

Differences in the SNARE evolution of fungi and metazoa

Nickias Kienle, Tobias H. Kloepper and Dirk Fasshauer

Research Group Structural Biochemistry, Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany
e-mail N.K.: nkienle@gwdg.de; , phone: +49-551-201 1621, Fax: +49-551-201 1499;
e-mail T.H.K.: tkloepper@gwdg.de
e-mail D.F.: dfassha@gwdg.de

Correspondence should be addressed to N.K

Key words: fungi SNARE, metazoa SNARE, membrane fusion, multicellularity, vesicle trafficking, evolution

Abbreviations used: SNARE, soluble N-ethylmaleimide-sensitive factor attachment protein receptor; TMR , transmembrane region; HMM, Hidden Markov Models; Npsn, novel plant SNARE; PX, Phox homology; Syb, Synaptobrevin; Syx, Syntaxin; Syp, Syntaxin of plants; PI(3)P, phospholipid phosphatidylinositol 3-phosphate.

Abstract

Proteins of the SNARE family are key factors in all vesicle fusion steps in the endocytic and secretory pathways. SNAREs can assemble into a tight four-helix bundle complex between opposing membranes, a process that is thought to pull the two membranes into close proximity. The complex-forming domains are highly conserved, not only between different species but also between different vesicle trafficking steps. SNARE protein sequences can be classified into four main types (Qa, Qb, Qc, and R), each reflecting their position in the four-helix bundle. Further refinement of these main types resulted in the identification of 20 distinct, conserved groups, which probably reflect the original repertoire of a proto-eukaryotic cell. We analyzed the evolution of the SNARE repertoires in metazoa and fungi and unveiled remarkable differences in both lineages. In metazoa, the SNARE repertoire appears to have undergone a substantial expansion, particularly in the endosomal pathways. This expansion likely occurred during the transition from a unicellular to a multicellular lifestyle. We also observed another expansion that led to a major increase of the secretory SNAREs in the vertebrate lineage. Interestingly, Fungi developed multicellularity independently, but in contrary to plants and metazoa, this change was not accompanied by an expansion of the SNARE set. Our findings suggest that the rise of multicellularity is not generally linked to an expansion of the SNARE set. The structural and functional diversity that exists between fungi and metazoa might offer a simple explanation for the distinct evolutionary history of their SNARE repertoires.

Introduction

The eukaryotic cell encompasses a large system of intracellular membrane delimited compartments. It is widely assumed that this extensive endomembrane system evolved as a result of a phagotrophic lifestyle by invagination of the plasma membrane. Vesicles, small, intracellular, membrane-enclosed sacs, are utilized as carriers to mediate material exchange between different compartments. During this process, vesicles bud from a donor organelle, target, and then fuse with an acceptor organelle. Proteins of the SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) family are established components of all vesicle fusion steps (reviewed by [1, 2]). They constitute a large family of cytoplasmic oriented membrane proteins anchored by a C-terminal transmembrane region (TMR). SNAREs assemble into a tight four-helix bundle complex between opposing membranes. This zipper-like assembly, from the N-terminus towards the C-terminus, is thought to pull the lipid bilayers into close proximity, thereby overcoming the repulsive forces between the membranes. Remarkably, the complex-forming SNARE domains are highly conserved, not only between different species but also between different vesicular trafficking steps, and allow the classification of SNARE proteins into four main types (Qa, Qb, Qc, and R). Interestingly, these main types reflect their different positions within the four-helix bundle [3-5].

In the past, classical approaches to define different subtypes of the SNARE family utilized only a few basic models (e.g. SMART or Pfam), with the primary goal to achieve high sensitivity. However, the major limitation of these approaches is that the specificity of the result is unknown. Additionally, these methods often fail to identify the putative SNARE motif itself. Early attempts to improve the basic classification of SNAREs employed either a crude clustering approach [6] or focused solely on a few organisms/sequences [5]. To shed more light towards the conservation of the endomembrane system, we have recently developed a highly sensitive and specific universal classification of SNARE proteins [7]. Our initial set comprised SNARE sequences from several model organisms. These were used to define a set of clusters based on their evolutionary history and sequence similarity. For each of the clusters we trained a Hidden Markov Model (HMM). With the set of HMMs, we screened protein databases and additional genome projects for candidate SNARE sequences. An iterative refinement of the classification, including eye-by-eye verifications to assure a high quality of the collected sequences, yielded in 20 distinct, conserved functional groups. These different SNARE groups are involved in distinct intracellular trafficking steps (Fig. 1). It is conceivable that this set represents the original SNARE repertoire of a proto-eukaryotic cell.

Closer dissection of the SNARE collection from several model organisms revealed that these species often possess more than 20 different SNARE proteins (*Arabidopsis thaliana* 62; *Homo sapiens* 44; *Caenorhabditis elegans* 30; *Drosophila melanogaster* 26; *Saccharomyces cerevisiae* 24), implying that their sets were modified during evolution. Furthermore, it appears that many SNAREs have been duplicated, followed by further diversification [7, 8]. To ensure accurate reconstruction of the evolutionary changes of the SNARE family, a point of comparison is indispensable. The eukaryotic ancestor would thus be the logical reference. However, since it is not available, a basal organism that is exposed to low evolutionary pressure would be an ideal substitute. Fortunately, basal organisms, with available genomes, that possess an almost unmodified SNARE repertoire exist for most kingdoms. This observation, together with our universal classification scheme, provides an excellent starting point to reconstruct the evolutionary history of the SNARE subfamilies in different eukaryotic lineages. To accomplish this, it is fundamental to include SNAREs from a wide variety of species. Thus, we analyzed the history of SNARE proteins in fungi and metazoa by collecting SNARE sequences from a large multiplicity of different species, including basal organisms [9, 10]. Interestingly, we found remarkable differences in their respective SNARE sets.

SNAREs in fungi

Having inspected the SNARE repertoire of more than 70 fungal species with completely sequenced genomes, our major finding indicates that the SNARE sets remain largely unchanged in fungi as compared to the assumed set of the proto-eukaryotic ancestor. In most fungi, each SNARE subgroup is comprised only of one member and thus we observed only little over 20 different SNAREs in most fungi species.

In general, fungi possess two Qc.III.c SNAREs, syntaxin 8 (Syx8) and Vam7 [11-14], whereas most other eukaryotes examined possessed only one Qc.III.c homolog. Vam7 is the only SNARE with an N-terminal Phox homology (PX) domain. The PX domain can interact with the phospholipid phosphatidylinositol 3-phosphate (PI(3)P) [15], which is specific for the membranes of the endosomal and vacuolar pathways, suggesting a role in endosomal trafficking. Furthermore, the acquisition of this novel membrane-binding domain by Vam7 possibly compensated for a loss of the C-terminal TMR. The Vam7 protein is a unique invention of the fungi kingdom and is the only SNARE with this specific domain structure. Thus, Vam7 is a defining feature (apomorphy) of the fungi lineage and can therefore be used as a criterion for the recognition of fungal species.

Remarkably, we also discovered that the basal fungi *Batrachochytrium dendrobatidis* and *Blastocladiella emersonii* both contain a sequence, classified as Qb.III.d (novel plant SNARE (Npsn)) [16]. Initially, Npsn was thought to be only present in plants [16] and protists [7]. Our analysis indicates that Npsn is absent in metazoa and other, more derived fungi.

Moreover, we observed few changes in the endosomal/vacuolar and secretory SNARE set of the Saccharomycotina lineage [10].

SNAREs in Metazoa

In contrast to the moderate changes of the SNARE set in fungi species, we detected marked changes in animals. However, we found that the unicellular choanoflagellate *Monosiga brevicollis*, which is closely related to metazoa, still contains a rather simple set of SNARE proteins.

Compared to *Monosiga brevicollis*, the SNARE set of lower metazoa (e.g. placozoa, cnidaria and several bilaterians) is already enlarged [9]. This suggests that a major expansion of the SNAREs occurred during the rise of multicellularity. It is possible that whole genome duplications at the base of the metazoa evolution led to an enlargement of the SNARE repertoire, with this event affecting mostly the endosomal SNAREs. An expansion in the R.III SNAREs gave rise to Vamp7-like and Vamp4 in addition to the ancestral Vamp7. Likewise, three different Qa.III.b SNAREs can be found, Syx7, Syx17 and Syx20. Interestingly, and in contrast to the common architecture of the SNARE proteins, Syx17 possesses two C-terminal TMRs. Another multiplication occurred in the Qbc.IV group, which contains three different homologs in basal metazoans, SNAP-25 [17], SNAP-29 [18, 19] and SNAP-47 [20]. Two additional factors, which do not possess the SNARE motif, but are clearly derived from SNARE proteins, emerged as well: Sec22-like (Sec22a) [21] is derived from the R.I SNARE Sec22 (Sec22b) and Lgl, which belongs to the R.Reg group, is a homolog of tomosyn [22, 23].

We observed another expansion that led to a major increase in the secretory SNARE set in the lineage of the vertebrates, described in detail in [9]. This expansion is probably due to two well-established consecutive rounds of whole genome duplications in the vertebrate lineage [24, 25].

Comparison and Hypotheses

It has been proposed that the transition from unicellular to multicellular organisms is concomitant with an expansion of the SNARE repertoire, in particular of the ones involved

in secretion and endosomal trafficking. In fact, a marked expansion has also been found in the SNARE sets of green plants [26], although the exact chronology of this event still needs to be determined. In addition, we have shown that the SNARE sets in metazoa underwent a substantial expansion during the transition to multicellularity. Most interestingly, fungi developed multicellularity independently, but in contrary to plants and metazoa, this change was not accompanied by an expansion of the SNARE set. The fungal SNAREs generally represent the types of the assumed proto-eukaryotic SNARE repertoire. Furthermore, we noted that the SNARE sets of a few single-cell eukaryotes are also markedly expanded [7]. For example, the ciliate *Paramecium tetraurelia* possesses around 70 SNAREs. *Paramecium tetraurelia* has a complex intracellular architecture containing multiple distinct organelles and therefore likely requires additional membrane trafficking pathways to sustain its functional homeostasis. Another example of a single-cell eukaryote with a markedly enlarged SNARE repertoire is the kinetoplastid *Trichomonas vaginalis*. We identified 46 SNAREs in this organism [7]. Taken together, these observations suggest that a SNARE expansion is not generally linked to the rise of multicellularity.

At a molecular level, gene/genome duplications are an important driving force for evolutionary changes. Initially, duplicated gene products carry out the same function, but, in the course of evolution, the two gene copies can specialize to perform complementary functions (subfunctionalization) or one of the copies can acquire a new function (neofunctionalization). In any case, as both factors specialize in their respective processes their sequences diverge. The species of the Saccharomycotina clade, for example, possess two Qa.III.b SNAREs, Pep12 and Vam3. In baker's yeast these proteins are believed to be involved in trafficking to late endosomes and vacuoles, respectively. As most fungi typically carry only one Qa.III.b SNARE, Pep12, it is possible that this ancestral protein mediates both trafficking steps. Another possibility exists in Saccharomycotina in which the more derived copy, Vam3, could have specialized in homotypic vacuolar fusion, a process that might not occur in other fungi. A further example of gene duplication was provided by a recent whole genome duplication in yeast. Interestingly, only the four secretory SNAREs were found to be maintained [10]. While the two secretory syntaxins (Sso1 and Sso2) and the two secretory R-SNAREs (Snc1 and Snc2) remain highly similar, the two Qbc.IV SNAREs Sec9 and Spo20 are remarkably diverse in their sequence identity as well as in their functionality. Whereas Sec9 is able to interact with both secretory syntaxins and R-SNAREs in secretion during vegetative growth, Spo20 is known to be necessary for the process of sporulation [27-29].

We noted that in most organisms inspected, SNAREs of the ER/Golgi pathways are usually present as singletons, suggesting that these basic secretory routes are, in particular, highly preserved. It appears that these fundamental pathways usually do not tolerate changes easily and thus no duplication prevailed. Interestingly, persistent SNARE duplications can be found largely in the endosomal and secretory pathways. For example, the first major expansion in the metazoa affected mostly endosomal SNAREs, whereas the majority that were affected by the second expansion were secretory SNAREs. Furthermore, the changes we observed in fungi, mainly within the Saccharomycotina lineage, largely impacted SNAREs involved in endosomal trafficking. Given that endosomal compartments function as a sorting hub for cargo of the exocytic and endocytic pathways, it is plausible that an extended endosomal SNARE set might result in more flexible trafficking routes between these compartments. As outlined above, the SNARE repertoires of fungi and metazoa are remarkably distinct, despite both lineages being multicellular. Thus, the question arises as to why fungi persist with a relatively simple SNARE set, whereas metazoa require an extended SNARE repertoire? How can these differences be explained? An answer might be simply gleaned from the marked

differences in lifestyle, as well as the diverse structural and functional conformations that exists between fungi and metazoa.

It is likely that the last common ancestor of fungi and metazoa, the proto-opisthokont, was a motile unicellular organism that engulfed prey through phagocytosis [30]. In the lineage of fungi, phagocytosis seems to have lost its general importance during evolution as most fungi species are detritivores. For nutrient uptake, most fungi form long, branching structures, so-called hyphae. Hyphae are divided into cells that are surrounded by a chitinous cell wall. Multiple interconnected hyphae form a mycelium. Hyphae grow at their tips towards the food supply. A fungi specific sub-cellular structure, the Spitzenkoerper, contains secretory vesicles that fuse with the membrane at the growing tip. The vesicles can secrete material to extend the cell wall and lytic exoenzymes to degrade larger organic compounds in the surrounding medium. The degraded products can then be absorbed by the fungal cell. Thus, fungi appear not to have a defined structural conformation. Metazoa on the other hand are motile heterotrophs that generally digest their food in an internal chamber. They have a defined body plan and their cells, which lack a cell wall, differentiate into separate tissues, each assigned with a specialized function. Some cell types for instance specialized in inter-cellular communication (e.g. neurons). In fact, a primitive neuronal system has been identified in basal organisms like the sea anemone *Nematostella vectensis* [31]. It is therefore highly plausible that the expansion of SNARE repertoires in basal metazoa confers an evolutionary advantage by allowing for the provision of more versatile endosomal and secretory trafficking pathways. These additional transport routes might have better facilitated the sorting of molecules to different areas of the plasma membrane. For example, tissue cells became capable of engulfing prey on the one side of the cell whilst secreting material on the opposing side to feed the neighboring cells. The expansion of the secretory set of SNAREs in vertebrates might have provided yet another level of control and fine-tuning of the trafficking pathways and hence allowed for the rise of specialized secretory cell types.

All in all, despite the recent classification of the SNARE repertoires in fungi and metazoa, much remains unknown about specific SNARE mediated processes in eukaryotic cells. The future culmination of complementary biological, biochemical and computational studies may help pave the way towards a better understanding of these enigmatic proteins.

Fig.1: Four-helix bundle structure of the neuronal SNARE complex, distribution of the basic SNARE repertoire within the cell and their allocation into functional units

The four-helix bundle structure of the neuronal SNARE complex is shown as ribbon diagram in the middle (blue, red, and green for synaptobrevin 2, syntaxin 1a, and SNAP-25a, respectively). The layers (-7 until +8) in the core of the bundle are indicated by virtual bonds between the corresponding C α positions. We divided SNARE proteins into four main groups and into 20 subgroups using an extensive classification analysis. Genuine complexes are composed of four different SNARE motifs each belonging to one of the four main groups ('QabcR' composition). Putative SNARE units have been assigned to the basic transport steps. In addition to the fusogenic SNARE proteins, a regulatory R-SNARE without membrane anchor, tomosyn, exists. The most commonly used names for the different SNARE types are given. For historical reasons, the names used for homologous SNAREs are often different in the different eukaryotic kingdoms. The different names used for metazoa (m), fungi (f), and plants (p) are listed. The names syntaxin and synaptobrevin (the secretory R-SNARE of metazoa that is also referred to as VAMP) are abbreviated by Syx and Syb, respectively. Several plant Q-SNAREs have been named Syntaxin of plants (Syp). Moreover, several, more special names of the markedly increased SNARE repertoire of vertebrates are not listed.

1. Hong, W., *SNAREs and traffic*. 2005. **1744**(3): p. 493-517.
2. Jahn, R. and R.H. Scheller, *SNAREs--engines for membrane fusion*. *Nat Rev Mol Cell Biol*, 2006. **7**(9): p. 631-43.
3. Weimbs, T., et al., *A conserved domain is present in different families of vesicular fusion proteins: a new superfamily*. 1997. **94**(7): p. 3046-51.
4. Fasshauer, D., et al., *Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs*. *Proc Natl Acad Sci USA*, 1998. **95**(26): p. 15781-6.
5. Bock, J.B., H.T. Matern, and A.A. Peden, *A genomic perspective on membrane compartment organization*. 2001. **409**(6822): p. 839-41.
6. Yoshizawa, A.C., et al., *Extracting sequence motifs and the phylogenetic features of SNARE-dependent membrane traffic*. *Traffic*, 2006. **7**(8): p. 1104-18.
7. Kloepper, T.H., C. Nickias Kienle, and D. Fasshauer, *An Elaborate Classification of SNARE Proteins Sheds Light on the Conservation of the Eukaryotic Endomembrane System*. *Mol Biol Cell*, 2007.
8. Dacks, J.B. and M.C. Field, *Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode*. *J Cell Sci*, 2007. **120**(Pt 17): p. 2977-85.
9. Kloepper, T.H., C.N. Kienle, and D. Fasshauer, *SNAREing the basis of multicellularity: Consequences of protein family expansion during evolution*. *Mol Biol Evol*, 2008.
10. Kienle, N., T.H. Kloepper, and D. Fasshauer, *Phylogeny of the SNARE vesicle fusion machinery yields insights into the conservation of the secretory pathway in fungi*. *BMC Evol Biol*, 2009. **In press**.
11. Sato, T.K., T. Darsow, and S.D. Emr, *Vam7p, a SNAP-25-like molecule, and Vam3p, a syntaxin homolog, function together in yeast vacuolar protein trafficking*. *Mol Cell Biol*, 1998. **18**(9): p. 5308-19.
12. Ungermann, C. and W. Wickner, *Vam7p, a vacuolar SNAP-25 homolog, is required for SNARE complex integrity and vacuole docking and fusion*. *EMBO J*, 1998. **17**(12): p. 3269-76.
13. Wada, Y. and Y. Anraku, *Genes for directing vacuolar morphogenesis in Saccharomyces cerevisiae. II. VAM7, a gene for regulating morphogenic assembly of the vacuoles*. *J Biol Chem*, 1992. **267**(26): p. 18671-5.
14. Wada, Y., Y. Ohsumi, and Y. Anraku, *Genes for directing vacuolar morphogenesis in Saccharomyces cerevisiae. I. Isolation and characterization of two classes of vam mutants*. *J Biol Chem*, 1992. **267**(26): p. 18665-70.
15. Cheever, M.L., et al., *Phox domain interaction with PtdIns(3)P targets the Vam7 t-SNARE to vacuole membranes*. *Nat Cell Biol*, 2001. **3**(7): p. 613-8.
16. Sanderfoot, A.A., F.F. Assaad, and N.V. Raikhel, *The Arabidopsis genome. An abundance of soluble N-ethylmaleimide-sensitive factor adaptor protein receptors*. *Plant Physiol*, 2000. **124**(4): p. 1558-69.
17. Oyler, G.A., et al., *The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal subpopulations*. *J Cell Biol*, 1989. **109**(6 Pt 1): p. 3039-52.
18. Steegmaier, M., et al., *Three novel proteins of the syntaxin/SNAP-25 family*. *J Biol Chem*, 1998. **273**(51): p. 34171-9.
19. Wong, S.H., et al., *GS32, a novel Golgi SNARE of 32 kDa, interacts preferentially with syntaxin 6*. *Mol Biol Cell*, 1999. **10**(1): p. 119-34.
20. Holt, M., et al., *Identification of SNAP-47, a novel Qbc-SNARE with ubiquitous expression*. *J Biol Chem*, 2006. **281**(25): p. 17076-83.

21. Hay, J.C., et al., *Protein interactions regulating vesicle transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells*. Cell, 1997. **89**(1): p. 149-58.
22. Fasshauer, D. and R. Jahn, *Budding insights on cell polarity*. Nat Struct Mol Biol, 2007. **14**(5): p. 360-2.
23. Pobbati, A.V., et al., *Structural basis for the inhibitory role of tomosyn in exocytosis*. J Biol Chem, 2004. **279**(45): p. 47192-200.
24. Holland, P.W., *More genes in vertebrates?* J Struct Funct Genomics, 2003. **3**(1-4): p. 75-84.
25. Panopoulou, G. and A.J. Poustka, *Timing and mechanism of ancient vertebrate genome duplications -- the adventure of a hypothesis*. Trends Genet, 2005. **21**(10): p. 559-67.
26. Sanderfoot, A., *Increases in the number of SNARE genes parallels the rise of multicellularity among the green plants*. Plant Physiol, 2007. **144**(1): p. 6-17.
27. Neiman, A.M., *Prospore membrane formation defines a developmentally regulated branch of the secretory pathway in yeast*. J Cell Biol, 1998. **140**(1): p. 29-37.
28. Neiman, A.M., *Ascospore formation in the yeast Saccharomyces cerevisiae*. Microbiol Mol Biol Rev, 2005. **69**(4): p. 565-84.
29. Neiman, A.M., L. Katz, and P.J. Brennwald, *Identification of domains required for developmentally regulated SNARE function in Saccharomyces cerevisiae*. Genetics, 2000. **155**(4): p. 1643-55.
30. Cavalier-Smith, T., *The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa*. Int J Syst Evol Microbiol, 2002. **52**(Pt 2): p. 297-354.
31. Putnam, N.H., et al., *Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization*. Science, 2007. **317**(5834): p. 86-94.

