ATP binding cassette importers in eukaryotic organisms

Cheri C. Choi1,2* and Robert C. Ford1

1 Faculty of Biology Medicine and Health, School of Biological Sciences, The University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.
2 Department of Biology, University of York, York, YO10 5DD, U.K.

ABSTRACT

ATP-binding cassette (ABC) transporters are ubiquitous across all realms of life. Dogma suggests that bacterial ABC transporters include both importers and exporters, whilst eukaryotic members of this family are solely exporters, implying that ABC import function was lost during evolution. This view is being challenged, for example energy-coupling factor (ECF)-type ABC importers appear to fulfill important roles in both algae and plants where they form the ABCI sub-family. Herein we discuss whether bacterial Type I and Type II ABC importers also made the transition into extant eukaryotes. Various studies suggest that Type I importers exist in algae and the liverwort family of primitive non-vascular plants, but not in higher plants. The existence of eukaryotic Type II importers is also supported: a transmembrane protein homologous to vitamin B12 import system transmembrane protein (BtuC), hemin transport system transmembrane protein (HmuU) and high-affinity zinc uptake system membrane protein (ZnuB) is present in the Cyanophora paradoxa genome. This protein has homologs within the genomes of red algae. Furthermore, its candidate nucleotide-binding domain (NBD) shows closest similarity to other bacterial Type II importer NBDs such as BtuD. Functional studies suggest that Type I importers have roles in maintaining sulphate levels in the chloroplast, whilst Type II importers probably act as importers of Mn²⁺ or Zn²⁺, as inferred by comparisons with bacterial homologs. Possible explanations for the presence of these transporters in simple plants, but not in other eukaryotic organisms, are considered. In order to utilise the existing nomenclature for eukaryotic ABC proteins, we propose that eukaryotic Type I and II importers be classified as ABCJ and ABCK transporters, respectively.

Key words: ATP binding cassette transporters, eukaryotes, membrane transporter, Type I importer, Type II importer, algae

CONTENTS

I. Introduction ................................................................. 1319
II. Type i abc transporters .................................................. 1319
   (1) Bacterial Type I ABC importers ..................................... 1319
   (2) Type I ABC importers in eukaryotes .............................. 1322
   (3) Evidence for Type I ABC importers in other eukaryotes .. 1323
III. Type ii abc importers .................................................. 1324
   (1) Bacterial Type II ABC importers .................................. 1324
   (2) Type II ABC importers in eukaryotes ........................... 1324
   (3) Other Type II ABC importers in eukaryotes .................. 1325
IV. Roles of eukaryotic type i and ii abc importers ............... 1326
V. Proposed naming of eukaryotic type i and type ii importers .. 1327
VI. Conclusions ............................................................. 1327
VII. Acknowledgements ................................................... 1328
VIII. References ............................................................ 1328
IX. Supporting information .............................................. 1329

* Address for correspondence (Tel: +07539996886; E-mail: cheri.choi@yahoo.co.uk)
I. INTRODUCTION

ATP binding cassette (ABC) transporters represent one of the largest families of transmembrane proteins and are found in all major divisions of living organisms (Higgins, 1992; Holland, 2011). These proteins share a highly conserved nucleotide-binding domain (NBD) with typical Walker A and B sequence motifs as well as an ABC transporter-specific ‘signature’ sequence (Higgins, 1992). By contrast, the transmembrane portions of the transporters can vary considerably and can be categorised into at least seven types according to their overall topology (Fig. 1) (Thomas et al., 2020). In humans, there are 48 known ABC proteins (Dean, Rzhetsky & Allikmets, 2001), segregating into seven distinct groups (ABCA, ABCB, ABCC, ABCD, ABCF and ABCG), nearly all of which have been associated with disease states, usually involving loss of function (Dean et al., 2001; Ford & Beis, 2019; Liu, 2019). Bacterial ABC transporters are classified into exporters and importers (Higgins, 2001; Ford & Beis, 2019). The bacterial ABC exporters can be categorised into three groups on the basis of the structures of their transmembrane domains (TMDs), and they have essential roles in the high-affinity uptake of various nutrients (Higgins, 2001). The bacterial exporters are associated with the general export of components across the inner membrane into the periplasm, the outer membrane or the exterior of the cell, and play important roles in antibiotic resistance (Higgins, 2001; Ford & Beis, 2019). One of the most surprising features of ABC transporters is the wide range of substances that they can transport across membranes, not only in terms of the chemical properties of the transported substances, but also their size (Higgins, 1992; Holland, 2011).

Until recently, eukaryotic ABC transporters were considered to function only as exporters. However some ABC transporters in the ABCC sub-family have evolved into regulators of other ion channels (ABCC3/SUR1 and ABCC9/SUR2), or in one case, a bone fide ion channel (ABCC7/CFTR). Furthermore, recent work on a human ABC transporter in the ABCA family has provided evidence that ABCA4 operates as an importer (Quazi, Lenevich & Allikmets, 2001), segregating into seven distinct groups (ABCA, ABCB, ABCC, ABCD, ABCF and ABCG), nearly all of which have been associated with disease states, usually involving loss of function (Dean et al., 2001; Ford & Beis, 2019; Liu, 2019). Bacterial ABC transporters are classified into exporters and importers (Higgins, 2001; Ford & Beis, 2019). The bacterial ABC exporters can be categorised into three groups on the basis of the structures of their transmembrane domains (TMDs), and they have essential roles in the high-affinity uptake of various nutrients (Higgins, 2001). The bacterial exporters are associated with the general export of components across the inner membrane into the periplasm, the outer membrane or the exterior of the cell, and play important roles in antibiotic resistance (Higgins, 2001; Ford & Beis, 2019). One of the most surprising features of ABC transporters is the wide range of substances that they can transport across membranes, not only in terms of the chemical properties of the transported substances, but also their size (Higgins, 1992; Holland, 2011).

Apart from the ECF importers, two other types of ABC importer have been identified in bacteria: the Type I importer has two TMDs that each consist of five core short transmembrane (TM) α-helices with a characteristic topology (Kadaba et al., 2008; Khare et al., 2009; Rice, Park & Pinkett, 2014) (Fig. 2). The Type II importer has two TMDs with 10 transmembrane α-helices each and with a topology different to the Type I importer (Locher, Lee & Rees, 2002; Pinkett et al., 2007) (Fig. 2). All bacterial ABC importers appear to utilise a substrate binding protein (SBP) component, but the Type I and Type II importer SBPs are primarily soluble proteins that are constrained either by the outer membrane (in gram-negative bacteria) or by a lipid anchor (in gram-positive bacteria) (Davidson & Chen, 2004; Davidson et al., 2008; Rice et al., 2014). A recent discussion on the origins of eukaryotic ABC transporters considered whether Type I and Type II bacterial ABC importers were also transferred into extant eukaryotes, like other types of bacterial ABC transporter (Ford & Beis, 2019; Ford et al., 2020). Here we discuss evidence for the existence of Type I and Type II importers in eukaryotic autotrophic organisms (algae and non-vascular plants). These putative importers seem to function in the transport of sulphate and manganese ions into the chloroplast where they support amino acid synthesis and oxygenic photosynthesis.

II. TYPE I ABC TRANSPORTERS

(1) Bacterial Type I ABC importers

Bacterial Type I ABC importers are involved in transport of metabolites that the cell requires in significant quantities such as sugars and amino acids (in rich growth media), or basic components for biomolecule synthesis such as sulphate and phosphate (ter Beck, Guskov & Slotboom, 2014). Because these metabolites may be limited outside the bacterial cell, high-affinity SBPs are employed in these transport systems.
to scavenge and shuttle the substrate to the ABC transporter (Davidson et al., 2008; Eitinger et al., 2011).

Like all ABC transporters, Type I importers have four core domains: two TMDs and two NBDs (Linton, 2007; Ford & Beis, 2019). The NBDs are typical of the whole ABC family, but some features do discriminate them from other ABC transporter NBDs. For example, in five Type I importer proteins for which there are available structures, four have a

Fig. 1. Structures of different types of bacterial ATP binding cassette (ABC) transporters: ABC transporters utilise a universal nucleotide-binding domain (NBD) bolted on to different transmembrane domains (TMDs), which can be classified according to their folds (Thomas et al., 2020). The three types of importers are shown on the right and the four types of exporters at the bottom, with individual TMD subunits coloured differently. Types I, II and III importers show completely different TMD topologies. Types V, VI and VII exporters share a somewhat similar topology for the first four core transmembrane spans, whereas type IV exporters have a completely different arrangement where the fourth and fifth transmembrane helices domain-swap to form contacts with the opposite side of the transporter. The approximate boundaries of the membrane are indicated with dashed lines.
C-terminal extension that forms a separate small domain below the NBD (Hollenstein, Frei & Locher, 2007; Oldham et al., 2007; Gerber et al., 2008; Maruyama et al., 2015). These have been suggested to have regulatory functions (Davidson et al., 2008), but may also help to maintain contact between the two NBDs that rock inwards and outwards during the catalytic cycle (Ford & Hellmich, 2020). The Type I importer TMD consists of a conserved core of five TM spans and a highly conserved cytoplasmic loop (often termed the EAA loop) that is present between the third and fourth TM spans (Mourez, Hofnung & Dassa, 1997; Wen & Tajkhorshid, 2011) of the core (Fig. 2). This loop forms contacts with the underlying NBD via the so-called Q-loop region of the NBD and is involved in assembly of the NBD onto the TMDs and in communication of the presence of a substrate from the NBD to TMD and vice-versa (Mourez et al., 1997; Eckey, Landmesser & Schneider, 2010; Wen & Tajkhorshid, 2011). In Type I importers with six TM spans, the extra span is at the N-terminus, and it forms extensive contacts with the opposing TMD, presumably stabilising the complex. This may explain why most of the Type I importer proteins for which we have structures to date display this additional TM helix, since structural studies of membrane proteins are facilitated by the use of highly robust proteins (Thonghin et al., 2018). A small number of Type I TMDs have additional membrane-spanning helices. For

Fig. 2. Topology of Type I and Type II ATP binding cassette (ABC) importer transmembrane domains (TMDs) illustrated in cartoon form (left) and with colouring of the three-dimensional structure (centre and right panels). Top panels: Type I TMD topology illustrated using the maltose transport system transmembrane protein MalG (PDB ID: 2R6G). Dark blue, N-terminal TM0; cyan, TM1; dark green, TM2; light green, TM3; yellow, cytoplasmic loop and TM4; red, TM5. In some Type I importers, TM0 (dark blue) is missing. Lower panels: Type II TMD topology illustrated using the vitamin B12 import system transmembrane protein BtuC (PDBID 4dbl). Dark blue, N-terminal TM1; light blue, TM2; dark green, TM3; light green, TM4; yellow, TM5; orange, TM6 preceded by an external loop that dips into the membrane-spanning region; orange-red, NBD-contacting loop and TM7; red, short/kinked helix TM8; cyan, TM9; purple, C terminal TM10. In the centre panel, contacting TM spans from the opposing TMD (grey) are also shown. The panels on the right show the structure after rotation of the models by 90 degrees around the membrane normal and from the perspective of the opposing TMD. The approximate boundaries of the lipid bilayer are indicated in the left panels by dashed lines.
example the transmembrane protein of the maltose importer MalF, which heterodimerises with another transmembrane protein MalG in the same importer, has eight TM spans, with TM helices 3–7 forming the core domain and TM spans 1, 2 and 8 being non-core regions (Davidson & Chen, 2004; Khare et al., 2009).

The transition from the nucleotide-free to ATP-bound state has been studied for the MalFGK2 importer (Khare et al., 2009). In the outward-facing state, with ATP and the SBP present, the formation of the NBD sandwich dimer forces together the two EAA loops between core TM spans 4 and 5 (Figs 2 & 3). This causes rotation of core TM helices 4 and 5 in each TMD and this seals off the cytoplasmic side of a central cavity through the membrane. The interactions between the four helices appear to be mostly via van der Waals interactions of hydrophobic side chains. In the nucleotide-free, inward-facing state, the outward movement of the NBDs after dissociation of the sandwich dimer pulls apart the two EAA loops and this causes rotation of core TM helices 4 and 5 so that they now seal the outside of the central cavity (Oldham & Chen, 2011a,b; Ford & Beis, 2019). This plug also involves the participation of core TM helix 3.

(2) Type I ABC importers in eukaryotes

Bioinformatics analyses of the most up-to-date sequence databases (as of August 2020) can be used to identify potential Type I importers in algae. For example, a recent PSI-BLAST search (Altschul et al., 1990) of Type I importer TMD sequences of MalG, MetI, a transmembrane protein from D-methionine transport system, and ModB, a transmembrane protein from a molybdenum transport system, against eukaryotic protein sequences revealed many potential hits in algal genomes, of which all are annotated as proteins associated with sulphate transport. The well-studied model alga *Chlamydomonas reinhardtii* has two transmembrane protein homologs: the sulphate transport proteins CysT/SULP1 and CysW/SULP2. These can potentially form a heterodimeric TMD complex that associates with two copies of a single nucleotide-binding domain (NBD) protein termed CysA/Sabc and a single SBP, CysP/Sbp (Melis & Chen, 2005). The characterisation and naming of these proteins in algae derives from earlier studies on cyanobacterial sulphate importers (Laudenbach & Grossman, 1991) that we now know to be Type I ABC importers because of their sequence homology with available structures of Type I importers (see Section II.1). The CysA NBD subunit has a large C-terminal

![Fig. 3. Homology models of target proteins, the sulphate transport system transmembrane protein CysT and the nucleotide-binding protein CysA from *Chlamydomonas reinhardtii*, intersecting with their template proteins from the maltose transport system, MalG, which heterodimerises with another transmembrane protein MalF, and the nucleotide-binding protein MalK, respectively, derived from the crystal structure of the Type I importer MalFGK2 from *Escherichia coli* (PDB ID: 4JBW). The view is from along the membrane plane. (A) Transmembrane domain (TMD) of CysT (Uniprot ID: Q8RVC7) model is coloured in lilac. Percentage sequence identity (%ID) of CysT and MalG is 20.9%. (B) Nucleotide-binding domain (NBD) of CysA (Uniprot ID: Q6QJE1) model is coloured in pink. %ID of CysA and MalK is 30.4%. (C) The homology models for the TMD of CysT and the NBD of CysA displayed separately. The green/blue atoms/spheres highlight two charged residues (R346 and E350) mentioned in the main text. Homology models were generated using Chimera (Petterson et al., 2004), where multiple sequence alignments were applied to construct a model of the target protein interactively with the three-dimensional structure of its homologous template proteins. All homology models can be obtained from R. Ford (robert.ford@manchester.ac.uk).](image-url)
extension that is homologous to the regulatory domain observed for Type I importer NBDs such as the nucleotide-binding protein in maltose importer MalK (Diederichs et al., 2000; Chen et al., 2003). The CysA protein also has an N-terminal extension that is not found in bacterial NBDs. This extension is likely to be involved in correct localisation of the protein in the cell to the plastid (Lee & Hwang, 2018), and is predicted by the TargetP 2.0 online bioinformatics tool (Emanuelsson et al., 2007; Almagro Armenteros et al., 2019) to be a leader sequence targeting the protein to the plastid, but not to the lumen of the thylakoids, in agreement with earlier predictions (Melis & Chen, 2005). Similarly, CysT has an N-terminal extension when compared to its bacterial homologs. This also appears to be a plastid-targeting sequence, as predicted by TargetP 2.0 (Almagro Armenteros et al., 2019). CysW is similarly extended at the N-terminus, but in this case it is predicted by TargetP 2.0 to be a signal peptide, and the score for plastid targeting is significantly lower. The SBP CysP does not have a predicted plastid N-terminal targeting sequence, which is interesting. Presumably it must at least deliver its substrate to the outer chloroplast envelope membrane but perhaps does not need to cross the outer membrane itself. Hence for CysW and CysP there remain missing details about the targeting of these proteins to the sulphate importer complex.

How strong is the evidence to support the conjecture that CysT, -W, -A and -P are components of a Type I importer complex? Sequence alignments alone appear to provide support, but unfortunately ABC transporter NBDs are highly conserved and hence may be less diagnostic for one class of ABC transporter versus another (Davidson et al., 2008). Conversely, the TMDs are much more variable in their sequences, and provide stronger discrimination; but here the simple requirement for 20–30 hydrophobic amino acid residues to span the lipid bilayer usually means that sequence identity in homologs of transmembrane proteins can be quite low, especially for polytopic membrane proteins with short extra-membranous loops (Stamm et al., 2013). Furthermore, evolutionary rates of divergence may be higher for membrane proteins compared to cytosolic proteins because of the increased selective pressure of the external environment (Sojo et al., 2016). The length and number of membrane-spanning helices could provide further discriminatory power in addition to sequence homology (Stamm et al., 2013).

The recent explosion in the number of available ABC transporter structures has allowed the number and location of transmembrane spans to be precisely identified for different types of transporter, providing further confidence in the assignment of the CysT, -W, -A, and -P homologs as Type I importers (Ford & Beis, 2019). Homology models are shown in Fig. 3 based on the MalG and MalK proteins. Models based on other Type I importers are available as supporting online information (Figs S1 and S2). CysT homology models can be compared: one based on the MalG structure (with six TM spans, Fig. 3) and two others based on the ModB structure (Fig. S1) and the MetI structure with five TM spans (Fig. S2). The CysT sequence can be readily threaded onto both bacterial Type I importer structures, but with superior correlation of the model generated using MalG. These models imply that the eukaryotic CysT protein has six rather than five transmembrane helices and that both the N- and C-termini are located in the chloroplast stroma. A test of homology models of membrane proteins is whether charged residues appear in the predicted transmembrane portions: for example, in the CysT model (Fig. 3), there is only one charged residue that is in a transmembrane position: R346 which lies in the penultimate TM span (Fig. 3C). However R346, is positioned towards the central cavity of the transporter and may be important for sulphate transport. In the model, R346 appears to be stabilised by interaction with E350 that is on the outer surface. Charge–charge interactions also stabilise the domain interactions in Type I importers, with negatively charged residues in the conserved EAA loop from the TMD interacting with a conserved positively charged residue in the first NBD helix (e.g. R47 in MalK interacts with E190 in MalG). In the models, the equivalent residues are R159 in the NBD and E308 and E311 in the EAA loop of the TMD. In summary, the homology models strongly support the idea that the C. reinhardtii CysT and CysA proteins are components of a Type I ABC importer complex.

In contrast to CysT, the CysW sequence produces a poorer homology model when aligned with MalG. However the models generated using ModB and MetI as templates are relatively good, with nearly all the secondary structural elements reproduced (Fig. S3). This is somewhat surprising since the sequence identity between CysW and MalG is lower than for ModB or MalG. It seems possible that modelling with MetI or ModB as templates may be more appropriate for CysW because of the correspondence between the lengths of the TM spans, coupled with the overall broad homology of amino acid types (hydrophobic versus polar and charged). Although the CysW model with MetI as a template displays three charged residues in the middle of the membrane span, two of these (R250, D194) are paired, but R300 is unpaired, and in the model protrudes into the central cavity of the membrane-spanning region. Interestingly, MetI also shows an unpaired arginine residue (R64) that protrudes into the central cavity, but from a different TM helix (Kadaba et al., 2008; Johnson et al., 2012; Nguyen et al., 2018). Hence it seems possible that the CysT, -W, -A, or -P complex may be composed of two TMDs with six TM spans for CysT and five or six TM spans for CysW.

(3) Evidence for Type I ABC importers in other eukaryotes

An early bioinformatics search (Melis & Chen, 2005) did not find any CysT or CysW homologs in higher plants. A similar search of the much-expanded modern sequence databases still reveals no convincing evidence for any Type I importers in higher plants, nor in any non-autotrophic eukaryotes. Some primitive plants such as hornworts and liverworts appear to have copies of the sulphate importer proteins,
and the CysA gene has been studied in liverworts (Wickett et al., 2011) with the conclusion that this transporter gene was being lost at a high rate from this family. However in species that retain the gene, it was still undergoing purifying selection (Wickett et al., 2011). The implication from these various bioinformatics analyses is that at least one Type I importer has moved from early prokaryotes (most likely via symbiotic cyanobacteria) and is present in extant algae and primitive plants (Margulis, 1971). However this type of importer does not appear to have been carried through into higher plants. Type I importers that may have been present in early mitochondria arising by symbiosis thus do not appear to have been retained in other modern eukaryotes. Potential reasons for these dead-ends in algae and non-vascular plants will be discussed in Section IV.

III. TYPE II ABC IMPORTERS

(1) Bacterial Type II ABC importers

These importers generally transport substances that are essential co-factors in biological processes in the cell (Davidson et al., 2008; ter Beck et al., 2014). For example vitamins and co-enzymes are transported by bacterial Type II ABC importers as well as transition metals or ions that are essential for metabolism such as Fe$^{3+}$, Zn$^{2+}$ and Mn$^{2+}$. The latter are usually imported as part of a siderophore complex. There is some overlap between Type II importers and the ECF importers in terms of their imported substrates (ter Beck et al., 2014).

The first high-resolution structure of an ABC transporter was obtained for the vitamin B12 importer BtuCD (Locher et al., 2002), which is an example of a Type II importer (see Fig. 1). BtuC is the TMD and BtuD the NBD. The SBP, BtuF was studied as part of the importer complex in later work (Hvorup et al., 2007; Korkhov, Mireku & Locher, 2012). Ten short membrane-spanning α-helices compose the TMD of these importers, although there is no swapping of TM helices from one TMD to the other (Fig. 2). There is much sequence variability for the N-terminal TM helix and first extracellular loop, although all the bacterial homologs of proteins with defined structures appear to have a hydrophobic region of ~20 residues at the N-terminal end, implying that the topology of 10 TM spans is well conserved. The available structures show very close structural similarity for the nucleotide-free state (Locher et al., 2002; Pinkett et al., 2007; Woo et al., 2012; Naoe et al., 2016). The TMDs have four cytoplasmic loops, three of which are very short. The longest loop (the third in the sequence) contacts the underlying NBD via the Q loop region, in a way that is very similar to Type I importers. Unusually for ABC transporters, some of the external loops are long and the third external loop dips back into the membrane-spanning region of the domain. The NBD is, as expected, very typical of the overall ABC family. Perhaps the only distinguishing feature in the deposited structures is a short C-terminal extension of ~30 residues that forms two antiparallel β-strands. This extension may be important in maintaining NBD–NBD contact; the C-terminal β-strands from each NBD can interact to form a loose four-stranded anti-parallel β-sheet in the assembled complex. The transition from the nucleotide-free to nucleotide-bound state can be compared for BtuCD where structures for both states exist (Korkhov et al., 2012) [e.g. Protein Data Bank (PDB) codes 4DBL (nucleotide free) to 4R9U (nucleotide bound)]. A rocking motion brings together the two NBDs to form the sandwich dimer configuration with the nucleotide cooperatively bound (Ford & Hellmich, 2020). This is associated with squeezing together of the two cytoplasmic loops of the TMDs and rearrangement of the TM helices to create a central cavity that spans the membrane. The presence of the SBP seals the outside of this cavity, whilst on the inside a few residues at the start of the third TM span seal the cytoplasmic exit. This plug is in a non-helical/kinked region of the TMD and presumably further conformational change is needed to allow exit of the substrate, perhaps associated with ATP hydrolysis (Korkhov et al., 2012). This is the opposite situation to that seen in ABC exporters where ATP binding is associated with the formation of an outward-facing central cavity (Rosenberg et al., 1997; Dawson, Hollenstein & Locher, 2007), hence the structural information seems to be consistent with idea of ATP-driven transport via an alternating access mechanism, as proposed long ago by Jardetzky (1966).

(2) Type II ABC importers in eukaryotes

For Type II importers, little has been published on systematic searches of eukaryotic sequence databases. However a PSI-BLAST (Altschul et al., 1990) search using bacterial TMDs of this type was able to identify a single eukaryotic protein as a hit in the Uniprot database (YCXC_CYAPA). This is a transmembrane protein found in the Ye23 – apcF intergenic region of the primitive alga Cyanophora paradoxa (Giddings, Wasmann & Stachelin, 1983; Pittenauer et al., 1993; Price et al., 2019). The same region of the genome also contains a sequence corresponding to an ABC NBD (YCXD_CYAPA). C. paradoxa is considered to be a primitive alga with features reminiscent of cyanobacteria. For example, the Cyanophora plastid retains a remnant peptidoglycan wall (Giddings et al., 1983; Pittenauer et al., 1993; Price et al., 2019).

As described above for Type I importers, 3D homology models can used to review the fidelity of the assignment of the C. paradoxa proteins to Type II importers. Models of the C. paradoxa transporter proteins (Fig. 4) testify to the close correspondence between their sequences and that of the nucleotide-free and outward-facing 3D structure of the bacterial heme transporter proteins HmuUV. An equi-covalent alignment for the first TM span (red-coloured ribbon, Fig. 4C) is reflected in the poor modelling of this region and seems discordant with the idea that the N-terminal 30 residues may constitute a signal peptide rather than a transmembrane
span. Hence it seems probable that the *C. paradoxa* TMD has nine rather than 10 TM spans, with the N- and C-termini on opposite sides of the membrane. Similarly, the C-terminal 10 residues in the *C. paradoxa* TMD (yellow-coloured ribbon, Fig. 4C) extend beyond the alignment with HmuU and hence are modelled as unstructured. As discussed above, a good test of the validity of a homology model for a membrane protein is the absence of charged residues from the membrane-spanning portions of the TMD. For the *C. paradoxa* TMD, the model meets this criterion except for two residues, D64 and R58 that are modelled into the unusual kinked start of TM span 3 which forms the inner cytoplasmic gate in the ATP-bound structure (see Fig. 4C and discussion above for BtuCD). It seems possible that D64 may be important for interaction with positively charged substrates such as Mn$^{2+}$. In the *C. paradoxa* model, D64 and R58 are close enough to allow charge–charge interactions. TMD–NBD interactions in the homology model should also be consistent with the Type II ABC transporter domain organisation. In the model, the cytoplasmic loop connecting the TMD to the NBD is stabilised by a salt-bridge between D183 in the TMD and R85 in the NBD. R85 is adjacent to the conserved glutamine Q84, which is part of the NBD ‘Q loop’ region (Wen & Tajkhorshid, 2011; Zolnerricks et al., 2014). The outward-facing central pathway displayed by the model is not lined with negatively charged residues, as might be predicted for a cationic substrate such as Mn$^{2+}$. However, in the model generated for the outward-facing nucleotide-free state, the cytoplasmic/predicted stromal surface of the TMDs has a central channel plugged by two positively charged lysine residues from each TMD (K109).

**(3) Other Type II ABC importers in eukaryotes**

On its own, the *C. paradoxa* protein and organism might be discounted as an ancient relic, but PSI-BLAST searches using the *C. paradoxa* transmembrane protein sequence against non-redundant protein sequences identified many more hits, the most homologous corresponding to predicted transmembrane proteins in red algae. These were annotated by the database as possible manganese permeases based on homology with bacterial sequences (Altschul et al., 1990). Caution must be exercised, however, when searching such databases; for example, there are also some spurious hits in this (less curated) non-redundant protein sequence database. These were mostly observed in insect genomes and are highly likely to correspond to contamination by bacterial sequences (Steinegger & Salzberg, 2020). For example the hit found in the deposited genome of the black garden ant (*Lasius niger*) showed 100% sequence homology to a Type II importer TMD from the bacterial pathogen *Serratia marescens*. This species is known to infect laboratory-raised insects (Inglis & Lawrence, 2001).

![Fig. 4. Homology models of target proteins, the transmembrane protein (YCXC_CYAPA) and the ATP-binding protein (YCXD_CYAPA) in the ycf23-apcF intergenic region of *Cyanophora paradoxa*, intersecting with their homologous template proteins, the hemin transport system transmembrane protein HmuU and the nucleotide-binding protein HmuV respectively derived from the crystal structure of the Type II importer HmuUV from *Yersinia pestis* (PDB ID: 4G1U). The view is from along the membrane plane.](image)
A BLAST search of the putative NBD from the *C. paradoxa* genome against the Uniprot database generated many hits across a wide range of organisms and ABC transporter types, as expected for this more conserved domain. However such searches show that the closest homologs in the protein data-bank are NBDs of Type II importers (e.g. HmuV and a metal chelate importer). As mentioned above, these have a short (~30 residue) C-terminal extension to the NBD that, in the available structures, appears to be involved in NBD–NBD contact. This region does not appear to be fully present in the *C. paradoxa* NBD (Fig. 4C). A similar PSI BLAST search of the *C. paradoxa* NBD sequence against non-redundant protein sequences showed that the closest homologs were in the red algae as well as in cyanobacteria. In contrast to the Type I importers, however, there was no strong prediction of a plastid-targeting N-terminal sequence for the identified proteins.

### IV. ROLES OF EUKARYOTIC TYPE I AND II ABC IMPORTERS

The existence of Type I and Type II importers in very early eukaryotic organelles would be a reasonable assumption, given the symbiotic theory for the origins of eukaryotes (Margulis, 1971). Hence, the lack of any PSI BLAST search hits in the genomes of non-autotrophic eukaryotes is unexpected and points to rapid elimination of these importers after symbiosis began (Margulis & Bermudez, 1985). An alternative explanation could be that homologs of Type I or Type II importers do exist, but that the sequence homology of the TMDs is too weak to identify them by PSI BLAST searches (Altschul *et al.*, 1990). For example, a PSI BLAST search using MalG as a query and the PDB sequences as a database identified MalF (Oldham *et al.*, 2007), AlgM2 (the TMD heterodimer in the alginate ABC importer) (Maruyama *et al.*, 2015) and MetI (Kadaba *et al.*, 2008) as homologs, but did not identify sufficiently high sequence homology to select ModB, nor the ModB importer TMD from *Methanosarcina acetivorans* (both of which are also Type I importers). If rapid mutation and divergence occurred during the development of eukaryotic ABC importer TMDs, then perhaps instead one might hope to find sequences corresponding to their NBDs. For example the additional regulatory domain of the Type I NBD may facilitate identification of eukaryotic Type I importers. However comparison of the available Type I ABC importer structures shows that the presence and structure of the additional C-terminal domain is quite variable. The additional domains of MalK (Oldham *et al.*, 2007), AlgM2 (Maruyama *et al.*, 2015) and the ModC molybdate/sulphate importer from *Methanosarcina acetivorans* (Gerber *et al.*, 2008) have very similar folds [<2.0 A root mean square deviation (rmsd) over ~100 backbone atoms]. However the MetN C-terminal domain has an entirely different fold (Kadaba *et al.*, 2008), whilst the ModC structure (2ONK) has only a short (~10 residues) helical extension (Hollenstein, Dawson & Locher, 2007a). Hence from the published structures, it would appear that the additional C-terminal domain may be dispensable. It is perhaps not surprising, therefore, that there is no current evidence to support the presence of Type I importer-like NBD extensions in higher plants and non-autotrophic eukaryotes.

We may therefore consider why Type I and Type II ABC importers only occur in algae and primitive plants. One might postulate that these importers became irrelevant in some organisms and hence were slowly lost, as implied from a study of the liverwort family (Wickett *et al.*, 2011). Some clues about the requirement for Type I ABC importers in algal plastids can be gleaned from the sets of species that were identified by bioinformatics. All CysT homologs derived from bacterial Type I importers are in algal species – green algae and soil algae, with either freshwater or terrestrial habitats (Melis & Chen, 2005). These include, among the *Helicosporidium* genus, the green algal subspecies *Simulium jonesii* which is a parasite inhabiting the gut of insects. A close relative of *Helicosporidium*, *Prototheca*, was also identified as having a CysT homolog, specifically *Prototheca wickerhamii* which can cause skin infections in humans (Melis & Chen, 2005).

It has been shown that low abundance of sulphate can be growth-limiting in freshwater cyanobacteria, and in some cases there has been selection for very abundant phycobiliproteins with reduced numbers of sulphur-containing amino acids (Mazel & Marlierre, 1989). Since the chloroplast is the principal site of amino acid synthesis in algae, and sulphur is necessary for the production of sulphur-containing amino acids, then it would appear to make sense to have a mechanism for active transport of sulphate into the chloroplast. It is possible that higher plants may have been able to dispense with these ATP-consuming chloroplast importers due to the relative abundance of sulphate in soil. For example a plastid-located sulphate transporter, SULTR3;1 has been identified in the model plant *Arabidopsis thaliana* (Cao *et al.*, 2013). Similar sulphate transporters are present in plant roots. Within the PDB, these sulphate transporters show strong structural homology to bacterial low-affinity sulphate transporters and also to the anion exchanger solute carrier family 26 (e.g. member 9, SLC26a9) (Chi *et al.*, 2020). Hence it seems feasible that less-energy-expensive transporters took over the role of sulphate import into the chloroplast as higher plants developed.

For the Type II importers that are found in the sequenced genomes of red algae, their function has been tentatively annotated as Zn$^{2+}$/Mn$^{2+}$ importers, based on homology to the ZnuB/ABC-3 protein family by InterPro at the European Bioinformatics Institute (Mitchell *et al.*, 2015). Again, it may be the case that manganese or zinc import against a concentration gradient became less important in increasingly developed plants that adapted to environments where these elements were not limiting. Manganese is essential in the chloroplast thylakoids for oxygenic photosynthesis, being part of the oxygen-evolving complex of photosystem II (Radmer & Kok, 1975; Barber & Murray, 2008). The chloroplast stroma also contains Mn$^{2+}$-dependent enzymes.
Occasionally Mn$^{2+}$ availability becomes limiting in land plants, and several transport processes involved in its distribution and storage (in the vacuole) have been characterised (Alejandro et al., 2020). Uptake of Mn$^{2+}$ into higher plant chloroplasts and the thylakoid lumen (where it is needed for oxygen evolution) is via H$^{+}$/Mn$^{2+}$ exchangers (Alejandro et al., 2020).

V. PROPOSED NAMING OF EUKARYOTIC TYPE I AND TYPE II IMPORTERS

Figure 5 provides an overview of the transit of prokaryotic ABC transporters into eukaryotic organisms, with the new naming system proposed by Thomas et al. (2020) mapped onto the eukaryotic alphabetical categorisation. We propose the extension of the latter categorisation system, with ABCJ and ABCK used for the eukaryotic homologs of Type I and Type II bacterial importers, respectively. However readers should note that ABCJ was also recently proposed as a categorisation of an NBD-only group of ABC proteins in mosquitoes (Figueira-Mansur et al., 2020). The relative merits of these proposals will be subject to the usual scrutiny of the scientific community.

VI. CONCLUSIONS

(1) Eukaryotic homologs of Type I bacterial ABC importers appear to be found in algae and the liverwort family of primitive non-vascular plants; but not in higher plants.

(2) Eukaryotic homologs of Type II bacterial ABC importers appear to be limited to red algae and glaucophytes such as *C. paradoxa*.

(3) The roles proposed for Type I and Type II importers in eukaryotes are in agreement with the general observation that Type I importers transport substrates essential for incorporation into the biomolecules of the organism (e.g. sulphate) and that Type II importers deliver crucial co-enzymes or co-factors (e.g. manganese and/or zinc ions).

(4) We propose that, according to the naming system of Thomas et al. (2020), ABCJ and ABCK are used for the eukaryotic homologs of Type I and Type II bacterial importers, respectively.

(5) Prokaryotic ABC importers are vitally important for bacterial growth and survival, and they have a significant impact on humans in terms of pathogenesis. We suggest that the Type I and II importers found in eukaryotes are equally important. Although these
VIII. REFERENCES

Alejandro, S., Holler, S., Meier, B. & Peiter, E. (2020). Manganese in plants: from acquisition to subcellular localization. Frontiers in Plant Science 11, 300.

Almagro Armenteros, J. J., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A. & Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. Life Science Alliance 2(5), e201900429.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology 215(3), 403–410.

Bar-On, Y. M., Phillips, R. & Milo, R. (2018). The biomass distribution on earth. Proceedings of the National Academy of Sciences of the United States of America 115(25), 6506–6511.

Barber, J. & Murray, J. W. (2008). The structure of the Mn4Ca2 cluster of photosystem II and its protein environment as revealed by X-ray crystallography. Molecular Cell 317(2), 409–419.

Berg, M. D. (2012). How many species of algae are there? Journal of Phycology 48(5), 1037–1063.

Guo, Z., Kang, S., Zhu, X., Xia, J., Wu, Q., Wang, S., Xie, W. & Zhang, Y. (2015). The novel ABC transporter ABC1H is a potential target for RNAi-based pest control and resistance management. Science Reports 5, 13728.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. Annual Review of Cell Biology 6, 67–113.

Higgins, C. F. (2001). ABC transporters: physiology, structure and mechanism—an overview. Research in Microbiology 152(4–5), 203–210.

Hollander, I. B. (2011). ABC transporters, mechanisms and biology: an overview. Antonie van Leeuwenhoek 99(4), 1–17.

Holmén, A. L., Hollenstein, K. R., Schuetz, J. P. & Locher, K. P. (2007a). Structure and mechanism of ABC transporter proteins. Current Opinion in Structural Biology 17(4), 412–418.

Holmén, K. F., Frei, D. C. & Locher, K. P. (2007b). Structure of an ABC transporter complex in eukaryote. Nature 447(7132), 213–216.

Inglis, G. D. & Lawrence, A. M. (2001). Effects of Saccharomyces cerevisiae on the F1 generation of laboratory-reared Heliothis virescens (Lepidoptera: Noctuidae). Journal of Economic Entomology 94(2), 362–366.

Jardetzky, O. (1966). Simple allosteric model for membrane pumps. Nature 211 (5052), 969–970.

Johnson, E., Nguyen, P. T., Yeates, T. O. & Rees, D. C. (2012). Inward facing conformations of the Mst2 linehomine ABC transporter: implications for the mechanism of translocation. Protein Science 21(1), 84–96.

Kadara, N. S., Kaiser, J. T., Johnson, E., Lee, A. & Rees, D. C. (2008). The high-affinity E.coli membrane ABC transporter: structure and allosteric regulation. Science 319(5868), 250–253.

Kang, J., Park, J., Choi, H., Buria, B., Kretzschmar, T., Lee, Y. & Martinou, E. (2011). Plant ABC transporters. Anabiosis 20, 41-53.

Khare, D., Oldham, M. L., Oreille, C., Davidson, A. L. & Chen, J. (2009). Alternating access in maltose transporter mediated by rigid-body rotations. Molecular Cell 33(4), 329–336.

Korovou, V. M., Mikkelsen, S. A. & Locher, K. P. (2012). Structure of AMP-PNP-bound variant B12 transporter BucDF. Nature 490(7420), 367–372.

Laudenbach, D. E. & Grossman, A. R. (1991). Characterization and mutagenesis of sulfate-regulated genes in a cyanobacterium: evidence for function in sulfate transport. Journal of Bacteriology 173(9), 2739–2740.
Eukaryotic ABC importers

Lee, D. W. & Hwang, I. (2018). Evolution and design principles of the diverse chloroplast transit peptides. *Molecules and Cells* 41(3), 161–167.

Linton, K. J. (2007). Structure and function of ABC transporters. *Physiology (Bethesda)* 22, 122–130.

Liu, X. (2019). ABC family transporters. *Advances in Experimental Medicine and Biology* 1141, 15–100.

Locher, K. P., Lee, A. T. & Rees, D. C. (2002). The C. elo BrocBD structure: a framework for ABC transporter architecture and mechanism. *Science* 296(5570), 1099–1108.

Margulis, L. (1971). Symbiosis and evolution. *Science* 172(3989), 401–409.

Mariya, Y., Itoh, T., Kaneko, A., Nishitani, Y., Mikami, B., Hashimoto, Y., & Moteki, K. (2020). Protein composition of the nuclear ABT transport involved in the import of an acidic polycationic alginic acid. *Structure* 28(9), 1643–1654.

Mazzol, D. & Marliere, P. (1989). Adaptor recognition of methionine and cysteine from cytoplasmic bacterial light-harvesting proteins. *Nature* 341(6239), 245–248.

Mitchell, A., Chen, L. & Chen, J. (2005). Chloroplasts of the marine transporter during ATP hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America* 102(37), 15132–15136.

Mitchell, A. L. & Chen, J. (2018). Crystal structure of the catalytic intermediate of the maltose transporter. *Nature* 540(7635), 515–521.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, C. E. & Ferrin, T. E. (2004). UCSF chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25(13), 1605–1612.

Pinkett, H. W., Lee, A. T., Lum, P., Locher, K. P. & Rees, D. C. (2007). An inward-facing conformation of a putative metal-chelator-type ABC transporter. *Science* 315(5810), 373–377.

Pittau, J. M., Cher, D., Quicho, F. A., Davidson, A. L. & Chen, J. (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450(7169), 515–521.

Pittenger, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, C. E. & Ferrin, T. E. (2004). UCSF chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25(13), 1605–1612.

Price, D. C., Goodfellow, U. W., Roth, R., Lee, J. H., Karayiawasam, T., Muthuk, M., Ferrari, C., Facchinelli, F., Ball, S. G., Cinci, U., Chan, C. X., Wagner, N. E., Yoon, H. S., Weber, A. P. M. & Bhattacharya, D. (2019). Analysis of an improved *Cyanophora paradoxa* genome assembly. *DNA Research* 26(4), 367–377.

Quazi, F., Lenevich, S. & Molnay, R. S. (2012). ABCA1 is an N-erythroidine-phosphatidylcholine and phosphatidylethanolamine importer. *Nature Communication* 3, 925.

Raimer, R. & Kor, B. (1975). Energy capture in photosynthetic photosystem II. *Annual Review of Biochemistry* 44, 409–433.

Rice, A. J., Park, A. & Pinkett, H. W. (2014). Diversity in ABC transporters type I and II. *Cellular and Molecular Life Science* 71(6), 935–947.

Rosenberg, M. F., Callaghan, R., Ford, R. C. & Higgins, C. F. (1997). Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis. *Journal of Biological Chemistry* 272(16), 10685–10694.

Slotboom, D. J. (2014). Structural and mechanistic insights into prokaryotic energy-coupling factor transporters. *Nature Reviews in Microbiology* 12(2), 79–87.

Sojo, V., Desimmoz, C., Pomiankowski, A. & Lane, N. (2016). Membrane proteins are dramatically less conserved than water-soluble proteins across the tree of life. *Molecular Biology and Evolution* 33(1), 2074–2084.

Stamm, M., Sturbezicher, R., Khiatizov, K. & Forrest, L. R. (2013). Alignment of helical membrane protein sequences using AlignMe. *Public Library of Science* 8(3), e57731.

Steenegger, M. & Salzberg, S. L. (2020). Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biology* 21(4), 115.

ter Reek, J., Guskov, A. & Slotboom, D. J. (2014). Structural diversity of ABC transporters. *Journal of General Physiology* 143(4), 419–435.

Van der Westhuizen, A. & Reis, A. J. (2019). ABC family transporters. *Annual Review of Physiology: Eukaryotic ABC importers* 1329.

van den Berg, B., Kolk, H. J., Van den Berg, H. B., Van den Berg, M. A. & Van den Berg, M. W. (2011). The InterProteome database: a classification resource after 15 years. *Nucleic Acids Research* 43(Database issue), D13–D21.

Voth von Voitenherg, L., Park, J., Stube, R., Lux, C., Lee, Y. & Philipp, K. (2019). A novel prokaryote-type ECF/ABC transporter module in chloroplast metal homeostasis. *Frontiers in Plant Science* 10, 1284.

Wang, P. & Gress-Reis, T. (2011). Structural and mechanistic insights into prokaryotic energy-coupling and the transmembrane domains in ABC transporters. *Biophysics Journal* 101(3), 680–690.

Wickett, N. J., Forrest, L. L., Bode, J. M., Shaw, B. & Goffinet, B. (2011). Frequent pseudogenization and loss of the plastid-encoded sulfate-transport gene distribution through the evolution of liverworts. *American Journal of Botany* 98(9), 1263–1273.

Woo, J. S., Zeltina, A., Goetz, B. A. & Locher, K. P. (2012). X-ray structure of the Yersinia pestis heme transporter HimU-HimV. *Nature Structural & Molecular Biology* 19(12), 1310–1315.

Yates, A. D., Achutah, P., Aranbi, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Arzam, I. A., Azov, A. G., Bennett, R., Bhai, J., Bellin, K., Boddou, S., Marugan, J. C., Cummins, C. E., et al. (2020). *Essentvlo 2020*. *Nucleic Acids Research* 48(D1), D602–D608.

Zhu, B., Aller, S. G., Akka, B. G., Sasse, M., Griba, P., Sekig, A. & Linton, K. J. (2014). The Q-loops of the human multidrug resistance transporter ABCB1 are necessary to couple drug binding to the ATP catalytic cycle. *FASEB Journal* 28(10), 4335–4346.

IX. Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article. *Fig. S1*. Homology models of target proteins, the sulphate transport system transmembrane protein CysT and the nucleotide-binding protein CysA from *Chlamydomonas reinhardtii*, intersecting with their template proteins, the molbdenum transport system transmembrane protein ModB and the nucleotide-binding protein ModC respectively, derived from the crystal structure of the Type I importer ModBC from *Arabidopsis thaliana* (PDB ID: 2ONK). The view is from along the membrane plane. (A) Transmembrane domain (TMD) of CysT (Uniprot ID: Q8RVC7) is modelled in lilac. Percentage sequence identity (%ID) of CysT and ModB is 25%. (B) Nucleotide-binding domain (NBD) of CysA (Uniprot ID: Q6QJE1) is modelled in light blue. %ID of CysA and ModC is 38.8%. (C) The two homology models for the TMD of CysT and NBD of Cys A displayed separately. The green/blue coloured atoms highlight residues R346 and E350 in the CysT protein. See Fig. 3 for methods.
**Fig. S2.** Homology models of target proteins, the sulphate transport system transmembrane protein CysT and the nucleotide-binding protein CysA from *Chlamydomonas reinhardtii*, intersecting with their template proteins, the D-methionine transport system transmembrane protein MetI and the nucleotide-binding protein MetN respectively, derived from the crystal structure of the Type I importer MetNI from *Escherichia coli* (PDB ID: 6CVL). The view is from along the membrane plane. (A) Transmembrane domain (TMD) of CysT (Uniprot ID: Q8RVC7) model is coloured in lilac. Percentage sequence identity (%ID) of CysT and MetI is 24.7%. (B) Nucleotide-binding domain (NBD) of CysA (Uniprot ID: Q6QJE1) model is coloured in light blue. %ID of CysA and MetN is 28.8%. (C) The two homology models for the TMD of CysT and the NBD of CysA displayed separately. See Fig. 3 for methods.

**Fig. S3.** Homology models of target transmembrane domain (TMD) protein CysW from the sulphate transport system of *Chlamydomonas reinhardtii* (Uniprot ID: Q6QJE2) intersecting with three different template proteins all derived from the crystal structure of Type I importers: (A) MalG transmembrane protein subunit from maltose transporter MalFGK2 (PDB ID: 4JBW); (B) ModB transmembrane protein subunit from molybdate ABC transporter ModBC (PDB ID: 2ONK); (C) MetI transmembrane protein subunit from methionine transporter MetNI (PDB ID: 6CVL). The view is from along the membrane plane. All CysW proteins are coloured in light blue. Percentage sequence identity (%ID) of CysW with MalG, ModB and MetI is 22%, 24.6% and 16.7%, respectively. (D) Individual CysW models derived from MalG, ModB and MetI top to bottom. The green/blue/red-coloured atoms highlight the interactions between charged residues D194 and R69 in the MalG-derived model, E304 and R300 in the ModB-derived model, and a pair of R250 and D194 and an unpaired R300 in the middle of the membrane spanning region of the MetI-derived model. See Fig. 3 for methods.

(Received 14 September 2020; revised 21 February 2021; accepted 23 February 2021; published online 2 March 2021)