BBA - Molecular Cell Research 1864 (2017) 1439–1449

Contents lists available at ScienceDirect

BBA - Molecular Cell Research

journal homepage: www.elsevier.com/locate/bbamcr

Tubular lipid binding proteins (TULIPs) growing everywhere☆

Louise H. Wong, Tim P. Levine*

UCL Institute of Ophthalmology, 11-43 Bath Street, London EC1V 9EL, UK

ARTICLE INFO

Keywords:
Anacetrapib
Cystic fibrosis transmembrane conductance regulator (CFTR)
Endoplasmic reticulum-mitochondria encounter structure (ERMES)
Extended-synaptotagmin (E-Syt)
Lipid traffic
Lipopolysaccharide (LPS)
PDZK8
Tether

ABSTRACT

Tubular lipid binding proteins (TULIPs) have become a focus of interest in the cell biology of lipid signalling, lipid traffic and membrane contact sites. Each tubular domain has an internal pocket with a hydrophobic lining that can bind a hydrophobic molecule such as a lipid. This allows TULIP proteins to carry lipids through the aqueous phase. TULIP domains were first found in a large family of extracellular proteins related to the bacterial permeability-inducing protein (BPI) and cholesterol ester transfer protein (CETP). Since then, the same fold and lipid transfer capacity have been found in SMP domains (so-called for their occurrence in synaptotagmin, mitochondrial and lipid binding proteins), which localise to intracellular membrane contact sites. Here the methods for identifying known TULIPs are described, and used to find previously unreported TULIPs, one in the silk polymer and another in prokaryotes illustrated by the *E. coli* protein YceB. The bacterial TULIP alters views on the likely evolution of the domain, suggesting its presence in the last universal common ancestor. The major function of TULIPs is to handle lipids, but we still do not know how they work in detail, or how many more remain to be discovered. This article is part of a Special Issue entitled: Membrane Contact Sites edited by Christian Ungermann and Benoit Kornmann.

1. Introduction

Evolution has created many situations in which hydrophobic, water-insoluble molecules such as lipids are transported through the aqueous phase by proteins. Lipid transport outside cells is important for scavenging or detecting lipids within the environment. Lipid transport inside cells is required to move lipids between membrane-bound compartments, even if these compartments are linked by vesicular traffic [1]. The lipid transfer proteins (LTPs) that mediate lipid transport both inside and outside cells have often been discovered through purification of an activity that carries lipid between artificial bilayers *in vitro* [2]. Structural studies of these LTPs show that they all possess a cavity that engulfs the hydrophobic portion of a lipid, if not the entire lipid molecule [3]. The number of lipids that organisms deal with is large, and the non-vesicular routes inside cells are also numerous, so it is interesting to find that over 100 LTPs are encoded in a mammalian genome [4], and that 17 different protein folds can act as LTPs for lipids [3]. This review describes the Tubular lipid binding proteins (TULIPs), a large superfamily of LTPs with a distinctive tubular fold [5]. Note that "family" is a technical term referring to a group of protein homologues that can be linked together by applying conventional sequence alignment tools, such as PSI-BLAST or PFAM. In contrast for TULIPs as a whole the correct term is "superfamily", which indicates multiple families sharing sequence similarity that is not detected by conventional tools, but is detected by more sensitive sequence comparison methods. Thus, new TULIP domains have been found that had could not have been identified by sequence homology with known domains. The discovery of TULIP structures in bacteria that is reported here shows that the superfamily is even more widespread than previously thought.

2. Classifying the full range of TULIPs

2.1. The main varieties of TULIP

Some of the first LTPs ever described are now known as founding members of the TULIP family [6,7]. These are extracellular mammalian proteins that solubilise a range of hydrophobic molecules, including vitamins, phospholipids and neutral lipids such as cholesterol ester and

Abbreviations: BPI, bacterial permeability-inducing protein; CETP, cholesterol ester transfer protein; CFTR, cystic fibrosis transmembrane conductance regulator; DALI, distance matrix alignment; DUF, domain of unknown function; ENaC, epithelial sodium channel; ER, endoplasmic reticulum; ERMES, ER-mitochondrial encounter structure; E-Syt, extended-synaptotagmin; JH, juvenile hormone; JHBP, JH binding protein; LTP, lipid transfer protein; LPS, lipopolysaccharide; LBP, LPS-binding protein; LAM, LTP anchored at a membrane contact site; PC, phosphatidylcholine; PS, phosphatidylserine; PLTP, phospholipid transfer protein; PLUNC, palate lung and nasal epithelial clone; RMSD, root mean square distance; SMP, synaptotagmin mitochondria and lipid binding protein; TULIP, tubular lipid binding protein

☆ This article is part of a Special Issue entitled: Membrane Contact Sites edited by Christian Ungermann and Benoit Kornmann.

* Corresponding author.
E-mail address: tim.levine@ucl.ac.uk (T.P. Levine).

http://dx.doi.org/10.1016/j.bbamcr.2017.05.019
Received 23 March 2017; Received in revised form 11 May 2017; Accepted 17 May 2017
Available online 26 May 2017

triacylglycerols. Both cholesteryl ester transfer protein (CETP) and phospholipid transfer protein (PLTP) transfer lipids between lipoproteins in the extracellular fluid. Two other members of this family bind lipopolysaccharide (LPS), the major lipid of the outer membrane of Gram-negative bacteria that induces toxic shock. One is the plasma protein LPS-binding protein (LBP), which mediates innate immune responses to LPS by presenting it to monocytes [8,9]. The other is Bacterial Permeability-Increasing protein (BPI) found in azurophilic granules of neutrophil polymorphs, which has a complex oxygen-independent bactericidal action that arises from LPS binding [10,11].

CETP, PLTP, LBP and BPI (hereafter referred to as CETP/BPI) were originally identified for their lipid solubilizing and transfer activities, with their TULIP folds discovered subsequently by X-ray crystallography. CETP/BPI are founders of a much wider protein family of sequence homologues that include the palate lung and nasal epithelial clone (PLUNC, also called BPI fold-containing family, BPIF) proteins. These are highly abundant components of the secretions of upper airways, nasopharynx and tears of mammals [12–14] and of egg albumen in birds [15,16]. In addition, there are full length homologues (homology spanning > 450 aa) widely dispersed in all branches of eukaryotic evolution, from fungi to protists, algae and plants. For example, CETP/BPI homologues are found in ascomycetes, even though they are missing from *S. cerevisiae*. All CETP/BPI family proteins are extracellular, and most are secreted proteins 350–500 aa long (Fig. 1A). However, a small number of protist and amoebal proteins are anchored in the plasma membrane by a C-terminal PqiA domain. This domain, which is usually found in prokaryotic proteins, has 9–12 transmembrane domains and a proposed function as a lipid permease (Fig. 1A) [17], which suggests that the associated CETP/BPI domains may handle the same lipid ligands as the PqiA domain.

2.2. Stumbling across double TULIPs and identifying more singles

In addition to the proteins related to CETP/BPI by BLAST and PSI-BLAST, there are some whose similarity was only discovered when their structures were solved. For any newly solved structure, the Cα backbone can be aligned in three dimensions to all other structures by tools such as DALI (distance matrix alignment) [18]. The first and most important discovery was made by solving the protein structure of one member of the family. This revealed that every CETP/BPI family member contains a large internal repeat [19], meaning that the basic unit of "TULIP-ness" is not the whole of CETP/BPI, but just half of it. Each half, *i.e.* both the TULIP-N and the TULIP-C domains on their own, have LTP activity and they each have tubular structures with separate lipid binding cavities (see next section). Thus, most CETP/BPI family members contain two LTPs as a tandem hetero-dimer. The main variation on this pattern is in the PLUNCs, where the TULIP-C domain can be missing, leaving just one domain (short PLUNC) (Fig. 1A) [20].

Another part of the CETP/BPI family that was originally identified purely from solved structures is found in arthropods, which appeared to have no sequence homologues of CETP/BPI. The first arthropod TULIPs identified were the juvenile hormone (JH) binding protein (JHBP) in haemolymph [21,22], and the regulatory protein Takeout [23], the structures of which turned out to be super-imposable on TULIP-N domains from CETP or BPI. JHBP and Takeout have hydrophobic ligands (JH and isoprenoids respectively), which the proteins bind in the extracellular milieu and deliver to cells, similar to the way LBP presents LPS. The same arthropod family includes the group 7 allergens, for example the house dust mite *Dermatophagoides pteronyssinus* allergen Der p 7. These are secreted proteins that may function in binding and presentation of polymyxin B, a bacterial lipopeptide, to the innate immune system [24]. As for the link between CETP/BPI and SPAC32A11.02c, Der p 7 is clearly in the same family as JHBP and Takeout, although PSI-BLAST searches must be initiated with Der p7, not JHBP, to reveal this (data not shown).

2.3. Profile-profile searches find TULIPs inside cells

The discovery of the two structurally similar domains within CETP/BPI indicated that an ancient duplication of the domain took place, followed by such wide sequence divergence as to make the homology almost invisible, although convergent evolution cannot be formally excluded [20,25]. Homologies within the TULIP superfamily can be completely missed by conventional sequence alignments. However,

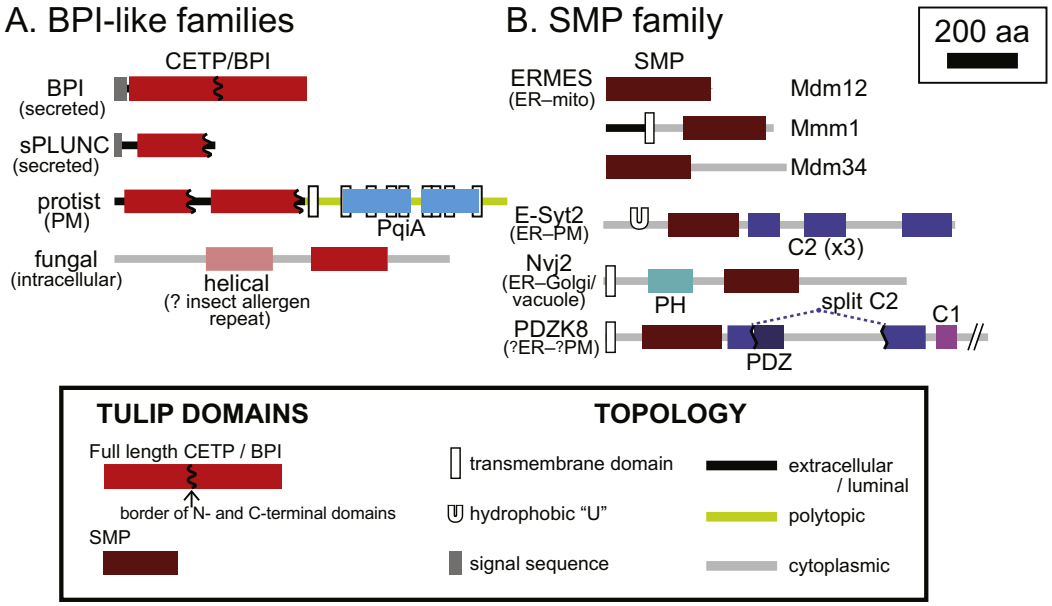

Fig. 1. Domain structure of TULIP proteins. TULIP domains are found in the two major families: CETP/BPI domains (A) and SMP domains (B). Family and membrane topologies (cytoplasmic *vs.* polytopic *vs.* secreted to become extracellular or luminal) are indicated as shown in the key. Location and accompanying domains are marked in the figure. The SMP proteins shown here are representatives of the four types that are widely spread through eukaryotes [20]. The fungal protein shown is SPAC32A11.02c from *S. pombe*. Its entire N-terminus is predicted to be helical, with residues 232–401 aligning well in HHsearch with residues 44–172 of Bla g 1, a helical insect allergen repeat protein crystalised as 4jrb. HHsearch does not make this prediction definitively, and other all-helical hits include bimodular sensor domains, which typically bind chemo-attractants. The C2 domain in PDZK8, split so that its two segments can assemble into one complete domain, was missed previously [34].