Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Transport protein evolution deduced from analysis of sequence, topology and structure

Milton H Saier Jr.

The vast majority of well studied transmembrane channels, secondary carriers, primary active transporters and group translocators are believed to have arisen vis intragenic duplication events from simple channel-forming peptides with just 1-3 transmembrane α-helical segments, found ubiquitously in nature. Only a few established channel-forming proteins appear to have evolved via other pathways. The proposed pathway for the evolutionary appearance of the five types of transport proteins involved intragenic duplication of transmembrane pore-forming peptide-encoding genes, giving rise to channel proteins. These gave rise to single protein secondary carriers which upon superimposition of addition protein domains and proteins, including energy-coupling proteins and extracytoplasmic receptors, gave rise to multidomain, multicomponent carriers, primary active transporters and group translocators. Some of the largest and best characterized superfamilies of these transmembrane transport proteins are discussed from topological and evolutionary standpoints.

Address
Department of Molecular Biology, University of California at San Diego, La Jolla, CA 92093-0116, USA

Corresponding author: Saier, Milton H (msaier@ucsd.edu)

Current Opinion in Structural Biology 2016, **38**:17–25

This review comes from a themed issue on **Sequences and topology**

Edited by **Aravind Iyer** and **Elizabeth Meiering**

http://dx.doi.org/10.1016/j.sbi.2016.05.001

Introduction

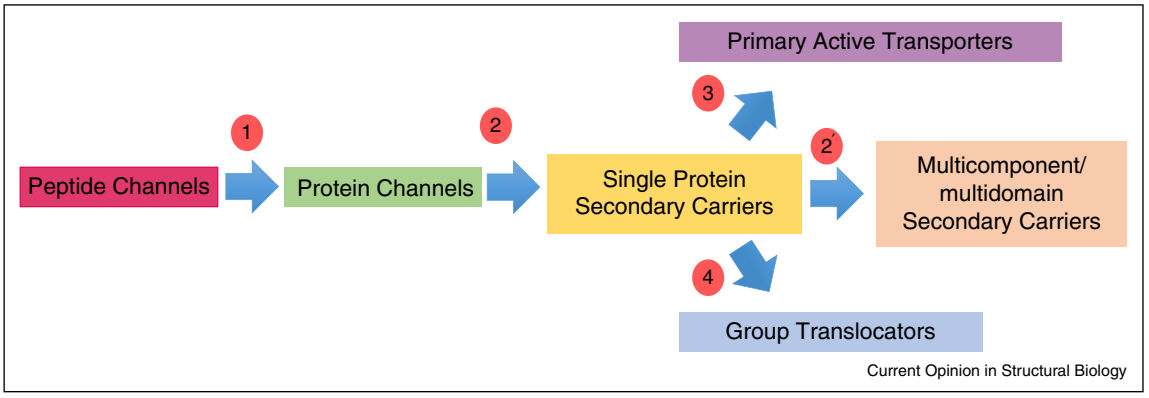
Transmembrane transport systems are vital for all aspects of cellular physiology including uptake and export of nutrients, end products of metabolism, drugs, toxins and macromolecules [1,2,3,4]. Our laboratory has studied these proteins for several decades, and have developed and maintain the IUBMB-approved Transporter Classification Database (TCDB; www.tcdb.org) [5,6,7,8]. TCDB provides a functional/phylogenetic system of classification with four well-defined classes, (1) Channels, (2) Secondary Carriers, (3) Primary Active Transporters and (4) Group Translocators. The first class includes both small oligomeric peptide channels and larger protein channels. The second class includes both single protein and multidomain, multicomponent, secondary carriers with the precursor channel protein providing the transport pathway. Primary active transporters and group translocators are usually multi-component systems in which the energy-coupling proteins are superimposed on the transporters. Only group translocators modify their substrates during transport, and these also usually require the participation of enzymes superimposed upon and mechanistically coupled to the transporters. However in a few recognized cases, enzymes have become transmembrane, gaining transport functions, and in other cases, transporters have evolved enzyme catalytic activities [9••]. Thus, we believe that most transporters evolved in a sequential fashion from class 1 through class 2 to classes 3 and 4 via the pathway shown in Figure 1 [42••]. This pathway and the functional diversity exhibited by members of a few large transport protein superfamilies will be the focus of this article [10•,11•].

In addition to α-type integral membrane transport proteins, transmembrane β-barrel proteins form a distinct class of channel proteins, called porins or outer membrane pore-forming proteins (OMPPs). They form channels in the outer membranes of many Gram-negative and Gram-positive bacteria, mitochondria, chloroplasts and certain other eukaryotic organelles such as peroxisomes [12,13]. The family and superfamily relationships of these OMPPs and the evolutionary pathways that gave rise to them have been considered in detail [14,15••] and will not be discussed further here.

The precursors of most α-helical type transporters: channel-forming peptides

Virtually all types of organisms synthesize peptides designed to insert into membranes of either self or non-self to create oligomeric ion-conducting pores. None of these small peptides form carriers or more complex transporters. Many families and superfamilies of peptide pore-formers are known and recorded in TCDB. They include bacteriocins and archaeocins, made by bacteria and archaea, respectively. Many bacteriocins function to attack other organisms (>12 families in TC subclass 1.C) [16,17,18]. Eukaryotic 'defensins' and 'cecropins' (>12 families, also in subclass 1.C) serve as the first line of bodily defense, destroying infectious agents of disease such as envelope viruses, bacteria, fungi and parasites [19]. Ubiquitous organismal holins, 'hole-formers', (TC subclass 1.E) insert into the cytoplasmic membranes of

Figure 1



Proposed pathway for the evolution of most transport proteins found in living organisms. Genes encoding simple oligomeric pore-forming peptides (red) underwent intragenic duplication, triplication or quadruplication as well as gene fusional events to add auxiliary domains, generating protein channels (green) with fewer subunits (step 1). These then mutated to achieve stereospecific recognition of solutes as well as an 'alternating access' transport mechanism by which the substrate binding site can shuffle between an outwardly open conformation and an inwardly open conformation with one or more intermediate occluded states, yielding a single protein secondary carrier (yellow) (step 2). In some cases, extra domains (for regulation and protein-protein interactions) and extra subunits (for substrate recognition, biogenesis and facilitation of catalysis) were added (step 2′) (tan). Finally, by the superimposition of energy coupling proteins (often enzymes) onto the carrier, primary active transporters (purple) and group translocating porters (blue) evolved (steps 3 and 4, respectively). In some cases, alternative pathways may have been followed, as for light-driven microbial rhodopsins and certain ion-translocating electron transfer proteins, where the primary active transport mechanism did not result from the superimposition of enzymes upon the carrier, but instead involved incorporation of the energy-coupling mechanism into the transport protein or vice versa. Such alternative events provided a much less common mechanism for generating primary active transporters and group translocators. Circles enclosing numbers indicate evolutionary steps leading to transporter types of the TC class, also indicated by the number.

the producing organisms, where they allow phage-induced cell lysis, programmed cell death, toxin release or development, depending on the system (>60 families) [20•]. Two holin families (TC#'s 1.E.18 and 1.E.36) have been found to contain both 2 and 4 transmembrane segment (TMS) homologues, and the latter have arisen from the former by intragenic duplication [21]. Viroporins and viral fusion pore-forming proteins (in TC subclasses 1.A and 1.G, respectively) (>40 families) act at various stages during the virus infection cycle. It is clear that such transmembrane peptides are found ubiquitously throughout living organisms and could have provided the starting materials for the generation of more complex transport systems.

Intragenic multiplication to generate large protein channels and pores

Many α-type channel proteins (TC subclass 1.A) and β-barrel porins (TC subclass 1.B) have apparently arisen via intragenic duplication events as well as gene fusion events [11•,12]. The best characterized of the former include members of the recently expanded Voltage-gated Ion Channel (VIC) Superfamily.

The voltage-gated ion channel (VIC) superfamily

Figure 2 shows our current conception of how the VIC Superfamily evolved to its present level of complexity. Some members (TC# 1.A.1 and 1.A.2) are small 2 TMS proteins, each with a central reentrant 'Pore' or 'P'-loop, and these small proteins form homo- or hetero-tetrameric channels with a total of 8 TMSs. Another family, the voltage-gated proton channel (VPC) family (TC# 1.A.51), consists of proteins with 4 TMSs that serve both as voltage sensors and H^+ channels [22]. Such 4 TMS domains evidently fused N-terminal to the 2 TMS VIC channels to yield 6 TMS proteins. The N-terminal 4 TMS domain confers voltage sensitivity to the C-terminal 2 TMS domain with central P-loop which retains its ion conductivity function. Although the sensor domains in the 6 TMS VIC Family proteins have lost their H^+ channel properties, a few amino acid substitutions in this 4 TMS domain restore primordial proton channel function [23].

Most VIC K^+ channels consist of tetramers of the basic 2 or 6 TMS subunit, but some of these channels have duplicated the 2 TMS element to give 4 TMS proteins that must dimerize to form a functional channel [24] (Figure 2). Animal Ca^{2+} channels are sometimes duplicated to give 12 TMS proteins, but most Ca^{2+} channels and all Na^+ channels of animals (but not prokaryotes) have duplicated the 12 TMS dimers to give full length 24 TMS monomeric channels [25]. We suggest that in animals, K^+ channels gave rise to Ca^{2+} channels which were the direct precursors of Na^+ channels. Further, several different fusional events have given rise to the large 6 or more TMS proteins of the Ryanodine/Inositol