Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

A topologically diverse family of fluoride channels

Christian B Macdonald[1] and Randy B Stockbridge[1,2]

Dual-topology proteins are likely evolutionary antecedents to a common motif in membrane protein structures, the inverted repeat. A family of fluoride channels, the Flucs, which protect microorganisms, fungi, and plants against cytoplasmic fluoride accumulation, has representatives of all topologies along this evolutionary trajectory, including dual-topology homodimers, antiparallel heterodimers, and, in eukaryotes, fused two-domain proteins with an inverted repeat motif. Recent high-resolution crystal structures of dual-topology homodimers, coupled with extensive functional information about both the homodimers and two-domain Flucs, provide a case study of the co-evolution of fold and function.

Addresses
[1] Program in Biophysics, University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109, USA
[2] Department of Molecular, Cellular, and Developmental Biology, University of Michigan, 930 N. University Avenue, Ann Arbor, MI 48109, USA

Corresponding author: Stockbridge, Randy B (stockbr@umich.edu)

Current Opinion in Structural Biology 2017, **45**:142–149

This review comes from a themed issue on **Membranes**

Edited by **Martin Caffrey** and **David Drew**

http://dx.doi.org/10.1016/j.sbi.2017.04.003

Introduction

The advent of the structural era in membrane protein biology revealed that many membrane transport proteins, built on common folds in classes as diverse as LeuT, CLC, MFS, aquaporin, NhaA, and GltPh, are comprised of two homologous domains, arranged antiparallel with respect to each other, in an arrangement known as an inverted repeat (discussed extensively elsewhere, for example, see Refs. [1–3]). This architecture is the vestige of a gene duplication and fusion so long ago that the structurally homologous domains no longer bear any discernable sequence identity. In almost all cases, that ancestral single domain gene has been lost to time. But dual-topology membrane proteins are rare examples in which that primitive architecture persists; they represent modern day, single-domain building blocks that assemble in the membrane with inverted architecture (Figure 1a).

The orientation of integral membrane proteins in bacteria can be predicted by the 'positive inside rule,' originally articulated by von Heijne [4], which describes a nearly universal feature of membrane proteins—a large excess of arginines and lysines on the cytoplasmic loops. In 2006, von Heijne described a rare exception to the rule: a class of small, unusually hydrophobic proteins that he called dual-topology proteins [5]. Monomers of dual-topology proteins are inserted into cell membranes with no inward/outward-facing bias and can dimerize in an antiparallel orientation *in vivo* [6•]. As the original controversy regarding this unusual pattern of membrane insertion has receded, dual-topology proteins have been recognized as likely evolutionary antecedents to inverted repeat [1,7,8••].

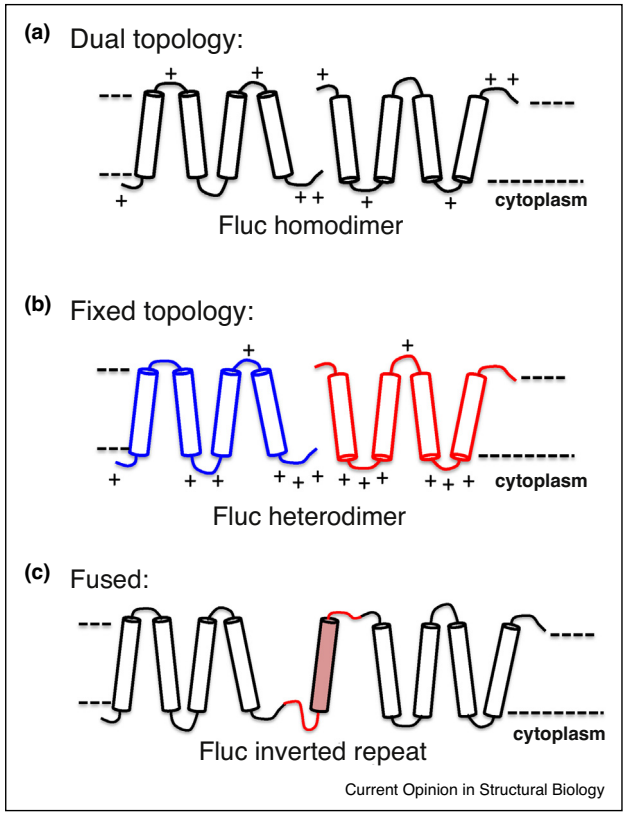
Topological diversity in Fluc family fluoride channels

Using sequence analysis, a number of putative dual-topology protein families have been proposed [5,9,10], but only two have been characterized in great biochemical depth: the multidrug efflux pump EmrE [11,12], and a class of fluoride channels called Flucs (also known as crcB or FEX), which protect against toxic environmental fluoride ion (F^-) in weakly acidic conditions [13•,14•,15••]. Flucs are ubiquitous in bacteria, and also found in archaea and free-living eukaryotes (where they are known as FEX proteins), including yeasts and protozoa, plants, and even simple marine animals such as sponges, sea anemones, and tunicates [16•]. Fluoride export function has been established by genetic knockout strains of *Escherichia coli*, *Candida albicans*, *Saccharomyces cerevisiae*, and *Neurospora crassa*, which are all rendered hypersensitive to fluoride at environmental concentrations [14•,16•].

With this many biological test kitchens, the Flucs have evolved into a topologically diverse family that it is not constrained only to dual-topology homodimers. Modern Flucs sample each evolutionary way station: they are present in genomes as single genes that code for antiparallel homodimers [17•] (Figure 1a), oppositely inserted pairs (~30% sequence identity) that transport fluoride as obligate heterodimers [15••] (Figure 1b), and in eukaryotes, fused two-Fluc proteins linked by a TM helix that forces the conjoined subunits into an antiparallel orientation: an inverted repeat [16•] (Figure 1c). This full assortment of topological states is rare, and unique among proteins with known function.

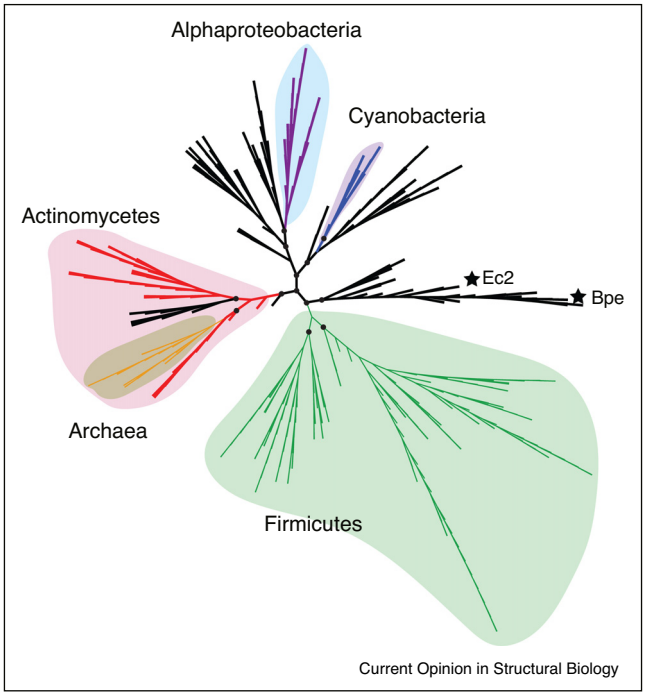
A phylogenetic tree constructed with ~500 representative bacterial and archaeal Fluc sequences shows that gene duplications leading to probable heterodimers have

Figure 1



Schematic illustrating subunit topologies found in the Fluc family of fluoride channels. (+) symbols indicate arginine and lysine residues of representative bacterial Fluc homologues. **(a)** Subunits of dual topology homodimers are inserted into the membrane in both inward- and outward-facing orientations, and can be identified by balanced (+) charge bias, as for this representative example from *Bordetella pertussis.* **(b)** For Fluc heterodimers, subunit pairs are inserted into the membrane with preferred, fixed orientations that can be predicted by the 'positive inside' rule [4], as in this representative heterodimer from *Lactobacillus acidophilus*. The termini of one subunit are always inward-facing (blue, positive charge bias on the termini and loop 2) and the termini of the other are always outward-facing (red, with positive charge bias on loops 1 and 3). **(c)** Cartoon of a fused Fluc, with a transmembrane 'inversion helix' (pink) enforcing the antiparallel domain architecture. Inverted repeat Flucs are found exclusively in eukaryotes, and have been experimentally determined to insert into the plasma membrane with the N-terminus facing the cytoplasm in *Saccharomyces cerevisiae* [24••]. The factors that determine topology in eukaryotic membrane proteins are more complex than in bacteria [41]; thus, the charge distribution is not shown for this homologue.

Figure 2



Unrooted phylogenetic tree constructed from ~500 representative Fluc proteins using maximum likelihood methods [42–44], with proposed duplication events indicated in color: Green: *Firmicutes* duplication; red: *Actinomycetes* duplication; blue: alphaproteobacteria duplication; violet: cyanobacteria duplication; yellow: archaeal duplication. Gene pairs were identified as likely heterodimers based on two criteria: the start of the second coding sequence falls within 100 nucleotides of the end of the first, and the two proteins have opposite loop charge biases, in accordance with the positive inside rule for membrane protein orientation, assessed by consensus prediction of membrane protein topology [45]. 92% of the pairs that met the first criterion also met the second. The remainder of pairs showed little positive charge bias, and thus prediction of the direction of insertion was uncertain. Amino acids likely to encode aqueous-exposed regions, inferred by structural alignment to the homodimeric Fluc structure, were trimmed from the alignment. This maneuver was essential to determine a phylogeny unbiased by the convergent evolution of positive charges on the inward-facing loops of heterodimers from divergent lineages. Nodes with >90% clade confidence according to approximate likelihood ratio test (aLRT [42]) are indicated by circles. Two structurally characterized homodimers, from *Bordetella pertussis* and a virulence plasmid isolated from *Escherichia coli*, are indicated by stars.

occurred at least four times in bacterial lineages: in firmicutes, in actinomycetes, in cyanobacteria, and in a handful of alphaproteobacteria (Figure 2). A fifth possible duplication occurred among a clade of archaea nestled among the actinomycetes. Archaeal lineages spring from bacterial clades here and elsewhere in the Fluc phylogeny, perhaps indicating that these proteins were transmitted via lateral gene transfer.

All eukaryotic Flucs are two-domain inverted repeats, and the eukaryotic clade is the sole example of a fusion event. Topological inspection reveals a simple explanation for the paucity of fused two-domain Flucs: to retain antiparallel topology for a protein with an even number of helices, the fusion event must also introduce a membrane spanning helix. Fusion events are more common for antiparallel proteins with an odd number of helices, as evidenced by the number of independent duplication/fusion events in the DUF606 (domain of unknown function) family [18••]. Sequence analysis suggests that the