Published in final edited form as:

Trends Biochem Sci. 2020 February ; 45(2): 137–148. doi:10.1016/j.tibs.2019.11.004.

Linker domains: why ABC transporters ‘live in fragments no longer’.

Robert C. Ford¹, Dominic Marshall-Sabey¹, John Schuetz²

¹Faculty of Biology Medicine and Health, the University of Manchester, Oxford Road, Manchester M13 9PL, UK.

²Pharmaceutical Sciences, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105-3678, USA.

Abstract

ATP-binding cassette (ABC) transporters are membrane proteins present in all kingdoms of life. We have considered the disordered region that connects the N- and C-terminal halves in many eukaryotic ABC transporters, allowing all four consensus functional domains to be linked. The recent availability of structures of ABC transporters containing linker regions has allowed us to identify the start and end points of the connectors as well as hinting at their localisation. We address questions such as: Where did the linker regions come from? Why do some ABC transporters have connectors and others not? What are the rules and roles of the linker regions? What are the consequences of mutations in these connector regions for disease in humans?

Keywords

ABC transporters; domain linker; structure-function; regulation; evolution

Only connect?

Only connect the prose and the passion and both will be exalted, and human love will be seen at its height. Live in fragments no longer. (EM Forster)

Many students of biochemistry will know the basics of structure/function in the **ATP binding cassette (ABC) transporter** (see Glossary) family. They consist of two transmembrane domains (TMDs) and two nucleotide-binding domains (NBDs) (Figure 1), and concerted nucleotide binding and hydrolysis by the latter is coupled to transport of substances across the membrane by the former [1]. Structural studies so far imply that the mechanism of transport fits with ideas proposed by Oleg Jardetzky many years ago - that transported **allocrites** have alternating access to the inside or outside of the membrane [2].

Correspondence : bob.ford@manchester.ac.uk.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

In ABC transporters, this access is governed by large, ATP-dependent conformational changes in the protein [3, 4].

In some prokaryotic ABC transporters, and as a minimum in nearly all eukaryotic ABC transporters, the TMD is fused to its cognate NBD and then two such entities must assemble to form a fully-functional ABC transporter. Typically, this will be a homodimeric assembly, but heterodimers are also allowed where two different ABC transporter genes, each with both TMD and NBD-encoding portions, generate the fully functional transporter. In plants, the ABCI sub-family is an exception; here, the functional transporters appear to resemble the bacterial ECF-type transporters where the TMDs and NBDs are all expressed on separate genes and must assemble together across the domain interfaces (Figure 2) [5].

In eukaryotes, but not prokaryotes, many ABC transporters have completely fused, resulting in all four domains being joined together into one functional entity. For this, a long linker sequence is employed that joins the C-terminal end of the first NBD to the N-terminal start of the second TMD. This region of the fused protein is disordered – in other words, it can adopt multiple transient structures and is highly flexible to the extent that no two ABC transporter linker regions will have exactly the same 3D conformation. Not all ABC transporters have such a linking region. More than half of the ABCB sub-family and for the most part, the ABCD, ABCG, and ABCH sub-families do not have linkers and are mostly homodimers. Yeast ABCG family homologs and plant ABCD1 orthologs do have linkers, however.

In some ABC proteins there is an additional TMD (TMD0) which precedes the first canonical TMD (TMD1). The TMD0 regions appear to be important for protein-protein interactions, as well as for subcellular trafficking and protein maturation [6, 7]. The TMD0 also has an internal linker sequence connecting it to TMD1. There are, so far, mostly low-resolution data for TMD0, but for ABCC8 there are good data and this shows the arrangement of the five TMD0 transmembrane helices, as well as its cytoplasmic linker [8]. In this case the linker is structured and connects to the ‘lasso’ structure at the start of TMD1. We will not consider further the TMD0-TMD1 linkers in this review, but instead will focus on the longer internal linkers between NBD1 and TMD2 which (so far) appear to be disordered.

Structural information for the internal linker in ABC transporters

In structural analyses where averaging of many molecules is needed (X-ray crystallography & cryo-electron microscopy (cryo-EM)), there is a smearing-out of the information for the disordered linker region. Weak, smeared density has been observed in some cases by cryo-EM. For example, an isolated cylinder-shaped density intervening between the NBDs was observed in structural studies of the cystic fibrosis transmembrane conductance regulator (CFTR/ABCC7) in the inactivated state [9]. Similarly, density connecting to TMD2 was observed for the ABCB1 linker in the inward-facing state [10]. NMR may be employed to study the dynamics of separately expressed and then purified linker peptides, and for ABCC7/CFTR this has proven very informative in understanding the likely interactions of the linker region with other purified domains and proteins [11, 12]. In the future, it may be

possible to use new NMR methods to study the dynamics of the mobile linker regions whilst still attached to their cognate ABC transporter [13].

Origins of the internal linkers

Eukaryotic ABC protein internal linkers are around 50-100 residues in length, with a preponderance of charged residues and with some regions of low complexity. This length and amino acid composition is typical for eukaryotic domain linker sequences that are structurally disordered [14]. The exception to the general rule *for linker length* is CFTR/ABCC7 which has a linker of ~200 residues and is a clear outlier when compared with other eukaryotic linkers [14]. The CFTR/ABCC7 linker sequence appears to be unique, although there is some weak homology between the C-terminal end of its linker and the ABCC8 linker. CFTR/ABCC7 is a phospho-regulated ion channel rather than a transporter [15], whilst ABCC8 regulates the Kir6.2 potassium channel as part of a large complex [8, 16]. ABC transporters in the A, C, and yeast G groups have linker regions that are related to the same sub-family, implying a linker insertion event at a point far in the past. In contrast, the ABCB sub-family linkers show much less similarity to each other. Divergence of linker sequences within ABC transporter orthologs appears to be greater than for the NBDs, but about the same as for the hydrophobic membrane-spanning regions of the TMDs.

In terms of overall structure, the ABCA sub-family differs markedly from the ABCC and ABCB sub-families, probably because it (and the structurally-related ABCG sub-family) emerged from a different branch of the prokaryotic ABC transporter tree, the 2×4 transmembrane spanning mechanotransducing ABC proteins and the 2×6 spanning glycerophospholipid and capsular polysaccharide exporters MlaEF and WzzM/T, respectively (Figure 2) [4, 17, 18].

In summary, for the ABCA group and most of the ABCC and yeast ABCG groups of transporters, it appears as though three different linkers were inserted at some point during the development of eukaryotes and then their sequences diverged at roughly the same rate as the TMDs. For the ABCB group members the situation is less clear: either (i), there has been much greater mutation and divergence in their linker sequences compared to the other domains, or (ii), functional divergence of half-transporters occurred first, followed by a later insertions of unrelated linker regions. The interesting example of the yeast ABCG linkers could be taken as evidence for a linker removal event pre-dating the evolution of mammalian ABCG transporters.

Insertion of a linker sequence implies a gene duplication followed by a gene fusion. For example, the *abcG5/G8* gene organization may represent an intermediate stage in this process. Additionally, the ABCB1 internal linker forms one exon in the gene, so we asked the question of whether there are traces of linker origins in the structure databases? At the limited significance levels implicit with such short sequences, there are echoes that have a Siren call deserving of discussion here. For example, a BLAST search against the protein data bank (PDB) using the human ABCB1 linker sequence yielded EsxB as the top non-ABCB1 hit. EsxB is a member of the prokaryotic WXG family of secreted helical hairpin peptides [19] (95% sequence coverage, E value 4.7, identity 30%). WXG proteins are

secreted by the type VII secretion system. This export system includes an ATPase subunit (EccC) with structural homology to ABC protein NBDs (RMSD between EccC and Mj0796 = 2.9Å, 87 comparisons). Cytoplasmic EccC is linked to transmembrane protein subunits that are required for EsxB export. If EsxB is a plausible structural model for the ABCB1 linker, it would imply that the linker could transiently take up a helical hairpin structure (Figure 3). The sequence alignment would predict an arginine-framed motif -RXX- at the apex of the hairpin rather than a -WXG- motif [19]. Going further towards *Sirenum scopuli*, if the ABCB1 linker is a functional echo of a transported toxin such as EsxB, then it would imply a substrate-like behaviour and a role in competitive regulation of the ABCB1 binding site for transported substrates in the inward-facing state (Figure 3). In this respect, many prokaryotic and eukaryotic ABC transporters are peptide transporters and there is evidence that ABCB1 also has the ability to transport linear and cyclic peptides [20-23] and a naturally-occurring substrate (the β -Amyloid peptide) [24]. A fusion event in an operon containing a prokaryotic ABC exporter and its peptide toxin substrate could be envisaged – perhaps as an origin for ABCB linkers. Although ABCB5 and ABCB11 linkers do not show any significant homology with the ABCB1 linker and EsxB, the ABCB5 linker shows weak homology with the heat-shock protein YbeY (and hence some structural homology to secreted proteases). The ABCB11 linker displays weak homology with the N-terminal region of a large secreted bacterial toxin (PaTox).

For ABCA, ABCC, plant ABCD, and yeast ABCG linkers, such a house of cards cannot be so easily built; a BLAST search of ABCA1 linker sequences also failed to identify any plausible hits in the PDB. However, a search of the bovine ABCC1 linker against the PDB yielded a short region of weak homology to human methionine aminopeptidase with an E value of 0.32 and an identity of 36%, but only over 52% of the (short) length of sequence. Moreover, this region is a well-structured part of the enzyme. Identification of the plant ABCD1 linker region was hampered by lack of structures, but the recent availability of the structure of human lysosomal ABCD4 [25] (PDBID-6JBJ) has allowed a better estimate for its location (the C-terminal ends of the NBDs are disordered in the 6JBJ atomic model, making this an estimate only). Taking residues 659-730 as the *Arabidopsis* ABCD1 linker yields no clue as to the origin of this linker. The top hit is the N-terminal end of the SecA ATPase (28% identity, 76% coverage). Yeast ABCG family members Pdr5, 10, and 15 are full transporters and have homologous linker regions of about 60-70 residues in length. Again, there are no clues from amino acid sequence databases as to the likely origin of these linkers.

Function(s) of the internal linker regions.

Ensuring correct ABC transporter assembly from the two halves.

There is a general trend towards the fusing of subunits in eukaryotes so that subunits of a functional complex now form linked domains of a single polypeptide [26]. We must therefore consider the hypothesis that domain fusion is an adaptation to improve the efficiency of ABC transporter folding and assembly in the increasingly complex protein mixtures in eukaryotes. The internally-linked ABCB1, ABCB4, ABCB5, ABCB11, and ABCA, as well as the ABCC sub-family members, appear to be mostly plasma membrane-

located where one might propose the membrane proteome to be at its most diverse in most cell types. Yeast ABCG family with linkers also have a predicted plasma membrane location. However, plant ABCD1 is at the peroxisomal membrane [27] and is unusual in its thioesterase activity.

Experiments where separate halves of ABCB and ABCC family members were co-expressed in cells appear to suggest that molecular recognition problems may be overcome without a linker, but the levels of the components in the membranes in these experiments may be higher than normal [28]. In the ABCB sub-family, the members that are homo- or heterodimers (ABCB2/3; ABCB6 through to ABCB10) are all proposed to be located in organelles rather than at the plasma membrane. We can speculate that this organellar segregation allows a greater chance of finding the ‘other half’ of the ABC transporter, obviating the need for an internal linker. ABCB2/3 (TAP1/2) is retained in the endoplasmic reticulum, whilst ABCB7 (TAP-like) and ABCB8 through to ABCB10 all appear to be mitochondrial transporters [29]. There is some uncertainty for ABCB6 where mitochondrial and/or lysosomal organelles have both been proposed as its home [29-32]. This transporter also has an (uncontested) presence in the plasma membrane of red blood cells where diversity in protein composition is unusually low [33, 34].

For the ABCG sub-family, all of which must dimerise to form a functional unit, there appears to be a plasma membrane location on the whole, including for the fused yeast ABCG family transporters (based on Uniprot annotation [29]). Hence, the above argument only holds for most of the eukaryotic ABC family, and if correct, suggests there is something still unknown about molecular recognition in most of the ABCG sub-family. That the heterodimeric ABCG5/G8 is encoded by back-to-back genes may provide some clues about evolution in this family [35]. For the plant ABCI family, there is likely to be a plastid location and the cellular locations of the little-studied ABCH family in insects remains unknown [36].

Regulation of ABC transporter function?

There is very strong evidence that the ABCC7/CFTR linker region is intimately involved in regulation of the channel function of the protein [37]. However, is this protein an outlier in the family? Can one propose an analogous function for linker regions in other ABC transporters? Such a hypothesis has been tested for ABCB1.

During the 1990's, a number of studies were carried out to determine the effect of phosphorylation on ABCB1. The up-regulation of protein kinase C (PKC) in selected drug resistant mammalian cancer cell lines [38-41] and the observation that PKC agonists correlated with increased levels of phosphorylated ABCB1 led to the hypothesis that PKC is a positive regulator of ABCB1 [42, 43]. To determine if the relationship between protein kinases and ABCB1 was due to direct post-translational control, mutants in which serine residues in the linker region were replaced with either alanine or the phospho-mimetic aspartate were generated. Several studies found there was no impact on the ATPase and transport activity of ABCB1 when multiple serine residues were mutated [44, 45]. However, Szabó et al. (1997) [46] found activity was reduced in wild type and aspartate mutants at lower concentrations of stimulatory drugs, whilst maximum activity was not affected.

In contrast, it was found that the ATPase activity of ABCB1-expressing Sf9 insect cells was increased by both verapamil and the addition of non-endogenous PKC, and this effect was reduced by the mutation of a single putative phosphorylation site, S671, to asparagine [47]. Interestingly, ablation of linker phosphorylation sites in transport-defective ABCB1 was shown to reverse the PKC-dependent regulation of ion channels with the suggestion that phosphorylation of ABCB1 could act as a regulator of osmotic stress responses, although this remains unconfirmed [48, 49]. For ABCB1, there is moderate conservation in mammals of the serine residue corresponding to human S661 (though in several instances it is proline at this position). There is strong conservation of the serines corresponding to human S667, S671, and S683, and most sequences examined have at least 2 proximal basic residues suggesting they are target sites for kinases [50].

To date, there is still no consensus on the effect of phosphorylation on mammalian ABCB transporters, and ABCB4 and ABCB11 appear to lack phosphorylation sites though observations of related proteins reviewed elsewhere suggest that phosphorylation is crucial in regulation [51, 52]. For example, the plant ABCB1 homolog is a selective transporter of auxin, and its activity is directly regulated by phosphorylation of the linker region [53]. These findings, the ongoing lack of consensus, and the clinical relevance of many ABC transporters perhaps warrants further tests of the hypothesis that the linker acts as a regulator of function.

Studies of various insertions and deletions of the human ABCB1 linker region shone further light on the potential role of either linker length or structure in regulating ABCB1 function. Importantly, deletion of 34 internal residues of the ~75 residue linker had no detrimental effect on either trafficking to the plasma membrane or substrate and ATP binding [54]. Unexpectedly, the catalytic activity of deleted ABCB1 was lost. This shortened linker, retaining about 30 residues in the mutated protein, seemed sufficient to span the distance between the end of NBD1 and the start of TMD2 in both the inward- and outward-facing conformations of ABCB1. This absence of catalytic activity was difficult to understand when put in the context of the activity of purified and lipid membrane-reconstituted protein; after the linker had been proteolytically cleaved into two portions [55] (Figure 3), catalytic activity was retained, but substrate selectivity altered.

Shortening of the linker in murine ABCB1 was also carried out by removing the equivalent 34 residues as above [56], and this construct was also crystallised and its structure determined. No further structural data were obtained for the linker, but the structures showed ATP binding only at NBD1. Shortening the murine ABCB1 linker also led to loss of drug stimulation of ATPase activity and drug transport activity was lost, but basal ATPase activity was still present. Although these experiments were carried out with the aim of restricting ABCB1 flexibility by shortening the linker, they suggest a role for the ABCB1 linker that goes further than a simple connector function and points to likely effects on the binding of ATP.

Taken together, these experiments suggest *ex vivo* the covalent connection the linker provides is dispensable for function, whilst its overall presence attached to either the N- or C-terminal halves of the transporter is required for correct function. However, in the cell, the

conformationally-sensitive ABCB1 antibody, UIC2 [57] demonstrated that the shortened linker ABCB1 (functionally inactive [54]) appeared locked in a substrate bound conformation. If the linker is required for correct conformation, this begs the question: what happens when the linker is either longer but unstructured, or of the same length but has an α -helical structure? Expanding the linker by insertion of an extra 17 residues, where the inserted peptide was expected to be structurally disordered, appeared to have only small effects on function that were consistent with a modestly altered conformation as detected by UIC2 [54]. In contrast, insertion of a hydrophobic, potentially helical, sequence of 17 Alanine residues at the same position led to a loss of ABCB1 expression at the plasma membrane which was not restored by addition of the substrate cyclosporine, an approach frequently used to promote folding of mutant ABCB1 and which therefore implies effects on normal membrane insertion [54]. Hence disorder and hydrophilicity may be essential components of the linker's characteristics, one that in a cellular context enables interactions promoting protein folding or other salutary interactions (but see below). Further, linker length might influence the inherent free energy with lower free energy available for the necessary functional conformations whereas a longer linker provides the 'entropic force' as has been shown to influence protein function and affinity for substrates [58]. Similarly, transient folding of the linker in the region of the NBDs may affect the binding and release of ATP and ADP/Pi, and such considerations may explain the effects of linker length.

Other workers have used the opposite approach to study linker region regulation by generating fused dimers of ABCG2 and LmrA using most of the ABCB1 linker region [59-61]. In both cases, the fused protein was active but a catalytic mutation in only one half of the fused protein became dominant negative, as expected.

If a role of the linker is to facilitate protein interactions, one might speculate that post-translational modifications such as phosphorylation will allow the ABCB1 motor to function properly at the membrane. Studies of the ABCB1 linker by Georges [62] demonstrated strong interactions with tubulin (a major cytoskeletal protein in eukaryotes and absent in prokaryotes) and noted that the region of the linker with the highest affinity harboured the three phosphorylation sites; however, it is unknown if changes in linker phosphorylation status alters ABCB1 interaction with tubulin or if this interaction is important for function of ABCB1 in cells.

ABCC1 interacts with other proteins and a robust ABCC1 protein interaction between ABCC1 and the protein ATP synthase was reported to be mediated by the internal NBD1-TMD2 linker [63]. While the interaction appears to be modulated by the phosphorylation status of linker, the biological relevance of such an interaction seemed inexplicable when first reported (the authors refer to it as "peculiar") as the ATP synthase is usually a mitochondrial protein. However, several reports have found ABCC1 lies outside of its typical plasma membrane residence. As it is reportedly functional when localized to the mitochondria [64, 65], one might speculate that the ATP synthase interaction is to chaperone ABCC1 to the mitochondria via the LI linker. Like ABCB1, the L1 linker of ABCC1 interacts with tubulin, an interaction that is modulated by phosphorylation [66]. Although the ABCB1 and ABCC1 L1 linkers are intrinsically disordered and of dissimilar sequence, one wonders if these two dissimilar linker sequences that bind tubulin do this to adopt a

structure that is biologically advantageous. If so, this would be considered an example of molecular mimicry.

For CFTR/ABCC7, the effects of removal of part of the linker have also been tested. In these experiments [67], a large portion of the linker was removed (between residues 708-835). In the light of recent structural data (PDB ID 5uak), we know that the linker left behind would be about 73 residues long (normal for an ABCC family linker length) and would have included about 63 residues from the N-terminal end (646-708) joined with 10 residues from the C-terminal end (834-843). The shortened construct retained about 30% constitutive channel activity in the presence of ATP, but this was not stimulated by phosphorylation by protein kinase A (PKA). Interestingly, the activity could be restored by exogenous addition of a separately expressed and purified linker polypeptide, but only after phosphorylation of the latter with PKA [67]. These data suggested that the linker was acting as a blocker of the channel, with block being relieved by phosphorylation. Finer dissection of the ABCC7/CFTR linker identified firstly the C-terminal end (residues 760-835) and then more precisely, residues 760-783 as being relevant for this channel blockade [68].

Consequences of mutations in the linker regions for human diseases.

Many disease-causing and disease-associated mutations have been characterised in the TMDs and NBDs of ABC transporters, and these are generally well-characterised. Perhaps surprisingly given the discussion above, some disease-associated mutations have also been found in the linker regions.

CFTR/ABCC7

For disease-associated mutations, the best-characterised ABC protein is CFTR/ABCC7 where over 2000 missense and nonsense mutations have been recorded in patients of predominantly European origins (<http://www.genet.sickkids.on.ca>). Of these, about 40 missense mutations have been identified in the 200-residue-long linker region. These mutations were identified by sequencing patient *cftr* genes, and because of the uncertainty over the effects of mutations in non-coding regions of the gene, as well as the variable symptoms exhibited, it remains possible that some, perhaps many, may be polymorphisms rather than disease-causing mutations.

Information for some of the more common missense linker mutations has been obtained by inserting them into test systems (mammalian cells where CFTR is fully stimulated). CFTR maturation and subsequent channel function can thus be measured [69, 70]. P750L and R668C were predicted to have mild disease consequences (~50% function). In contrast, V754M, I807M, and D836Y were examples of mutations that did not appear to have functional consequences, at least under the test conditions employed. This is an important caveat because CFTR/ABCC7 activity was measured with full activation of the channel (using forskolin to activate cAMP-dependent kinases for cell-based assays or using protein kinase A (PKA) for membrane patches). This can be compared to testing a car engine with the accelerator pedal pressed to the floor, and it is possible that missense mutations affecting regulation of the channel may be missed under these extreme conditions. For CFTR/ABCC7, the current consensus is therefore that linker-region missense mutations are

relatively mild, but the data for P750L and R668C suggest that these residues may be more significant than any single linker-located phosphorylated residue. Hence, the CFTR/ABCC7 mutation databases hint at some function for the linker that goes beyond purely phospho-regulation.

ABCB1

This transporter of many hydrophobic compounds is important at the blood-brain barrier and for the clearance of xenobiotics [71], therefore one might expect to find a plethora of disease-causing mutations in humans. The paucity of such may be explained by a ready functional replacement of ABCB1 by other ABC transporters such as ABCG2 and ABCC1. Indeed, ABCB1 knockout model organisms are relatively unaffected (except when exposed to some neurotoxic drugs) [72]. In contrast, gain-of-function missense mutations have been characterised for this transporter in humans [73], with two located in the linker region (L662R, R669C). Such booster mutations result in worse prognosis in cancer patients because of increased multi-drug resistance and their prevalence in the human population suggest that they may have conferred some advantage under conditions where dietary intake of xenobiotics was high in the past.

ABCC8 and 9 (SUR1/SUR2)

ABCC8 and 9, like ABCC7/CFTR, are not transporters, but form part of an ATP-sensitive potassium channel (K_{ATP}) that links the metabolic activity of cells to their membrane potential [16]. The K_{ATP} forms an inward rectifying K^+ channel composed of four K^+ channels (Kir6.1 or 6.2) and four regulatory sulphonylurea receptors (with SUR1/ABCC8 or SUR2/ABCC9). The SUR proteins regulate the gating of the Kir6 pore. Although not discussed in detail above, ABCC8 and ABCC9 have significantly longer linking regions connecting TMD1 with NBD2 which seem to be intrinsically disordered. Importantly for this discussion, SUR2B has a heme-binding function localised to this linker [74], and mutations here cause cardiac disease. SUR1 (ABCC8) lacks the heme binding domain found in the linker of SUR2. This finding raises the possibility that small molecule interactions at the TMD-NBD linkers in ABC transporters might also regulate their activity.

Missense mutations in ABC transporter linker regions that have been associated with human disease states are listed in Table I. This list is likely to grow and will become more predictive of disease as the sequencing of patient genomes and *in-vitro* testing becomes more routine. That single amino acid changes in the sequence of a linker lead to changes in function suggests that these linkers do not 'only connect'. It is our hope that this review will stimulate further research on the functions of the disordered linker regions of ABC transporters and their origins. It is possible that knowing more about the linker regions may be crucial for making progress with some of the outstanding questions in ABC transporter mechanism (Figure 4, Key Figure).

Concluding Remarks

Fusion of all four domains in eukaryotic ABC transporters via linkers may be an adaptation to the increased complexity of the eukaryotic cell and the dual challenges of correct subunit

assembly coupled to navigating to the correct destination in the cell. The linkers joining NBD1 and TMD2 appear crucial from several perspectives. Linker length appears to be important, and this constraint may be related to thermodynamics. Recent structural data showing the linker between the two halves of ABC transporters (but only in the inward-facing conformation) raises the possibility of volume exclusion effects (Figure 4), constraining the conformational landscape of the linker region. Such 'entropic force' may regulate the function of the transporter (or channel as in CFTR/ABCC7).

Despite these considerations, other experiments on the effects of single amino acid changes in the linker region imply that length, the amino acid sequence, its charged and hydrophilic nature and the structure in specific regions of the linker may also be important for function. Linker interactions with proteins or small molecules might be facilitated by modifications such as phosphorylation that would play a role in such thermodynamic forces (Figure 4), but so far links between linker phosphorylation and ABC protein regulation have been limited to a few examples. Unfortunately, no highly homologous sequences in other proteins have so far been identified that might throw light on sequence/function relationships for the ABC transporter linkers. However, some hints from prokaryotic peptide export systems point to a possible regulatory function via competitive inhibition (Figure 3). Our conclusion is that ABC transporter internal linkers do not 'only connect' but have additional regulatory roles that we are starting to understand. There remain, however, many unresolved areas of uncertainty (see Outstanding Questions).

Acknowledgements:

We thank Professor Susan Cole (Kingston University, Canada) for reading a first draft of the review and for invaluable comments and suggestions for improvement. We also thank Dr Jim Warwicker (University of Manchester, UK) and Dr Sarah Rottet (CSIRO, Australia) for helpful discussions and advice on bioinformatics. JDS is supported by NIH grants R01CA194057, R01CA194206 (JS), P30CA21745 and ALSAC. Lastly we acknowledge the superb contributions the three anonymous reviewers of this manuscript made. Without their help it would have been incomplete and out of date before the ink was dry.

Glossary:

Allocrite:

A transported substance.

ATP-binding cassette (ABC) transporter:

A family of membrane protein importers and exporters with two cytoplasmic domains, each contributing residues to two nucleotide binding pockets that are formed by domain dimerisation.

BLAST:

Basic local alignment search tool, widely employed to navigate through millions of nucleotide and amino acid sequences available in sequence databases. Cyber home to the Sirens, in this review.

ECF:

Energy-coupling factor (ECF) transporters are a category of ABC importers with a substrate-specific 'S' transmembrane domain and a general transport 'T' transmembrane domain

linked to two nucleotide binding domains. Different ‘S’ domains may interact with a ‘T’ domain.

Linker:

(As used in this review): A sequence of amino acids linking the first nucleotide binding domain with the second transmembrane domain, typically having little or no stable 3D structure.

PDB:

Protein databank. Repository of available protein structural models, listed according to a 4-character code. No longer includes homology models.

Sirenum scopuli:

The mythical islands, home to the Sirens. Upon their rocks, mariners were lured to their fate (shipwreck).

Volume exclusion effects/entropic force:

A disordered linker that is in some way tethered has a restricted (excluded) volume to explore and has limitations on available conformations. This is an entropic penalty that may be considered as a ‘force’ exerted on the system that will be influenced by the linker length.

References

- Higgins CF (1992) ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* 8, 67–113. [PubMed: 1282354]
- Jardetzky O (1966) Simple allosteric model for membrane pumps. *Nature* 211 (5052), 969–70. [PubMed: 5968307]
- Jardetzky O and Snell FM (1960) Theoretical Analysis of Transport Processes in Living Systems. *Proc Natl Acad Sci U S A* 46 (5), 616–22. [PubMed: 16590649]
- Ford RC and Beis K (2019) Learning the ABCs one at a time: structure and mechanism of ABC transporters. *Biochem Soc Trans* 47 (1), 23–36. [PubMed: 30626703]
- Eitinger T et al. (2011) Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol Rev* 35 (1), 3–67. [PubMed: 20497229]
- Bock C et al. (2018) Structural and functional insights into the interaction and targeting hub TMD0 of the polypeptide transporter TAPL. *Sci Rep* 8 (1), 15662. [PubMed: 30353140]
- Fernandez SB et al. (2002) Role of the N-terminal transmembrane region of the multidrug resistance protein MRP2 in routing to the apical membrane in MDCKII cells. *J Biol Chem* 277 (34), 31048–55. [PubMed: 12060660]
- Martin GM et al. (2017) Cryo-EM structure of the ATP-sensitive potassium channel illuminates mechanisms of assembly and gating. *Elife* 6.
- Liu F et al. (2017) Molecular Structure of the Human CFTR Ion Channel. *Cell* 169 (1), 85–95 e8. [PubMed: 28340353]
- Thonghin N et al. (2018) Novel features in the structure of P-glycoprotein (ABCB1) in the post-hydrolytic state as determined at 7.9 Å resolution. *BMC Struct Biol* 18 (1), 17. [PubMed: 30545335]
- Baker JM et al. (2007) CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat Struct Mol Biol* 14 (8), 738–45. [PubMed: 17660831]
- Bozoky Z et al. (2013) Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions. *Proc Natl Acad Sci U S A* 110 (47), E4427–36. [PubMed: 24191035]

13. Kay LE (2016) New Views of Functionally Dynamic Proteins by Solution NMR Spectroscopy. *J Mol Biol* 428 (2 Pt A), 323–331. [PubMed: 26707200]
14. Wang M et al. (2011) Reductive evolution of proteomes and protein structures. *Proc Natl Acad Sci U S A* 108 (29), 11954–8. [PubMed: 21730144]
15. Riordan JR (2008) CFTR function and prospects for therapy. *Annu Rev Biochem* 77, 701–26. [PubMed: 18304008]
16. Bryan J and Aguilar-Bryan L (1999) Sulfonylurea receptors: ABC transporters that regulate ATP-sensitive K(+) channels. *Biochim Biophys Acta* 1461 (2), 285–303. [PubMed: 10581362]
17. Kamischke C et al. (2019) The *Acinetobacter baumannii* Mla system and glycerophospholipid transport to the outer membrane. *Elife* 8.
18. Bi Y et al. (2018) Architecture of a channel-forming O-antigen polysaccharide ABC transporter. *Nature* 553 (7688), 361–365. [PubMed: 29320481]
19. Fan Y et al. (2015) EsxB, a secreted protein from *Bacillus anthracis* forms two distinct helical bundles. *Protein science : a publication of the Protein Society* 24 (9), 1389–400. [PubMed: 26032645]
20. Sharom FJ et al. (1998) Linear and cyclic peptides as substrates and modulators of P-glycoprotein: peptide binding and effects on drug transport and accumulation. *Biochem J* 333 (Pt 3), 621–30. [PubMed: 9677321]
21. DiDiodato G and Sharom FJ (1997) Interaction of combinations of drugs, chemosensitizers, and peptides with the P-glycoprotein multidrug transporter. *Biochem Pharmacol* 53 (12), 1789–97. [PubMed: 9256153]
22. Sharom FJ et al. (1996) Synthetic hydrophobic peptides are substrates for P-glycoprotein and stimulate drug transport. *Biochem J* 320 (Pt 2), 421–8. [PubMed: 8973548]
23. Sharom FJ et al. (1995) Interaction of the P-glycoprotein multidrug transporter with peptides and ionophores. *J Biol Chem* 270 (17), 10334–41. [PubMed: 7730340]
24. van Assema DM and van Berckel BN (2016) Blood-Brain Barrier ABC-transporter P-glycoprotein in Alzheimer's Disease: Still a Suspect? *Curr Pharm Des* 22 (38), 5808–5816.
25. Xu D et al. (2019) Cryo-EM structure of human lysosomal cobalamin exporter ABCD4. *Cell Res*.
26. Doolittle RF (1995) The multiplicity of domains in proteins. *Annu Rev Biochem* 64, 287–314. [PubMed: 7574483]
27. Carrier DJ et al. (2019) Mutagenesis separates ATPase and thioesterase activities of the peroxisomal ABC transporter, Comatose. *Sci Rep* 9 (1), 10502. [PubMed: 31324846]
28. Moriya H (2015) Quantitative nature of overexpression experiments. *Molecular biology of the cell* 26 (22), 3932–9. [PubMed: 26543202]
29. (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic acids research* 47 (D1), D506–D515. [PubMed: 30395287]
30. Fukuda Y et al. (2011) Conserved intramolecular disulfide bond is critical to trafficking and fate of ATP-binding cassette (ABC) transporters ABCB6 and sulfonylurea receptor 1 (SUR1)/ABCC8. *J Biol Chem* 286 (10), 8481–92. [PubMed: 21199866]
31. Krishnamurthy P and Schuetz JD (2011) The role of ABCG2 and ABCB6 in porphyrin metabolism and cell survival. *Curr Pharm Biotechnol* 12 (4), 647–55. [PubMed: 21118089]
32. Kiss K et al. (2015) Role of the N-terminal transmembrane domain in the endo-lysosomal targeting and function of the human ABCB6 protein. *Biochem J* 467 (1), 127–39. [PubMed: 25627919]
33. Helias V et al. (2012) ABCB6 is dispensable for erythropoiesis and specifies the new blood group system Langereis. *Nat Genet* 44 (2), 170–3. [PubMed: 22246506]
34. Fukuda Y et al. (2016) The severity of hereditary porphyria is modulated by the porphyrin exporter and Lan antigen ABCB6. *Nat Commun* 7, 12353. [PubMed: 27507172]
35. Hubacek JA et al. (2001) Mutations in ATP-cassette binding proteins G5 (ABCG5) and G8 (ABCG8) causing sitosterolemia. *Hum Mutat* 18 (4), 359–60.
36. Guo Z et al. (2015) The novel ABC transporter ABCH1 is a potential target for RNAi-based insect pest control and resistance management. *Sci Rep* 5, 13728. [PubMed: 26333918]
37. Hegedus T et al. (2009) Role of individual R domain phosphorylation sites in CFTR regulation by protein kinase A. *Biochim Biophys Acta* 1788 (6), 1341–9. [PubMed: 19328185]

38. Fine RL et al. (1988) PHORBOL ESTERS INDUCE MULTIDRUG RESISTANCE IN HUMAN-BREAST CANCER-CELLS. *Proceedings of the National Academy of Sciences of the United States of America* 85 (2), 582–586. [PubMed: 3422442]
39. Posada J et al. (1989) PROTEIN KINASE-C IN ADRIAMYCIN ACTION AND RESISTANCE IN MOUSE SARCOMA 180 CELLS. *Cancer Research* 49 (23), 6634–6639. [PubMed: 2819714]
40. Gollapudi S et al. (1992) PROTEIN-KINASE-C ISOFORMS IN MULTIDRUG RESISTANT P388/ADR CELLS - A POSSIBLE ROLE IN DAUNORUBICIN TRANSPORT. *Cancer Letters* 62 (1), 69–75. [PubMed: 1347251]
41. Posada JA et al. (1989) Human multidrug resistant KB cells overexpress protein kinase C: involvement in drug resistance. *Cancer communications* 1 (5), 285–92. [PubMed: 2577153]
42. Hamada H et al. (1987) PHOSPHORYLATION OF THE MR170,000 TO 180,000 GLYCOPROTEIN SPECIFIC TO MULTIDRUG-RESISTANT TUMOR-CELLS - EFFECTS OF VERAPAMIL, TRIFLUOPERAZINE, AND PHORBOL ESTERS. *Cancer Research* 47 (11), 2860–2865. [PubMed: 3567906]
43. Chambers TC et al. (1990) PROTEIN KINASE-C PHOSPHORYLATES P-GLYCOPROTEIN IN MULTIDRUG RESISTANT HUMAN KB CARCINOMA-CELLS. *Journal of Biological Chemistry* 265 (13), 7679–7686. [PubMed: 1970571]
44. Goodfellow HR et al. (1996) Protein Kinase C-mediated Phosphorylation Does Not Regulate Drug Transport by the Human Multidrug Resistance P-glycoprotein. *Journal of Biological Chemistry* 271 (23), 13668–13674. [PubMed: 8662768]
45. Germann UA et al. (1996) Characterization of phosphorylation-defective mutants of human P-glycoprotein expressed in mammalian cells. *Journal of Biological Chemistry* 271 (3), 1708–1716. [PubMed: 8576173]
46. Szabo K et al. (1997) Phosphorylation site mutations in the human multidrug transporter modulate its drug-stimulated ATPase activity. *Journal of Biological Chemistry* 272 (37), 23165–23171. [PubMed: 9287320]
47. Ahmad S et al. (1994) MODULATION OF P-GLYCOPROTEIN BY PROTEIN-KINASE C-ALPHA IN A BACULOVIRUS EXPRESSION SYSTEM. *Biochemistry* 33 (34), 10313–10318. [PubMed: 7915139]
48. Hardy SP et al. (1995) PROTEIN KINASE-C-MEDIATED PHOSPHORYLATION OF THE HUMAN MULTIDRUG-RESISTANCE P-GLYCOPROTEIN REGULATES CELL VOLUME-ACTIVATED CHLORIDE CHANNELS (VOL 14, PG 68, 1995). *Embo Journal* 14 (8), 1844–1844.
49. Gill DR et al. (1992) Separation of drug transport and chloride channel functions of the human multidrug resistance P-glycoprotein. *Cell* 71 (1), 23–32. [PubMed: 1382860]
50. Pearlman SM et al. (2011) A mechanism for the evolution of phosphorylation sites. *Cell* 147 (4), 934–46. [PubMed: 22078888]
51. Aryal B et al. (2016) Learning from each other: ABC transporter regulation by protein phosphorylation in plant and mammalian systems (vol 43, pg 966). *Biochemical Society Transactions* 44, 663–673. [PubMed: 27068986]
52. Stolarczyk EI et al. (2011) Regulation of ABC Transporter Function Via Phosphorylation by Protein Kinases. *Current Pharmaceutical Biotechnology* 12 (4), 621–635. [PubMed: 21118091]
53. Henrichs S et al. (2012) Regulation of ABCB1/PGP1-catalysed auxin transport by linker phosphorylation. *Embo Journal* 31 (13), 2965–2980. [PubMed: 22549467]
54. Hrycyna CA et al. (1998) Structural flexibility of the linker region of human P-glycoprotein permits ATP hydrolysis and drug transport. *Biochemistry* 37 (39), 13660–73. [PubMed: 9753453]
55. Sato T et al. (2009) Functional role of the linker region in purified human P-glycoprotein. *FEBS J* 276 (13), 3504–16. [PubMed: 19490125]
56. Esser L et al. (2017) Structures of the Multidrug Transporter P-glycoprotein Reveal Asymmetric ATP Binding and the Mechanism of Polyspecificity. *J Biol Chem* 292 (2), 446–461. [PubMed: 27864369]
57. Vahedi S et al. (2018) Mapping discontinuous epitopes for MRK-16, UIC2 and 4E3 antibodies to extracellular loops 1 and 4 of human P-glycoprotein. *Sci Rep* 8 (1), 12716. [PubMed: 30143707]

58. Keul ND et al. (2018) The entropic force generated by intrinsically disordered segments tunes protein function. *Nature* 563 (7732), 584–588. [PubMed: 30420606]
59. Shilling RA et al. (2003) A new dimer interface for an ABC transporter. *Int J Antimicrob Agents* 22 (3), 200–4. [PubMed: 13678821]
60. Bhatia A et al. (2005) Oligomerization of the human ABC transporter ABCG2: evaluation of the native protein and chimeric dimers. *Biochemistry* 44 (32), 10893–904. [PubMed: 16086592]
61. van Veen HW et al. (2000) The homodimeric ATP-binding cassette transporter LmrA mediates multidrug transport by an alternating two-site (two-cylinder engine) mechanism. *EMBO J* 19 (11), 2503–14. [PubMed: 10835349]
62. Georges E (2007) The P-glycoprotein (ABCB1) linker domain encodes high-affinity binding sequences to alpha- and beta-tubulins. *Biochemistry* 46 (25), 7337–42. [PubMed: 17530867]
63. Yang Y et al. (2012) Human ABCC1 interacts and colocalizes with ATP synthase alpha, revealed by interactive proteomics analysis. *J Proteome Res* 11 (2), 1364–72. [PubMed: 22188235]
64. Roundhill EA and Burchill SA (2012) Detection and characterisation of multi-drug resistance protein 1 (MRP-1) in human mitochondria. *Br J Cancer* 106 (6), 1224–33. [PubMed: 22353810]
65. Jungsuwadee P et al. (2009) Mrp1 localization and function in cardiac mitochondria after doxorubicin. *Mol Pharmacol* 75 (5), 1117–26. [PubMed: 19233900]
66. Ambadipudi R and Georges E (2017) Sequences in Linker-1 domain of the multidrug resistance associated protein (MRP1 or ABCC1) bind to tubulin and their binding is modulated by phosphorylation. *Biochem Biophys Res Commun* 482 (4), 1001–1006. [PubMed: 27908733]
67. Ma J et al. (1997) Function of the R domain in the cystic fibrosis transmembrane conductance regulator chloride channel. *J Biol Chem* 272 (44), 28133–41. [PubMed: 9346969]
68. Baldursson O et al. (2001) Cystic fibrosis transmembrane conductance regulator Cl-channels with R domain deletions and translocations show phosphorylation-dependent and-independent activity. *J Biol Chem* 276 (3), 1904–10. [PubMed: 11038358]
69. Raraigh KS et al. (2018) Functional Assays Are Essential for Interpretation of Missense Variants Associated with Variable Expressivity. *American journal of human genetics* 102 (6), 1062–1077. [PubMed: 29805046]
70. Salinas DB et al. (2015) Benign outcome among positive cystic fibrosis newborn screen children with non-CF-causing variants. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 14 (6), 714–9. [PubMed: 25824995]
71. Terasaki T and Ohtsuki S (2005) Brain-to-blood transporters for endogenous substrates and xenobiotics at the blood-brain barrier: an overview of biology and methodology. *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics* 2 (1), 63–72. [PubMed: 15717058]
72. Roulet A et al. (2003) MDR1-deficient genotype in Collie dogs hypersensitive to the P-glycoprotein substrate ivermectin. *European journal of pharmacology* 460 (2-3), 85–91. [PubMed: 12559367]
73. Jeong H et al. (2007) Function-altering SNPs in the human multidrug transporter gene ABCB1 identified using a *Saccharomyces*-based assay. *PLoS Genet* 3 (3), e39. [PubMed: 17352537]
74. Burton MJ et al. (2016) A heme-binding domain controls regulation of ATP-dependent potassium channels. *Proc Natl Acad Sci U S A* 113 (14), 3785–90. [PubMed: 27006498]

Highlights

- The ATP-binding cassette (ABC) proteins are essential in all living organisms and are mostly transporters of substances across biological membranes.
- Some eukaryotic ABC transporters have disordered linker regions that join their amino- and carboxy-terminal halves together, perhaps as an adaptation to eukaryotic complexity.
- The evolutionary origins of eukaryotic ABC transporters may provide insight about where the linker regions came from and their function(s).
- The linker region may possess a number of functions and may be regulated by post-translational modifications, such as phosphorylation.
- Mutations to the linker regions have been observed in a number of human diseases.

Outstanding questions:

1. What are the mechanisms by which the disordered linker regions and their phosphorylation status regulate ABC transporters?
2. How do single missense mutations in disordered linker regions of ABC transporters lead to disease?
3. Where did the linker regions come from?
4. What selective pressure favoured the insertion of linkers into some, but not all, eukaryotic ABC transporter genes?
5. What happened to the type I and II bacterial ABC importers? Did they ever leap across into eukaryotes?
6. Are plant chloroplasts the only eukaryotic location of ECF-like ABC transporters?

Figure 1:

ABC transporters with internal linkers: Disordered linkers are indicated by the black dashed lines. The N-terminal and C-terminal halves of the protein are shown in green and red, respectively. Weak/low resolution density proposed to be part of the ABCC7 linker was fitted as a helix as indicated in dark blue. All structures were determined in the absence of nucleotide. Structural data for another C family member with a linker (ABCC8) resembles ABCC1 and is not shown. The approximate location of the lipid bilayer is indicated in blue. Note that the connector sequences are much longer than needed to span the distance indicated by the lines. ABCC7 has the longest linking sequence, with almost 200 residues, whilst the others shown have linkers of 60-90 residues in length.

Figure 2:

Proposal for the evolution of ABC transporters: In prokaryotes (left), associations between a progenitor NBD and different transmembrane (TM) subunits evolved to yield different types of ABC importers or exporters. Apart from the Type I and II importers (blue, top); all the other prokaryotic ABC transporters appear to have crossed over into eukaryotes (right) where they evolved to form the current eukaryotic ABC transporter sub-families (ABCA to ABCI). Some divergence of function away from transport is indicated for the ABCC proteins ABCC7-9 as well as the NBD-only families (ABCE and ABCF).

Figure 3:

ABCB1 linker modeled on EsxB: The sequence alignment between human ABCB1 and *B. anthracis* EsxB is shown below with the WxG motif and a possible RxR motif highlighted in yellow. Trypsin (black arrows), chymotrypsin (red arrow) and V8 protease (blue arrow) cleavage sites are indicated [55]. The N-terminal (green) and C-terminal (red) halves of the ABCB1 structures (PDBID-6Q81, 6C0V) are shown as ribbons (above). The EsxB hairpin-shaped polypeptide (blue) was manually docked into the ABCB1 central cavity. The left-handed superhelical twist of EsxB brings the ends of the hairpin close to the ABCB1 nucleotides (right, spacefill representation). Un-modeled regions of linker are indicated by the blue lines.

Figure 4, Key Figure:

Schema for considering the possible regulatory roles of the linker region in ABC transporters. (i) Here the linker is interacting with ABC transporter domains and – perhaps via excluded volume (entropic) effects is forcing a quiescent state – either by blocking the allocrite binding site in the TMDs or the ATP-binding site in the NBDs. (ii) Some regulatory event such as phosphorylation allows the clearance of the linker from its inhibitory location and frees the allocrite site in the inward-facing conformation. (iii) ATP binding to the inward-facing state can now trigger a conformational switch to the outward-facing state and successful transport of the allocrite across the membrane. Whether the linker plays a role in modulating the outward-facing state is a moot point.

Table 1:

Mutations/polymorphisms identified in ABC transporter linker regions that are associated with disease states. Data were obtained from the Uniprot database (<https://www.uniprot.org>), the CFTR mutation database (<http://www.genet.sickkids.on.ca>) and the CFTR2 database (<https://www.cftr2.org>).

Transporter Linker Mutation	Effect	Disease
<i>ABCB1</i>		
L662R	GoF ^a	Cancer/MDR ^b
R669C	GoF	Cancer/MDR
<i>ABCC6</i>		
R881S	LoF ^a	PXE ^c
<i>ABCC7</i>		
R668C	Mild LoF (FA ^d)	CF ^f
P750L	Mild LoF (FA)	CF
V754M	None (FA)	CF
I807M	None (FA)	CF
D836Y	None (FA)	CF
+35 others	Suspected LoF/U ^e	CF
<i>ABCA1</i>		
D1289N	LoF	HDLD1 ^g
<i>ABCA4</i>		
L1250P	LoF	STGD ^h
T1253M	LoF	FFM ⁱ
R1300Q	LoF	STGD

^aGoF/LoF – gain/loss of function

^bMDR – multi-drug resistance

^cPXE – pseudoxanthoma elasticum

^dFA – tested with full activation

^eU – untested

^fCF – cystic fibrosis

^gHDLD1 – High density lipoprotein deficiency one

^hSTGD – Stargardt disease

ⁱFFM – Fundus flavimaculatus