABC type uptake systems known are the nickel- and (oligo)peptide transporters, in which each domain is present as a separate polypeptide, giving a total of (at least) five unique proteins required for transport (Fig. 1A and B; Nik and Opp). In many cases, two copies of the same NBD plus two different TMDs (e.g., the maltose, histidine, and glucose transporters) or, less frequently, two copies of one NBD plus one TMD are present (the vitamin B12 transporter Btu). Occasionally, the...
NBDs or TMDs are fused (e.g., siderophore/haem and ribose transporters) such that in total only three unique proteins are used (Fig. 1A). The smallest number of unique subunits of which ABC uptake systems are composed, known to date, is two (Fig. 1B). This is the case for the glycine betaine and glutamine/glutamic acid transporters in which the SBPs are fused to either the N- or C-terminus of their cognate TMDs, and the second unique subunit is formed by the NBD [6].

3. Architecture of ABC efflux systems

In contrast to ABC uptake systems, which are present in the cytoplasmic membrane of prokaryotic species (bacteria and archaea), ABC efflux systems are found in every life-form ranging from microorganisms to man. ABC efflux pumps can be found in the cytoplasmic membrane of prokaryotes and eukaryotes, as well as in the organellar membranes of the higher organisms, that are, the endoplasmic reticulum, inner mitochondrial membrane, peroxisomal and vacuolar membranes. Although the common four-domain structure of the translocator (two TMDs plus two NBDs) is conserved, ABC efflux pumps are somewhat less complex compared to ABC uptake systems in that they do not require a SBP for function (Fig. 1C and D). The direction of transport can be deduced from the presence or not of this latter component. An exception to this rule has yet to be found. Even more frequently than in ABC uptake systems, the four core-domains of the translocator are fused, yielding only one or two unique polypeptides forming the functional unit (Fig. 1C and D). The most common organization is the “half-size transporter” in which a single TMD is fused to the N- or C-terminus of a NBD. The functional complex is either a homo- or heterodimer of two half-transporters (e.g., the lipid flippases, multi-drug transporters, cysteine exporter, and peptide transporters; Fig. 1C and D). Finally, all four core-domains (two half-transporters) can also be fused as is the case for the multi-drug transporters (e.g., P-gp, MRP1 and 2, and PDR5), and the ABC type chloride channel CFTR (Fig. 1D), and these are termed “full-length transporters.”

4. Transmembrane domains

The TMDs are two highly hydrophobic proteins or protein domains that create the channel through which the substrate passes during translocation (Fig. 1). In ABC uptake systems, they form the docking site for the SBP and transduce a signal to the NBDs to bind and hydrolyze ATP. The TMDs are bundles of α-helices that traverse the membrane multiple times in a zig-zag fashion. Many typical ABC transporters conform to the paradigmatic “two-times-six” α-helices structure, yielding
a total of 12 transmembrane segments (TMSs) per functional unit (Table 1). However, the number of TMSs may vary from anywhere between 5 and 11 for each individual TMD.

Whereas the NBDs show considerable sequence identity (see below), the TMDs (as well as the SBPs) have been less well conserved during evolution, reflecting the large diversity of substrates transported. TMDs of ABC uptake systems contain the so-called EAA motif (L-loop), which usually is found in the cytoplasmic loop between TMS 5 and 6 [7]. This motif is absent in ABC exporters.

In ABC efflux systems, the TMDs harbor the substrate-binding site(s). Evidence for the existence of a substrate-binding site in the TMDs of ABC uptake systems came from the analysis of SBP-independent mutants [8–11]. Mutations in the NBD of the histidine transporter (HisP) or the TMDs of the maltose transporter (MalF and G) resulted in an uncoupled phenotype, in which the SBP no longer interacted with the TMDs and the NBDs hydrolyzed ATP constitutively. In addition, a SBP-independent maltose transporter with a truncated form of MalF had its substrate specificity changed from maltose to lactose, albeit with very high affinity constants [12]. Finally, the crystal structure of the BtuCD translocator revealed a cleft in the periplasmic half of the TMDs that could accommodate a vitamin B12 molecule and represent a genuine substrate-binding site [13].

### 5. Nucleotide-binding domains

The engines of ABC transporters are the NBDs as these power substrate translocation by ATP-binding and hydrolysis. Whereas the SBPs and TMDs are poorly conserved, the NBDs are more closely related and group the different transporter families in the ABC superfamily. They contain the ATP-binding cassettes, hence the name ABC transporters. Many crystal structures of isolated NBDs are available, yielding detailed information about the overall fold and the nucleotide-binding site(s) [14]. A NBD consists of two domains, a RecA-like and a

<table>
<thead>
<tr>
<th>TMD</th>
<th>Transport substrate</th>
<th>Number of TMSs</th>
<th>Total number of TMSs</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABC uptake systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HisQ</td>
<td>Histidine</td>
<td>5</td>
<td>10 (2 x 5)</td>
<td>[133]</td>
</tr>
<tr>
<td>HisM</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GlnP</td>
<td>Glutamine/glutamic acid</td>
<td>5$^a$</td>
<td>10 (2 x 5)</td>
<td>[52]</td>
</tr>
<tr>
<td>ProW</td>
<td>Gycine betaine/Proline betaine</td>
<td>2 x 5</td>
<td>14 (2 x 7)</td>
<td>[26]</td>
</tr>
<tr>
<td>OpuABC</td>
<td>Gycine betaine</td>
<td>2</td>
<td>16 (2 x 8)</td>
<td>[6]</td>
</tr>
<tr>
<td>OppB</td>
<td>Oligopeptidases</td>
<td>6</td>
<td>12 (2 x 6)</td>
<td>[134]</td>
</tr>
<tr>
<td>OppC</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MalF</td>
<td>Maltose$^c$</td>
<td>2</td>
<td>14 (8 + 6)</td>
<td>[54,135]</td>
</tr>
<tr>
<td>MalG</td>
<td>maltodextrins</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RbsC</td>
<td>Ribose$^d$</td>
<td>10$^e$</td>
<td>20 (2 x 10)</td>
<td>[136]</td>
</tr>
<tr>
<td>BtuC</td>
<td>Vitamin B12</td>
<td>10</td>
<td>20 (2 x 10)</td>
<td>[13]</td>
</tr>
<tr>
<td>FhuB</td>
<td>Siderophores/haem/vitamin B12</td>
<td>10 + 2</td>
<td>20</td>
<td>[137]</td>
</tr>
<tr>
<td><strong>ABC efflux systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Half-transporters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmrA</td>
<td>Drugs</td>
<td>6$^d$</td>
<td>12 (2 x 6)</td>
<td>[138]</td>
</tr>
<tr>
<td>BmrA</td>
<td>Drugs</td>
<td>6$^d$</td>
<td>12 (2 x 6)</td>
<td>[139]</td>
</tr>
<tr>
<td>MsbA</td>
<td>Lipids</td>
<td>6</td>
<td>12 (2 x 6)</td>
<td>[16]</td>
</tr>
<tr>
<td>CydC</td>
<td>Cysteine/GSH</td>
<td>6</td>
<td>12 (2 x 6)</td>
<td>[140]</td>
</tr>
<tr>
<td>CydD</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HlyB</td>
<td>Haemolysin$^f$</td>
<td>(2)$^f$</td>
<td>12–16$^f$</td>
<td>[141,142]</td>
</tr>
<tr>
<td>TAP1 (ABC2)</td>
<td>Antigenic</td>
<td>4</td>
<td>19 (10 + 9)</td>
<td>[59]</td>
</tr>
<tr>
<td>TAP2 (ABC3)</td>
<td>peptides</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Full-length transporters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-gp (MDR1/ABCB1)</td>
<td>Drugs</td>
<td>6 x 2</td>
<td>12</td>
<td>[143]</td>
</tr>
<tr>
<td>BCRP (ABCG2)</td>
<td>Drugs</td>
<td>6 x 2</td>
<td>12</td>
<td>[144]</td>
</tr>
<tr>
<td>MRP1 (ABCG2)</td>
<td>Drugs</td>
<td>5</td>
<td>17 (5+12)</td>
<td>[144]</td>
</tr>
<tr>
<td>MRP2 (ABCC1)</td>
<td>Drugs</td>
<td>5</td>
<td>17 (5 + 12)</td>
<td>[144]</td>
</tr>
<tr>
<td>SUR1 (ABCC8)</td>
<td>None/unknown$^i$</td>
<td>5</td>
<td>17 (5 + 12)</td>
<td>[145]</td>
</tr>
<tr>
<td>CFTR (ABCC7)$^g$</td>
<td>Chloride</td>
<td>6 x 2</td>
<td>12</td>
<td>[146]</td>
</tr>
<tr>
<td>Stec</td>
<td>Mating pheromone</td>
<td>6 x 2$^bh$</td>
<td>12</td>
<td>[147]</td>
</tr>
<tr>
<td>ABCR (ABCA4)</td>
<td>Retinal/retinal phosphatidylethanolamine</td>
<td>6 x 2$^l$</td>
<td>12</td>
<td>[24]</td>
</tr>
</tbody>
</table>

$^a$Two SBDs are fused in tandem to the first TMS. The N-terminal signal sequence, preceding the SBDs, is cleaved off.
$^b$The SBD is fused to a C-terminal signal anchor sequence (the 8th TMS).
$^c$Initially claimed to have 6 TMSs [148]; later shown to possess 10 TMSs [136].
$^d$No experimental evidence; topology based on homology with other multidrug-efflux systems.
$^e$HlyB is thought to contain 6-8 TMSs; the presence of two N-terminal TMSs has not been unambiguously determined.
$^f$Sulfonylurea receptors (SURx) associate with Kir6.6.x to form ATP-sensitive potassium channels.
$^g$CFTR is a chloride channel rather than transporter although it has the classical two TMSs plus two NBDs structure observed for genuine ABC transporters.
$^h$Experimental data only for the N-terminal TMD.
$^i$Two large extracytoplasmic glycosylated domains are inserted between the first and second TMS of both the N- and C-terminal TMDs.
helical domain, the latter being specific for ABC transporters. In the NBD monomer, the ATP-binding site is formed by two conserved sequences, the Walker A (P-loop) and Walker B motifs, present in many ATP-binding proteins and not restricted to ABC proteins [15]. These motifs bind the phosphates of ATP and ADP, and coordinate a Mg$^{2+}$-ion via H$_2$O in the nucleotide-binding site, respectively. A glutamate directly following the Walker B motif might be the catalytic base for hydrolysis, since it binds to the attacking water and Mg$^{2+}$-ion, and mutation of this residue leads to loss of ATPase activity [14]. The Q-loop also forms hydrogen bonds with Mg$^{2+}$ and the attacking water. A conserved histidine in the H motif hydrogen bonds with the γ-phosphate of ATP. In the NBD dimer, the signature sequence, LSGGQ, (the motif that distinguishes ABC proteins from other nucleotide-binding proteins) of one NBD binds to the γ-phosphate of ATP bound to the other NBD, thereby completing the nucleotide-binding sites [14]. Thus, two ATPs are sandwiched in head-to-tail orientation between two NBDs in the catalytically active dimer.

6. ABC transporter mechanism

Given the conserved architecture of two TMDs plus two NBDs across the entire ABC transporter superfamily, one could expect a common translocation mechanism for ABC uptake and efflux systems. Indeed, from the culmination of vast amounts of functional data, in combination with the recent X-ray structures of complete ABC transporters (for both efflux and uptake systems; [13,14,16–21]), a consensus mechanism is emerging. The simplest model of SBP-dependent uptake encompasses four steps (Fig. 2): (i) substrate binds to the SBP; (ii) the closed, liganded SBP docks onto the TMDs and signals the NBDs to cooperatively bind ATP; (iii) upon ATP-binding, a binding site in the TMDs is made available to the outside, the SBP opens and substrate is transferred to the binding site in the TMDs; and (iv) ATP hydrolysis dissociates the NBD dimer and results in reorientation of the binding site from an outward- to an inward-facing conformation. The substrate arrives in the cytoplasm and the SBP is released from the TMDs. Alternatively, it cannot be ruled out that upon ATP binding the substrate is transferred directly to the cytoplasm (step iii) and ATP hydrolysis in combination with dissociation of ADP and inorganic phosphate is solely used to reset the system for another translocation cycle. Whereas step (i) is unique to SBP-dependent uptake systems, steps (ii–iv) could be envisaged for ABC efflux systems as well. In step (ii), a substrate bound to the inward-facing binding site would signal the NBDs to bind ATP. Subsequently, step (iii), ATP-binding reorients the binding site and substrate is released to the outside. (iv) ATP-hydrolysis resets the system by converting the TMDs back to the inward-facing conformation.

Although ABC uptake and efflux systems could share a common translocation mechanism, there are some indications that things might be more complicated. First of all, although in nature TMDs and NBDs appear as separate modules or fused together in almost any possible combination (Fig. 1), to the best of our knowledge, no ABC uptake systems are known that have a TMD fused to a NBD as is the case with half- or full-length efflux systems. For ABC efflux systems, this latter architecture seems to be a universal feature. Second, the conserved EAA motif (see above) is found exclusively in ABC uptake systems. This motif is thought to be important for the interaction of the TMDs with the NBDs. Third, the structures of the ABC efflux system MsbA showed an additional intracellular domain between the TMD and NBD that was not present in the structure of the SBP-dependent vitamin B12 uptake system BtuCD [13,16–18]. Fourth, NBD2 of CFTR was shown, in addition to hydrolyzing ATP (Mg-ATP + H$_2$O → Mg-ADP + Pi), to display adenylate kinase activity (Mg-ATP + AMP ↔ Mg-ADP + ADP), which influenced channel activity [22]. It remains to be established whether the adenylate kinase activity is unique to CFTR or ABC efflux systems, but some important structural differences between ABC uptake and efflux systems seem to exist. The significance of these observations in terms of a general ABC transport mechanism awaits further experimentation.

7. Accessory domains in ABC transporters

In addition to the ubiquitous four core-domains, that are two TMDs plus two NBDs, accessory domains can be found in several ABC transporters (Fig. 3). These extra domains often perform different functions and can be found separate from, attached to, or integrated into different parts of the core ABC transporter. On the basis of their location (extrinsic or intrinsic to the membrane) and function (regulatory or catalytic) within the cognate transporters, we divide the well-characterized accessory domains of ABC transporters into four different groups: (i) extracytoplasmic domains; (ii) membrane-embedded domains; (iii) cytosolic regulatory domains and (iv) cytosolic catalytic domains.

7.1. Extracytoplasmic domains

One of the characteristics of accessory domains is that they are not part of the general ABC transporter core complex. There are three main examples of known extracytoplasmic
domains or proteins that can be considered accessory units (i) the large extracytoplasmic domains (ECDs) found in members of the ABCA family; (ii) the extracellular N-terminal domain of ProW (and closely related homologues), the membrane component of the ProU system from Escherichia coli; and (iii) the SBDs or proteins found in prokaryotes.
Members of the ABCA family are full-length transporters consisting of two TMDs of 6 TMSs each and two NBDs, each located C-terminal of one of the TMDs (Table 1; [23]). The two ECDs are predicted after the first TMS of the N- and C-terminal half of the transporter (Fig. 3C) and are 600 and 275 amino acids long in ABCA4 (ABCR) [24]. Eight N-glycosylation sites were mapped by mutagenesis onto the ECDs of ABCA4, four in the N-terminal half and four in the C-terminal half of the protein [24]. Removal of these N-glycosylation sites resulted in reduced expression levels, although the mutants retained the capacity to bind ATP. It has been postulated that the two ECDs are linked by at least one disulfide bridge. Intramolecular crosslinking of cysteine residues in an ECD was shown for ABCG2 (BCRP) [25], a half-transporter with the ECD located between TMS 5 and TMS 6. Two ABCG2 subunits were linked by a disulfide bridge, and two remaining cysteine residues formed intramolecular disulfide bonds within the ECD. Although mutations in the ECDs have been shown to influence the substrate specificity of the transporters, a distinct function is not known for these relatively large ECDs.

The second example of an ECD is found in the proline and glycine betaine transport system (ProU) of E. coli. Here the membrane component, ProW, is composed of an extended N-terminal region of about 100 amino acid residues followed by 7 TMS (Table 1). The function of the N-terminal region is unknown, but a phoA and LacZ fusion study showed that its location is periplasmic [26].

The best-studied ECDs or proteins are the SBDs or proteins of ABC-type uptake systems. These domains or proteins are considered to be accessory, because these components are not intrinsic to the translocation process as has been demonstrated for ABC transporter mutants that function in the absence of SBPs (see Section 4). SBPs are the main determinants of substrate specificity of ABC uptake systems and confer high-affinity on the transport process with dissociation constants ($K_d$) in the sub(micromolar range. They have been a gold-mine for structural studies due to their well behavior during overexpression, purification and crystallization trials. Although some early experiments suggested that SBPs might self-associate to form dimers or higher order oligomers [27–30], crystal structures of numerous SBPs have led to the generally accepted view that they are monomers with one substrate-binding site per molecule [31]. Even though systems are known that use multiple SBPs to broaden the substrate specificity of the transporter (several peptide transporters; [32–34]), only a single SBP is required for transport function. The high resolution structure of the vitamin B12 importer BtuC determined that there is space for only a single SBP to dock onto the TMDs [35]. The ABC transporters with one or two SBP have been fused to the TMD are thought to function similarly, except for that the presence of a second, third or fourth SBP might affect the kinetics of the translocation process [36] or perhaps broaden the substrate specificity.

The archetype SBP consists of two domains or lobes connected by a flexible linker. The two lobes close and engulf the ligand upon substrate-binding (Venus’s Flytrap model; [37]). Structures are available of SBPs in open-unliganded, closed-unliganded, open-liganded, and closed-liganded forms [38–42], of which the latter conformation is thought to productively interact with the membrane complex [43–45]. In some cases, unliganded SBPs also have access to the TMDs, thereby inhibiting ABC transporter function (e.g., in case of the maltose- and histidine-binding proteins; [46,47]). These observations, however, could not be confirmed in similar experiments performed with the oligopeptide ABC transporter [45]. Mutational and suppressor analyses suggested that each lobe of the SBP binds to one of the TMDs [8,48]. Site-directed mutagenesis and pre-steady-state kinetic analysis indicated that donation of the substrate by the SBP to the TMDs is the rate-limiting step during the ABC uptake catalytic cycle of the oligopeptide transporter Opp [49,50].

Although a typical SBP consists of two lobes (domains) that are connected by a flexible hinge [31], exceptions to this rule are the SBPs of the nickel/oligopeptide transporter family (Fig. 3A). OppA contains an extra, third domain [38]. The function of this additional domain is unknown, although in AppA, an OppA homologue from Bacillus subtilis, it has been shown to form part of the binding site for nonameric peptides [51].

Two ABC transporter families (OTCN and PAO) are known in which one or two SBPs are fused to either the N- or C-terminus of one TMD, yielding a total of two or four SBDs per functional complex (Fig. 1B, OpuA and GlnPQ; [6]). Two SBDs are present in OpuA, the osmoregulatory ABC transporter from Lactococcus lactis, although for activity only one SBD is needed. The SBDs of OpuA interact in a cooperative manner by stimulating either the docking of one domain onto the TMDs or substrate donation to the TMDs (Fig. 2, steps (ii) and (iii), respectively; [36]). However, this cooperative interaction has not been established for other SBP dependent ABC transporters, e.g., the glutamate/glutamine transporter GlnPQ from L. lactis, containing four SBDs [52]. In GlnPQ, the SBD in the primary sequence nearest to the translocator domain (SBD2) is used as high affinity receptor that delivers the substrate to the TMDs. SBD1, on the contrary, is not required for high affinity transport and is not sufficient for transport but does bind glutamine with a low affinity (dissociation constant in the submillimolar instead of micromolar range; Schuurman–Wolters, unpublished).

7.2. Membrane-embedded domains

The majority of ABC transporters are predicted to have six TMSs per TMD core (Table 1). Thus, in a typical transporter a total of 12 TMSs are present. Exceptions are known for ABC uptake systems that contain 5 or 10 TMSs as a single TMD core that is both essential and sufficient for transport. In addition to this somewhat variable number of TMSs in the TMDs, several ABC transporters contain extra TMSs that are not part of the core TMD and can be considered as accessory domains. Although this phenomenon frequently occurs, the function of these extra domains is in many cases unknown. For example, within the OTCN family of ABC transporters [53], sequence alignments reveal a core TMD of 5 TMSs but homologues can have in total either 5, 6 or 8 TMSs. The core TMSs are generally much better conserved than the accessory TMSs ([6]; unpublished). Within the OTCN family, some polypeptides (e.g., the OpuABC subunit of OpuA; Fig. 3B, [6]; and the ProW subunit of ProU, [26]) have two extra TMSs N-terminal of the TMD core but their function is unknown. In addition, some members have an 8th TMS (or 6th TMS in case the first two are missing), C-terminal of the TMD core, that serves as signal anchor sequence for the translocation of the extracytoplasmic SBD. Also, members of the OSP family to which the maltose transporter belongs can have two extra N-terminal
TMSs (MalF; [54]). For the maltose transporter, the extra TMSs of MalF have been proposed to be an export signal, but most of the two TMSs could be removed without any effect on maltose transport activity [54–56].

A system in which a clear function could be assigned to the accessory TMSs is the transporter associated with antigen processing (TAP), present in the endoplasmic reticulum membrane. Based on hydrophobicity analysis and sequence alignments 10 (4 accessory plus 6 core) and 9 (3 accessory plus 6 core) TMSs were predicted for human TAP1 and TAP2, respectively (Table 1 and Fig. 3D; [57,58]). Although the core-TAP complex (the C-terminal 6 TMSs from TAP1 and TAP2 combined) represented the minimal functional unit for transport, the N-terminal domains of TAP1 and TAP2 were essential for recruitment of tapasin [58,59], a glycoprotein known to mediate association between TAP and MHC class I molecules [60]. Consequently, the extra domains were proposed to mediate the assembly of the macromolecular peptide-loading complex. Subsequently, it was shown that in TAP2 the tapasin-binding site is exclusively located in the N-terminal domain. Upon deletion of the N-terminal domain of TAP1, tapasin could still interact with the TAP1 TMD core [61]. TAP variants lacking the N-terminal domain of TAP2, but not of TAP1, were disturbed in the quality control of MHC I loading, and formed peptide loading complexes that were disturbed in their physical interaction with calreticulin, calnexin, and ER60.

Instead of a few extra TMSs, some ABC systems, including the sulfonylurea receptors SUR1 and SUR2, and the multidrug-resistance proteins MRP1, 2, 3, 6 and 7, have 5 accessory TMSs that may form a domain by itself (Table 1 and Fig. 3E; [62–64]). These transporters belong to the ABC transporter superfamily. The core complex in this subfamily is a full-length transporter composed of two TMDs (TMD1 and TMD2, each containing 6 TMSs) and two NBDs (NBD1 and NBD2) that follow TMD1 and TMD2, respectively. The extra TMSs, indicated as TMD0, are linked via a cytoplasmic loop (L0) to the N-terminus of TMD1 (Fig. 3E). In MRP1, the entire TMD0 could be deleted without having any effect on transport activity or on the trafficking and insertion of the protein into the basolateral membrane [65,66]. This contrasted results obtained with SUR1 and MRP2, where deletion of TMD0 affected channel activity, processing and trafficking [67–71]. Recent studies revealed that TMD0 is needed for trafficking if the C-terminal region of MRP1 is mutated or deleted [72]. It seems that in MRP1 TMD0 and the C-terminal region contain redundant trafficking signals that become essential when one or the other region is missing or mutated.

In SUR, TMD0 is required for activation of the inwardly rectifying potassium channel (KIR) subunits and anchoring of the KIR subunits to the ABC transporter core of SUR, thereby generating functional KIR channels [67,68]. Thus, similar to the TAP system, the accessory membrane-embedded domains of SUR provide additional protein interaction sites and serve a regulatory purpose.

### 7.3. Cytosolic regulatory domains

Regulation of many ABC transporters takes place at the level of gene expression. Most ABC transporter operons encode a transcriptional regulator that connects the level of protein to environmental conditions. Expression of the osmoregulatory ABC transporter OpuA from L. lactis is, for example, regulated in response to osmotic stress via the transcription factor OpuR [73]. However, the activity of ABC transporters, once expressed, may also be regulated. For this purpose, additional (autoregulatory) domains in the transporter itself are required. Examples are the L0 linker in MRP1 and SUR, the C-terminal domain of MalK of the maltose transporter from E. coli and S. typhimurium, the tandem cystathionine-β-synthase (CBS) domains of OpuA from L. lactis and other members of the OTCN family, and the R-domain of CFTR (Fig. 3E, F, G and H, respectively). Although these domains all regulate ABC transporter activity, they use different mechanisms to do so.

The maltose/maltodextrin transporter from E. coli and S. typhimurium is directly involved in transcriptional regulation of, amongst others, itself. The regulation takes place at the C-terminal domain of MalK, the NBD of the transporter (Fig. 1A (Mal) and 3F; [74]). Crystal structures are available of MalK [75], and GlcV, the NBD of the glucose transporter from Sulfolobus solfataricus [76]. These NBDs contain an accessory C-terminal domain with a common tertiary fold. The C-terminal domain of MalK is involved in two regulatory processes, at least in enteric bacteria [77]. First, via its C-terminal domain, MalK is subject to regulation by IIA^abc, a component of the bacterial phosphoenolpyruvate-dependent sugar:phosphotransferase system (PTS) [75,78,79]. IIA^abc can be in the unphosphorylated and phosphorylated state, depending on the availability of PTS sugars. In the unphosphorylated form, i.e., when a PTS sugar such as glucose is present, IIA^abc binds to MalK and thereby prevents maltose uptake. The inhibition is relieved when glucose is exhausted and IIA^abc becomes phosphorylated. This regulatory process determines the hierarchy of sugar utilization and is known as inducer exclusion mechanism. GlcV is unlikely to be involved in this type of regulation, since archaea lack a PTS. Second, MalK inhibits the activity of MalT, the positive regulator of the maltose regulon, by competing with MalT for maltotriose binding [75,80–82]. Binding of maltotriose to MalT is required for oligomerization of the transcription factor, a prerequisite for the interaction of MalT with its DNA binding sites [83,84]. In addition to MalK, two other proteins, MalY and Aes, can also bind MalT [85,86], creating an intricate regulatory network that will not be detailed further in this paper. MalK is thought to possess two distinct protein binding sites, one for MalT and one for IIA^abc, because mutations in MalK affecting inducer exclusion do not interfere with MalT inactivation [78,87]. By mapping known mutations onto the structure of MalK (the one from Thermococcus litoralis), the MalT interaction site has been proposed to be located on an exposed surface of the C-terminal regulatory domain. Point mutations that confer an inducer exclusion insensitive phenotype form a patch adjacent to and oriented perpendicularly to the MalT interaction site [75].

The osmoregulatory ABC transporter OpuA protects the cell against hyperosmotic stress, ultimately plasmolysis, through uptake of the compatible solute glycine betaine, and thereby reverses the osmotic shrinkage of the cell [88]. When reconstituted in proteoliposomes, OpuA is activated by increasing internal ionic strength [89]. The threshold for activation by ions is dependent on the ionic lipid content (charge density) of the membrane, indicating that the ionic signal is transduced to the transporter via critical interactions of protein domains with membrane lipids [90]. The ABC transporter
OpfU, as do other members of the OTCN family, has two CBS domains (Fig. 3G) in tandem at the C-terminal end of the ATPase subunit OpfUA. With two ATPase subunits per translocation unit, a total of four CBS domains are present in OpfU. This so-called CBS module of OpfU constitutes the sensor that switches the transporter between an inactive, electrostatically locked and an active, thermally relaxed state, by interacting with the negatively charged membrane surface in response to the ionic strength [89]. The switching mechanism is an effective means for cells to respond to osmotic shifts, as an increase in medium osmolality will result in a decrease in cell volume, and the accompanying increase in cytoplasmic ionic strength will activate the transporter. Glycine betaine accumulation, followed by water influx, will increase the cell volume and decrease the ionic strength in due time, and the transporter will ultimately be switched off. This inherent feedback mechanism prevents over-accumulation of glycine betaine and couples transporter activity with ionic strength in the cell.

CFTR is an ABC-type chloride channel that is regulated by protein kinase A-mediated phosphorylation of its cytoplasmic regulatory (R) domain [91]. The R-domain is located between NBD1 and TMD2 of the full-length transporter (Fig. 3G), and contains a number of charged residues and multiple consensus phosphorylation sites [92,93]. The unphosphorylated R-domain is thought to inhibit channel activity by interaction with the CFTR core transporter. Release of the R-domain by phosphorylation would activate the channel [94–96]. Thus far, phosphorylation dependent release of the R-domain has not been demonstrated unequivocally. Correspondingly, a recent study strongly suggested that phosphorylation results in enhanced association of the R-domain with the transporter rather than in its release [97]. The mechanism of regulation suggested by Chappe et al. was that binding of the R-domain to the transporter induces intramolecular rearrangements that allow nucleotide-induced conformational changes to be transmitted from the NBDs to the TMDs. A second system that was shown to be regulated by phosphorylation of a cytoplasmic domain is ABCA1 [98]. This half transporter is a member of the ABCA family in which the R-domains (R1 and R2) are located C-terminal of the NBDs of the C- and N-terminal half of the transporter, respectively (Fig. 3C). Phosphorylation of the R1 domain of ABCA1 diminished the phospholipid flip-flop, apolipoprotein binding, and cholesterol/lipid extrusion activities.

The N-terminal tail of CFTR (N-tail; Fig. 3H), next to its R-domain, also plays a role in regulating CFTR channel activity. It modulates the rate of channel openings, probably by interacting with components of CFTR (R-domain or NBD), or with the pore itself [99]. Next to a suggested physical interaction with a portion of the R-domain [100], the N-tail and the C-terminal tails of CFTR have been shown to bind several inhibitors and stimulators (physical modulators) of CFTR activity [101]. These interaction partners either bind directly to CFTR or mediated through various PDZ domain-containing proteins. One of the PDZ containing proteins is the Na^+H^- exchanger regulatory factor (NHERF) that can form homo- and hetero-oligomers between CFTR and other membrane proteins, e.g., the β2-adrenergic receptor [102,103]. It has been recently shown that the assembly of the C-terminal domains of CFTR is regulated by ezrin, an actin binding protein [104]. The dimeric CFTR-NHERF-ezrin interaction would provide a pathway to regulate CFTR activity via the cytoskeleton actin network. Moreover, there is evidence that when two subunits of CFTR are brought in close proximity by NHERF, the open probability of the channels is increased [105]. Finally, the increasing number of proteins interacting with CFTR suggests that the chloride conductor is not only regulated itself, but that CFTR might also regulate the activities of other ion channels, receptors, or transporters.

The N-tail of CFTR contains a cluster of acidic residues which are located onto one face of a predicted α-helix [100,106]. A similar amphipatic α-helix has been predicted for the L0 linker in the sulfonlurea receptor, SUR [107]. This linker is located between TMD0 and TMD1 (Fig. 3E), and consequently links TMD1 and TMD2 of SUR to the TMD0/KIR complex. Interaction of the N-terminus of KIR with L0 has a bidirectional effect on the open probability of the K_ATP channel, i.e. structurally distinct segments of the linker can saturate or attenuate the open probability of the channel [68,108,109]. Additionally, the L0 linker has been reported to be involved in the binding of the sulfonlurea glibenclamide which inhibits K_ATP channels [110]. Similar to CFTR, there is an increasing number of proteins known to interact with K_ATP channels [111]. Although these proteins do not bind directly to the ABC component, the sulfonlurea receptor, they form part of the same supramolecular complex. A variety of glycolytic enzymes, adenylate kinase and creatine kinase have been shown to be part of the K_ATP channel complex and regulate its function.

The requirement of the L0 linker for transport function has also been demonstrated for MRPI [107]. In MRPI, two α-helices have been predicted in L0, although the first one with low probability [66]. L0, and especially the second helix in the loop, are absolutely essential for MRPI function [65,112–114], and might be required for correct folding and trafficking of the protein [66]. L0 has been proposed to interact both with the core domains of MRPI and with the hydrophobic membrane regions of other proteins, or with the membrane itself [66,107]. In addition, it has been suggested that at least part of the glutathione binding site is located in L0, although there is no direct evidence for this notion [115].

7.4. Accessory catalytic domains

The type I secretion pathway is used to export competence-stimulating peptides, bacteriocins, haemolysins, and non-proteinaceous material, independent of the Sec system [116–119]. In this pathway, ABC transporters have dual functions, that are, export of polypeptides and removal of leader peptides. Similar to the Sec pathway, the (poly)peptides are synthesized ribosomally and targeted for export by a N-terminal sequence of between 15 and 30 amino acids [120], that is removed upon secretion across the membrane. Removal of this signal sequence is taken care of by a peptidase domain associated with dedicated ABC transporters. The signal sequences with the consensus LXXXEXXXI-XGG are usually cleaved after the two conserved glycine residues [121] and are therefore called double-glycine-type leader peptides. The polypeptides that are released from the cell, after processing of the leader peptide, range in length between 17 and 80 amino acids [120] and often require post-translational modification prior to secretion to become biologically active. The secreted peptides rarely share significant homology, which is reflective of their
functions as molecules that act in a strain- or species-specific manner, typically as competence-stimulating peptides (ComA in Streptococcus pneumoniae) or as class I (lantibiotics) or class-II (non-lantibiotics) bacteriocins.

The secretion of ComA and bacteriocins is mediated by dedicated ABC half transporters that possess an accessory N-terminal domain of approximately 160 amino acids to remove the leader peptide [122]. The peptidase domain has characteristics of a cysteine protease, including conserved cysteine and histidine residues in the active site that are essential for proteolytic activity [123]. How the peptidase domain interacts with the other components of the transporter to promote secretion is unknown. For LcnC, the ABC transporter responsible for export of the bacteriocin LcnA, the peptidase domain was located at the cytoplasmic face of the membrane on the basis of PhoA and LacZ fusion studies (Fig. 3I; [124]).

In Gram-negative bacteria, the type I secretion systems are composed of a protein complex that spans the cytoplasmic membrane (ABC transporter), periplasm (membrane fusion protein) and outer membrane (porin-like structure). The best-characterized system is the one for secretion of HlyA, a 110 kDa haemolytic toxin, in E. coli. This system is composed of the ABC transporter HlyB, the membrane fusion protein HlyD, and the outer membrane factor TolC. Contrary to the type I secretion systems for peptides and bacteriocins, HlyB does not contain a peptidase like domain and HlyA does not contain a leader peptide. Instead, type I proteins like haemolysin contain a poorly conserved C-terminal secretion signal [125]. The secretion signal domain of HlyA has been shown to interact with the NBD in vitro [126], but the possibility cannot be excluded that the signal region of HlyA may interact with the TMD of HlyB. Also for the interaction between HlyB and HlyD there is not a clearly defined region or domain in the ABC transporter.

8. ABC transporter domains in other systems

Whereas accessory domains in ABC transporters carry out important regulatory and catalytic functions, (core) components of ABC transporters can have dual functions or be integrated into completely different systems where they are vital for function. The ABC Walker A and B motifs are found in many proteins [15], the best known examples probably being myosin, and the DNA repair enzymes Rad50 and MutS. Moreover, Rad50 and MutS not only contain the Walker A and B motifs but also the ABC signature sequence LSGGQQ and are genuine ABC proteins, albeit without transport function. In fact, a subset of ABC proteins is involved in DNA maintenance and protein synthesis, e.g., recombination, DNA repair, chromosome condensation and segregation, and translation elongation. These latter proteins exert their functions in the cytoplasm and/or nucleus rather than in the membrane [127].

SBPs of bacterial uptake systems have been shown to act as chaperones [128], and can be used as chemotaxis receptors [129]. In addition, SBP-dependent secondary transporters have been discovered [130], but an evolutionary relatedness of these SBPs to those of ABC transporters is not evident. Finally, domains homologous to prokaryotic SBPs function as receptors in ionotropic glutamate receptors present in the central nervous system of higher eukaryotes [131,132].

9. Conclusions

The core complex of an ABC transporter is composed of two NBDs and two TMDs, each in most cases with five or six transmembrane α-helical segments. A notable exception to this rule is the vitamin B12 transporter BtuCD with two times ten TMSs in the TMDs, that, on the basis of the crystal structure, all seem critical for the formation of the translocation path. The RbsC and FhuB transporters may also conform to this exception. Additional TMSs associated with the core ABC complex seem to play regulatory roles as is best documented for the transporter involved in antigen presentation and the TMD0 domain of the sulfonlurea receptors.

The SBPs are accessory components of the ABC uptake systems that have evolved to signal substrate availability to (the activation of) the ATPase, following the binding of liganded SBPs to the TMDs. Moreover, the SBPs increase the affinity of the transporters for their substrate(s) by several orders of magnitude. For a number of SBPs, functions are known that are not related to transport, including chaperoning and signal transduction. Several ABC transporters have evolved to play crucial roles in processes such as cell volume regulation, control of channel activity and antigen presentation and therefore acquired accessory regulatory functions. The CBS domains, protein modules linked to a wide variety of proteins (channels, transporters, enzymes and transcription factors), in OpuA and ProU add osmosensing and osmoregulatory functions to the transporters. Finally, ABC transporters (type I secretion systems) can have a peptidase domain that cleaves off the signal sequence prior to translocating the mature peptide (bacteriocin) or protein.

Acknowledgments: This work was supported by a “Top-subsidie” grant (NWO-CW, Grant No. 700-50-302). We thank Barry Holland for valuable discussion on the type I secretion systems.

References


[54] Boyd, D., Manoil, C. and Beckwith, J. (1987) Determinants of specificity and ionic regulation of GlnPQ from *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lactococcus lactis* the oligopeptide transport system of *Lacto
[99] Baldursson, O., Ostedgaard, L.S., Rokhlina, T., Cotten, J.F. and Csana´dy, L., Chan, K.W., Seto-Young, D., Kopsco, D.C.,
[102] Roosbeek, S., Peelman, F., Verhee, A., Labeur, C., Caster, H., Lensink, M.F., Cirulli, C., Grooten, J., Cohet, C., Van-


