Tracking the origin and divergence of cholinesterases and neuroligins: the evolution of synaptic proteins

Nicolas Lenfant, Thierry Hotelier, Yves Bourne, Pascale Marchot, Arnaud Chatonnet

Tracking the Origin and Divergence of Cholinesterases and Neuroligins: The Evolution of Synaptic Proteins

Nicolas Lenfant, Thierry Hotelier, Yves Bourne, Pascale Marchot & Arnaud Chatonnet

Volume 51 Number 3, November 2013 ISSN: 0895–8696

Journal of
Molecular
Neuroscience

Editor-in-Chief: ILLANA GOZES, PhD

Issue Theme
Mechanisms

ONLINE FIRST

🕿 Springer

🕿 Springer

Springer

Tracking the Origin and Divergence of Cholinesterases and Neuroligins: The Evolution of Synaptic Proteins

Nicolas Lenfant · Thierry Hotelier · Yves Bourne ·
Pascale Marchot · Arnaud Chatonnet

Abstract A cholinesterase activity can be found in all kingdoms of living organism, yet cholinesterases involved in cholinergic transmission appeared only recently in the animal phylum. Among various proteins homologous to cholinesterases, one finds neuroligins. These proteins, with an altered catalytic triad and no known hydrolytic activity, display well-identified cell adhesion properties. The availability of complete genomes of a few metazoans provides opportunities to evaluate when these two protein families emerged during evolution. In bilaterian animals, acetylcholinesterase co-localizes with proteins of cholinergic synapses while neuroligins co-localize and may interact with proteins of excitatory glutamatergic or inhibitory GABAergic/glycinergic synapses. To compare evolution of the cholinesterases and neuroligins with other proteins involved in the architecture and functioning of synapses, we devised a method to search for orthologs of these partners in genomes of model organisms representing distinct stages of metazoan evolution. Our data point to a progressive recruitment of synaptic components during evolution. This finding may shed light on the common or divergent developmental regulation events involved into the setting and maintenance of the cholinergic versus glutamatergic and GABAergic/glycinergic synapses.

N. Lenfant · A. Chatonnet (✉)
Dynamique Musculaire et Métabolisme, INRA, Université
Montpellier, Place Viala, Montpellier, France
e-mail: chatonne@supagro.inra.fr

N. Lenfant · Y. Bourne · P. Marchot
Architecture et Fonction des Macromolécules Biologiques, CNRS/
Aix-Marseille Université, Campus Luminy, Marseille, France

T. Hotelier
Unité Informatique INRA-AgroM, Montpellier, France

Introduction

The large COesterase (PFAM PF00135) family encompasses a number of proteins characterized by a canonical alpha/beta hydrolase fold domain preceded by an amino-terminal extension containing a disulfide bond. The family is composed mainly of carboxylesterases, cholinesterases, and lipases; it also contains a number of proteins with an altered catalytic triad and no known hydrolytic activity (Krejci et al. 1991). Some of the later ones, such as the neuroligin, neurotactin, gliotactin, and glutactin ectodomains and the thyroglobulin carboxy-terminal domain, participate in homophylic or heterophylic protein–protein interactions (Schulte et al. 2003; Olson et al. 1990; Darboux et al. 1996; Marchot and Chatonnet 2012; De Jaco et al. 2012). Members of this family are highly represented in opisthokonts (fungi/metazoans) (Johnson and Moore 2012b; Pezzementi and Chatonnet 2010; Lenfant et al. 2013a). Acetylcholinesterase hydrolyzes the neurotransmitter acetylcholine at cholinergic synapses of all bilaterian animals (vertebrates, insects, nematodes, etc.). However, some cholinesterases display poor or no catalytic activity and some acetylcholinesterases are located outside of the cholinergic system, two features raising a question about their true function(s) (Silman and Sussman 2005; Kim et al. 2012; Kim and Lee 2013; Johnson and Moore 2012a). The extracellular domains of neuroligins and acetylcholinesterases display ~35 % sequence identity and high structural homology (root mean square deviation value of ~1.35 Å for at least 330 Cα atoms (Leone et al. 2010)). Similarities and differences between neuroligin and acetylcholinesterase are highlighted

on a schematic topological structure (Fig. 1). Neuroligins have no catalytic activity, but they bind to various partners (neurexins, MDGA, etc.). They are present at excitatory glutamatergic and inhibitory GABAergic/glycinergic synapses as well as in cholinergic synapses as more recently reported (Takács et al. 2013). The role of neuroligins in synapse functioning has been analyzed in various bilaterian organisms (Biswas et al. 2008; Choi et al. 2011; Knight et al. 2011; Hu et al. 2012; DeJaco et al. 2012). Experimental acetylcholinesterase overexpression in the nervous system impairs glutamatergic transmission, a feature suggesting that acetylcholinesterase and neuroligin could share some functional properties (Grifman et al. 1998; Soreq and Seidman 2001; Dong et al. 2004). It is therefore interesting to decipher when the cholinesterases and neuroligins emerged from their common ancestor in the metazoan lineage and when they acquired distinctive functions. To correlate the divergence of neuroligin and acetylcholinesterase with the divergence of neurotransmitter-specific synapses (e.g., cholinergic versus glutamatergic or GABAergic/glycinergic), we devised a program that automatically retrieves and compares the most likely orthologs of synaptic proteins from organisms representative of major lineages of metazoans for which complete genomes are documented.

Materials and Methods

Phylogenetic Tree of Cholinesterases, Neuroligins, and Related Proteins

The COesterase (PFAM PF00135) family contains about 6,000 entries in the ESTHER database (Lenfant et al. 2013b) (this is already a compressed value as many sequences from different strains are collapsed in one entry). For analysis, we excluded sequences of bacteria and fungi and short sequences (pseudo genes, fragments, etc.). We conserved only those sequences that have a clear second cysteine residue involved in the first disulfide bond situated in close variants of the consensus SEDCLYLN (PROSITE PS00941) sequence. These criteria ensure that the sequence is quite complete at the amino-terminus and contains the signature of the family. Sequences were aligned with ClustalO (Sievers et al. 2011), and the alignment was subjected to minor corrections to conform to known structural features of COesterases and exclude amino- and carboxy-terminal extensions found only in non-catalytic proteins. The evolutionary history was inferred using the neighbor-joining method (Saitou and Nei 1987). The optimal tree with a sum of branch lengths equal to 298 is shown (Fig. 2 and Supplementary Material). The tree is drawn to scale, with branch lengths in the same unit as the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson's correction method (Zuckerkandl and Pauling 1965) and are reported as the number of amino acid substitutions per site. The analysis involved 1,190 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 3,256 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). The full alignment and the tree can be retrieved from ESTHER and imported in the MEGA5 suite.

Evolution of Other Synaptic Proteins in Metazoan

We restricted our analysis of orthologs to a limited number of representative species and proteins. Concerning species, we chose three vertebrate species: two primates *Homo sapiens* (Hs) and *Pan troglodyte* (Pt), in which the orthology of the studied proteins is not questionable, and *Danio rerio* (Dr). Three other species, *Caenorhabditis elegans* (Ce), *Drosophila melanogaster* (Dm), and *Schistosoma mansoni* (Sm), represent some of the other major triploblastic/bilaterian phyla. *Nematostella vectensis* (Nv) is a cnidarian, and *Trichoplax adhaerens* (Ta) is a primitive metazoan. The choanoflagellida *Monosiga brevicollis* (Mb) is a unicellular opisthokont. This species represents one of the closest relative to metazoans whose genome includes a few genes that encode cell adhesion and signaling proteins (King et al. 2008).

The list of analyzed proteins is presented in Table 1. We limited the analysis to proteins that co-localize with acetylcholinesterase and cholinergic receptors, such as MuSK (Karmouch et al. 2013), Lynx (Miwa et al. 2006), MDGA (Pettem et al. 2013), LRP4 and agrin (Zhang et al. 2011), VAChT, and ChAT (Erickson et al. 1994) or with neuroligins and glutamate or GABA/glycine receptors, such as GKAP and PSD-95 (Kim et al. 1997; Irie et al. 1997), Shank (Sheng and Kim 2000), gephyrin (Giannone et al. 2013) collybistin (Kuhse et al. 2012), and MDGA (Pettem et al. 2013). We excluded proteins only composed of multiple repetitive short domains found in many other proteins including non-synaptic proteins.

The deltaBLAST alignment program with search requests using URLAPI-encoded commands and parameters (Boratyn et al. 2012) was used. We retrieved sequences homologous to each of the selected list of protein from GenBank on the NCBI server (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The search was limited to the selected species. The best hit in each species was used to perform recursively deltaBLAST search within all the other species databases. In parallel, all best hits were analyzed on the PFAM database (http://pfam.sanger.ac.uk/) for domain composition (Punta et al. 2012). Raw data are available as Supplementary Material (http://bioweb.ensam.inra.fr/ESTHER/SupData.pl?paper=JMolNeurosci2013).

In the raw Supplementary Data file, the first table shows the list of protein selected for analysis. Links in the table lead directly to the results for the chosen protein. For each protein selected, four tables are provided (Tables S1, S2, S3, and S4
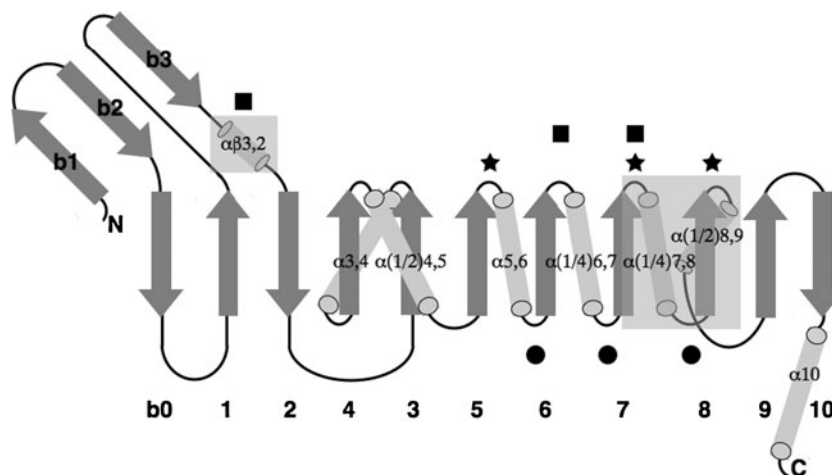
Fig. 1 Topology diagram of the AChE and NL common elements (adapted and labeled from Cygler et al. 1993). The beta-sheets are displayed as *gray arrows* and alpha-helices by *rods*. *Stars* indicate the positions of the catalytic triad residues in AChE: Ser (replaced by Gly in NL) is located after beta5, Glu is after beta 7, and His is after beta 8. *Filled squares* indicate surface regions involved in the binding of non-competitive AChE inhibitors at the active site gorge entry while *filled circles* denote those involved in the binding of neurexin at the NL surface. *Shaded areas* indicate regions of large conformational rearrangements between AChE and NL

indexed with the name of the test sequence), followed by a tree built from the results (Figure S1 with same index). Table S1 shows the list of proteins recovered with best blast hit in the different species. Table S2 shows the name and identification numbers of the reciprocal best hits. Table S3 shows the *e* value of the alignment of the sequence indicated as head of column with the best hit in each species (identification number in equivalent cell from Table S2). Table S3 shows the domain composition of the sequences. The names of domains are colored according to the query sequence (black if domain composition is similar, red if a domain is missing, blue if a domain is added. As for the tree (Figure S1), the name of each branch shows the GenBank identification number and the initials of the species in capital letter (e.g., HS for *Homo sapiens*) followed by the initials of all other species for which a best hit was found (only the order initial is in capital, e.g., Dm for *Drosophila melanogaster*). The branches are colored according to the domain composition relative to the query sequence (black if domain composition is similar, red if a domain is missing, blue if a domain is added, purple if both a domain is missing and a domain is added, and grey if there are no available data in the PFAM database). Finally, subtrees where branches are black and most sequences show multiple species best hits are indicative of orthology.

All the results are summarized in Table 2 (A–C) in the text. Results with *P. troglodyte* are not shown as this species served as control, and orthology is always observed with human genes. The best *e* value is indicated in each cell. Italicized entries indicate strong probability of orthology both from domain composition analysis and strong homology. Entries in bold indicate possible orthology. Those without emphasis indicate absence of homologous protein with similar domain composition.

Results

Place of Neuroligins and Cholinesterases in the COesterase Phylogenetic Tree

Acetylcholinesterase and neuroligins belong to the COesterase family of proteins (PFAM PF00135). A conserved SEDCLYLN motif in the amino-terminal extension (PROSITE: PS00941) is the signature of the family. The family is highly represented only in the animal/fungi embranchment (more specifically in the Opisthokonta) but absent in plants (Pezzementi and Chatonnet 2010). Figure 2 depicts the relationships of proteins of the COesterases family. The figure is a simplified tree. A more detailed tree can be found as Supplementary Material on ESTHER together with the alignment used to build the tree (http://bioweb.ensam.inra.fr/ESTHER/SupData.pl?paper= JMolNeurosci2013). Subtrees are compressed and colored according to known function and taxonomic groups. The level of details is variable, and we annotated the clades of insect and vertebrate carboxylesterases according to previously defined annotations (Claudianos et al. 2006; Teese et al. 2010; Holmes et al. 2010; Gilbert and Auld 2005). The most basal branches are not well supported by bootstrap analysis and cannot be considered as true evolutionary divergences.

A few lines can be drawn from this analysis. Non-catalytic functions (associated to a loss of catalytic activity) appeared independently on several occasions. Glutactins and neurotactins are truly specific proteins of Ecdysozoa (insects, nematodes, etc.). Gliotactins seem to be present in cnidarians, platyhelminths, insects, and nematodes as well as in the hemichordate *Saccoglossus*
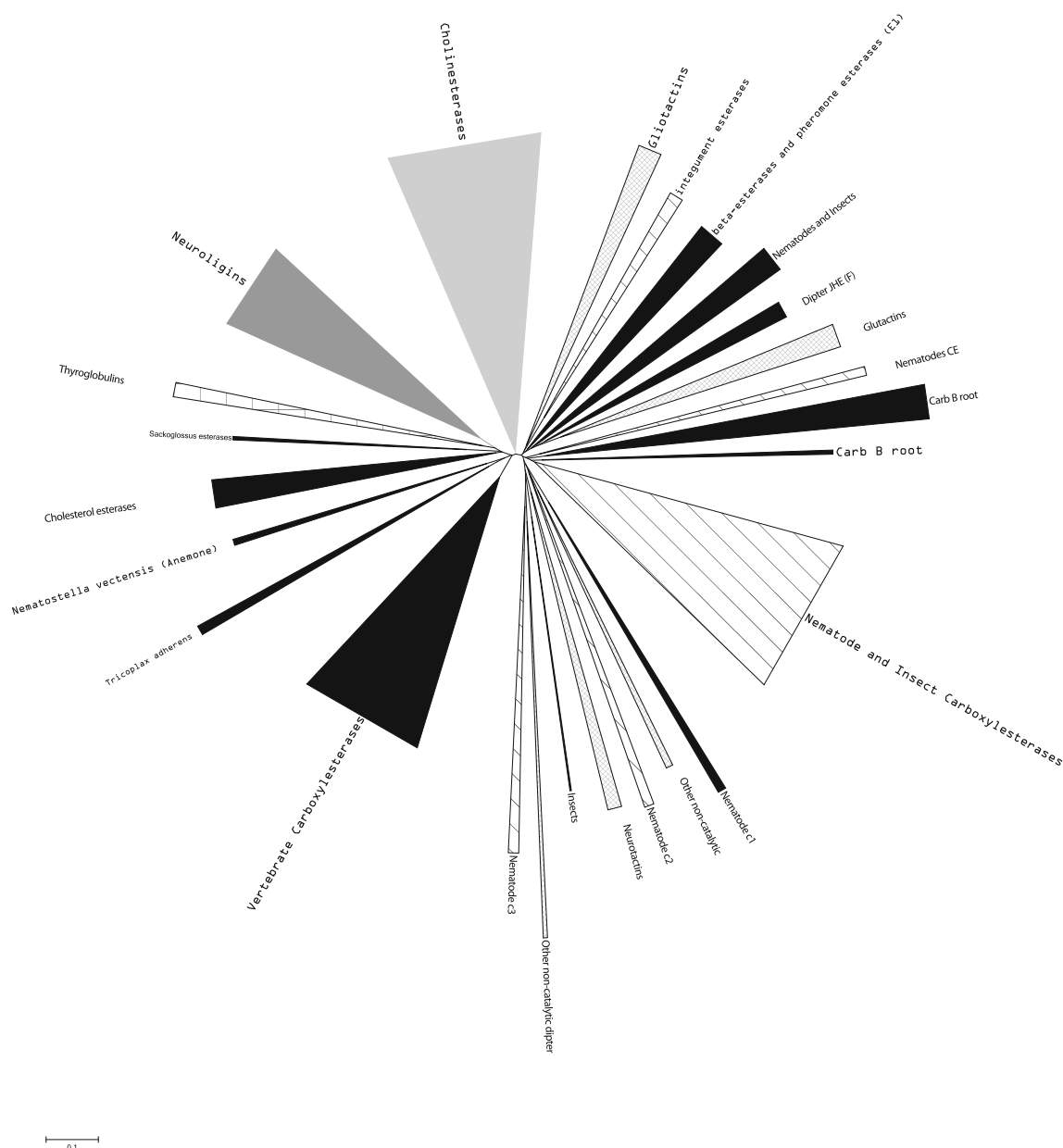
Fig. 2 Phylogenetic tree of the COesterase family (PFAM PF00135). Sequences were aligned with ClustalO (Sievers et al. 2011). The evolutionary history was inferred using the neighbor-joining method (Saitou and Nei 1987). All taxa in a subfamily are condensed as a single triangle. Apparition of Insect non-catalytic proteins such as Gliotactins and Neurotactins (grid triangles) occurred after triploblastic radiation that led to Insects and Nematodes (hatched triangles). Cholinesterases and Neuroligins (light and dark grey triangles) have basal anchoring between whole Carboxylesterases (black triangles) and Insects/Nematodes (hatched triangles). These evolutionary analyses were conducted in MEGA5 (Tamura et al. 2011). The full tree and the alignment are available as Supplementary Materials

kowalevskii but are to be lost in chordates. Cholinesterases and neuroligins from bilaterian animals form homogeneous subtrees, but no sequence from cnidarians or more primitive metazoans can be definitively associated with those subtrees. However, one sequence of *N. vectensis* (RefSeq XP_001631050 UniProt A7SBE0) not only shows homologies to cholinesterases, but it also presents a carboxy-terminal sequence homologous to the H peptide found in bilaterian animals for the posttranslational attachment of the glycolipid anchor that links the protein to the

extracellular membrane. Some other sequences of *Nematostella* are homologous to neuroligins (RefSeq XP_001631073 and XP_001639055; UniProt A7SBD9 and ANV3). These proteins have conserved the catalytic triads of carboxylesterases and possess long carboxy-terminal extensions corresponding to predicted transmembrane and disordered cytoplasmic domains. However, these extensions are not homologous to the carboxy-terminal extensions found in neuroligins from bilaterians (Paz et al. 2008).

Table 1 List of human proteins used as baits to search orthologs in databases of model organism sequences

GenBank	UniProt	Name	Gene
NP_000656.1	P22303|ACES_HUMAN	Acetylcholinesterase	ACHE
NP_000741.1	P30926|ACHB4_HUMAN	Neuronal acetylcholine receptor subunit beta-4	CHRNB4
NP_000730.1	P20309|ACM3_HUMAN	Muscarinic acetylcholine receptor M3	CHRM3
NP_940978.2	O00468|AGRIN_HUMAN	Agrin	AGRIN
NP_001166950.1	O43307|ARHG9_HUMAN	Rho guanine nucleotide exchange factor 9	ARHGEF9
NP_006530.1	O60359|CCG3_HUMAN	Voltage-dependent calcium channel gamma-3 subunit	CACNG3
NP_114102.2	P62955|CCG7_HUMAN	Voltage-dependent calcium channel gamma-7 subunit	CACNG7
NP_001136401.1	P28329|CLAT_HUMAN	Choline *O*-acetyltransferase	CHAT
NP_001119526.1	O14936|CSKP_HUMAN	Peripheral plasma membrane protein CASK	CASK LIN2
NP_001122299.1	P78352|DLG4_HUMAN	Disks large homolog 4	PSD95
NP_001003809.1	O14490|DLGP1_HUMAN	Disks large-associated protein 1	GKAP
NP_000100.2	P11532|DMD_HUMAN	Dystrophin	DMD
NP_000802.2	Q16445|GBRA6_HUMAN	Gamma-aminobutyric acid receptor subunit alpha-6	GABRA6
NP_001019389.1	Q9NQX3|GEPH_HUMAN	Gephyrin	GPHN
NP_741960.1	Q9UQM7|KCC2A_HUMAN	Calcium/calmodulin-dependent protein kinase type II subunit alpha	CAMKA
NP_002325.2	O75096|LRP4_HUMAN	Low-density lipoprotein receptor-related protein 4	LRP4
NP_056379.1	O43300|LRRT2_HUMAN	Leucine-rich repeat transmembrane neuronal protein 2	LRRTM2
NP_076435.1	Q9BZG9|LYNX1_HUMAN	Ly-6/neurotoxin-like protein 1	LYNX1
NP_705691	Q8NFP4|MDGA1_HUMAN	MAM domain-containing glycosylphosphatidylinositol anchor protein	MDGA1
NP_001159752.1	O15146|MUSK_HUMAN	Muscle, skeletal receptor tyrosine-protein kinase	MUSK
NP_055747.1	Q8N2Q7|NLGN1_HUMAN	Neuroligin-1	NLGN1
NP_619635.1	O60391|NMD3B_HUMAN	Glutamate receptor ionotropic, NMDA 3B	GRIN3B
NP_001129131.1	Q9ULB1|NRX1A_HUMAN	Neurexin-1	NRXN1
NP_068751.4	Q9NPQ8|RIC8A_HUMAN	Synembryn-A	RIC8A
NP_068587.1	Q9GZV3|SC5A7_HUMAN	High affinity choline transporter 1	CHT1
NP_057232.2	Q9Y566|SHAN1_HUMAN	SH3 and multiple ankyrin repeat domain protein 1	SHANK1
NP_003046.2	Q16572|VACHT_HUMAN	Vesicular acetylcholine transporter	VACHT

Search for Orthologous Sequences of Selected Synaptic Proteins in Species Representative of Major Transitions of Metazoan Evolution

The results of iterative BLAST and tree construction (low *e* value and reciprocal identical best hits) allowed us to isolate a group of orthologous proteins, for each synaptic participant. These findings are summarized in Table 2 (A) for the cholinergic synapse, Table 2 (B) for the glutamatergic excitatory synapse, and Table 2 (C) for the GABAergic/glycinergic inhibitory synapse (http://bioweb.ensam.inra.fr/esther/PapersSupData/JMolNeurosci2013/Table2.png). Italicized entries indicate a strong probability for the existence of an ortholog in the considered species. A star placed after the protein name denotes absence of strong correlation between sequence similarity and conservation of domain composition. In those cases, the italicized entries are kept only in species where both criteria are valid. For example, in agrin, the SEA, NtA, and EGF domains are missing in the sequence showing the best BLAST hits in non-triploblastic species (even if these domains are found separately in other proteins from almost all metazoan). Indeed,

the best candidates homologous to agrin, described in *C. elegans* (Hrus et al. 2007) or *Drosophila melanogaster* (Husain et al. 2006), do not play any role in synapses. Unexpectedly, our analysis suggests that collybistin and gephyrin may have homologues in cnidarians and early metazoans but were lost in nematodes, insects, and platyhelminths. A similar loss was observed for other human genes (Sullivan and Finnerty 2007).

Compared to the above-mentioned proteins, acetylcholinesterase and neuroligin, starred twice, required a deeper analysis due to the high conservation rate of sequences in proteins of the COesterase family, which results in a low *e* value for candidates in all species. Analysis of the candidate sequences in *N. vectensis*, *T. adhaerens*, or *M. brevicollis* showed that none of these sequences possess hallmarks of a true acetylcholinesterase (choline binding site, aromatic lining of the catalytic gorge) or true neuroligin (putative neurexin binding site).

Our data indicate that the first cholinergic molecules that appeared are muscarinic acetylcholine receptors and choline acetyltransferase, two molecules that indeed could be sufficient for cell-to-cell communication using acetylcholine as a mediator/transmitter. The synaptic proteins that appeared later rendered

Table 2 Probability of the presence of putative orthologs of human synaptic proteins in the selected species

		Metazoa						Placozoa	Choanoflagellida
		Eumetazoa							
		Bilateria					Cnidaria		
		Tripoblastics		Arthropods	Nematodes	Platyhelminths			
		Homo sapiens	Danio rerio	Drosophila melanogaster	Caenorhabditis elegans	Schistosoma mansoni	Nematostella vectensis	Trichoplax adhaerens	Monosiga brevicollis
A. Cholinergic synapse	mAChR	0	0	e-81	e-75	e-42	e-28	**e-22**	1
	ChAT	0	e-156	e-109	e-149	e-136	e-123	**e-71**	0
	LRP-4*	0	0	0	0	e-13	0	**0**	e-7
	CHT	0	e-176	e-162	e-123	e-156	**0**	0	0
	AChE**	0	0	e-118	e-142	e-93	**e-96**	e-97	e-67
	nAChR	0	0	e-106	**e-104**	e-62	**e-76**	1	e-07
	VAChT	0	0	e-147	e-124	**e-149**	e-14	e-13	e-07
	hCASK*	0	0	0	0	**0**	0	0	e-78
	MuSK*	0	0	**e-118**	e-64	e-64	e-78	e-69	e-61
	Agrin*	0	**0**	e-63	e-106	e-63	e-71	e-95	e-4
	SLURP-2	0	3.	>10	>10	7	>10	>10	>10
B. Excitatory synapse	Neurexin*	0	0	0	e-152	e-61	e-29	e-31	**e-17**
	CAMKA	0	0	0	0	0	0	0	**e-166**
	GKAP	0	0	e-17	e-11	e-15	e-23	**e-16**	>10
	Neuroligin**	0	0	e-81	e-58	e-56	**e-82**	e-74	e-40
	LRRTM2*	0	0	e-27	e-25	**e-15**	e-24	e-26	e-27
	NMDAR*	0	0	**e-72**	**e-75**	**e-63**	e-59	e-35	1
	PSD-95*	0	0	0	e-156	**e-97**	e-100	0	e-109
	hCASK*	0	0	**0**	**0**	**0**	0	0	e-75
	MDGA1	0	**0**	e-25	e-27	e-26	e-37	e-33	e-16
	AMPAR & Stargazin	0	**e-170**	e-11	e-7	1	e-1	e-2	1
	Shank1*	0	**0**	e-102	e-91	e-90	e-103	e-10	e-75
C. Inhibitory synapse	Neurexin*	0	0	0	e-152	e-61	e-29	e-31	**e-17**
	Gephyrin*	0	0	e-70	e-77	e-27	e-161	**e-158**	e-47
	GABAR	0	0	e-96	e-82	e-55	**e-78**	>10	e-16
	Collybistin*	0	0	e-47	e-16	e-17	**e-143**	e-100	e-30
	Neuroligin**	0	0	e-81	e-58	e-56	**e-82**	e-74	e-40
	LRRTM2*	0	0	e-27	e-25	**e-15**	e-24	e-26	e-27
	hCASK*	0	0	0	0	**0**	0	0	e-75

Proteins of the cholinergic synapse, proteins of the glutamatergic excitatory synapse, and proteins of the GABAergic/glycinergic inhibitory synapse. The e value of the best BLAST score is indicated in each case. For each protein, cells corresponding to species where an orthologous protein is detected with low e value score and conservation of domain composition are shaded (italicized when both criteria are fulfilled and boldface when the probability is low)

*Proteins for which the two criteria are not fulfilled simultaneously

**Proteins for which information on sequence requirements for function was also used as a marker of orthology

cholinergic transmission more efficiently. Acetylcholinesterase probably evolved simultaneously with choline transporters, an event favoring neurotransmitter recycling. Proteins involved in the efficient functional positioning of receptors, such as agrin, appeared later only, in vertebrates.

Concerning structural or functional proteins of glutamatergic excitatory and GABAergic/glycinergic inhibitory synapses, a few putative orthologs can be retrieved in the *M. brevicollis* genome (LRRTM2, neurexins, PSD-95, CAMKA), a feature indicating that these proteins performed other functions before being recruited in synapses. The fact that neurexins and neuroligins do not appear concomitantly could indicate that their partnering at the synapse was not their primary function.

Discussion

Most basal bilaterians such as the acoel *Symsagittifera roscoffensis* unambiguously possess a cholinergic nervous system with very efficient acetylcholinesterase activity (Bery and Martinez 2011). Cnidarians are considered the most basal metazoans possessing a primitive nervous system. Analysis of the genome of this organism has already shown that most of the neurotransmitters are already present in cnidarians (Anctil 2009), and acetylcholine-hydrolyzing activity was described in *Hydra magnipapillata* (Takahashi and Hamaue 2010) and *Clytia hemisphaerica* (Denker et al. 2008). Apart from vertebrates, neuroligins have been functionally characterized only in a few other classes of bilaterian animals, such as insects (Biswas et al. 2008; Knight et al. 2011), nematodes (Hu et al. 2012), and mollusks (Hu et al. 2012). Quite homologous sequences were also found in platyhelminth genomes. However, neuromediators, receptors, and synapse-specific proteins have been recruited progressively for neurotransmission during metazoan evolution (Buznikov et al. 1999).

Orthologous proteins (in different species that evolved from a common ancestor and perform similar functions) can be detected by sequence similarity search. However, when the reference protein has similarity to not only one but many paralogous proteins, every cross-species protein pair is technically orthologous. Many methods have been devised to cope with this difficulty for large-scale comparison of genomes. For our restricted and focused analysis, our program bypasses this difficulty in selecting only the best blast hit and hence cannot find the true ortholog, yet the level of diversity of chosen species (representing highly divergent branches of metazoan) precludes a sure automatic detection of orthologs. Including information on domain structure and conservation allows us to provide a probability of existence of orthologs or preadapted gene products. Our analysis of acetylcholinesterase and neuroligin shows that knowledge of specific structural characteristics is still necessary to solve these difficulties. Out of the hundreds of proteins that have been found to be localized in synaptic structures, we selected a small number of proteins representing synthesis, binding, or degradation of neurotransmitter or structural proteins located either in the synaptic cleft or the post- or presynaptic compartment. Likewise, we restricted our analysis to a limited number of species. However, our program is available, upon request, to whom might be willing to extend the analysis to other molecules and species.

Conclusions

Compared to other synaptic proteins, true acetylcholinesterase and neuroligins appeared relatively late in the evolution of metazoans. However, two ancestral carboxylesterases probably acquired first proper localization clues to synapses and then evolved toward a high catalytic activity for one of them and the capacity for protein recognition and interaction for the other one. In particular, the candidate genes for neuroligins in cnidarians could still code for proteins with a hydrolytic activity.

References

Anctil M (2009) Chemical transmission in the sea anemone *Nematostella vectensis*: a genomic perspective. Comp Biochem Physiol D Genomics Proteomics 4:268–289

Bery A, Martinez P (2011) Acetylcholinesterase activity in the developing and regenerating nervous system of the acoel *Symsagittifera roscoffensis*. Acta Zool 92:383–392

Biswas S, Russell RJ, Jackson CJ, Vidovic M, Ganeshina O, Oakeshott JG, Claudianos C (2008) Bridging the synaptic gap: neuroligins and neurexin I in *Apis mellifera*. PLoS One 3:e3542

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL (2012) Domain enhanced lookup time accelerated BLAST. Biol Direct 7:12

Buznikov GA, Shmukler YB, Lauder JM (1999) Changes in the physiological roles of neurotransmitters during individual development. Neurosci Behav Physiol 29:11–21

Choi YB, Li HL, Kassabov SR, Jin I, Puthanveettil SV, Karl KA et al (2011) Neurexin-neuroligin transsynaptic interaction mediates learning-related synaptic remodelling and long-term facilitation in *Aplysia*. Neuron 70:468–481

Claudianos C, Ranson H, Johnson RM, Biswas S et al (2006) A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. Insect Mol Biol 15:615–636

Cygler M, Schrag JD, Sussman JL, Harel M, Silman I, Gentry MK, Doctor BP (1993) Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases, and related proteins. Protein Sci 2:366–382

Darboux I, Barthalay Y, Piovant M, Hipeau-Jacquotte R (1996) The structure-function relationships in Drosophila neurotactin show that cholinesterasic domains may have adhesive properties. EMBO J 15:4835–4843

De Jaco A, Comoletti D, Dubi N, Camp S, Taylor P (2012) Processing of cholinesterase-like α/β-hydrolase fold proteins: alterations associated with congenital disorders. Protein Pept Lett 19:173–1799

Denker E, Chatonnet A, Rabet N (2008) Acetylcholinesterase activity in *Clytia hemisphaerica* (Cnidaria). Chem Biol Interact 175:125–128

Dong H, Xiang YY, Farchi N, Ju W, Wu Y, Chen L et al (2004) Excessive expression of acetylcholinesterase impairs glutamatergic synaptogenesis in hippocampal neurons. J Neurosci 24:8950–8960

Erickson JD, Varoqui H, Schäfer MK, Modi W, Diebler MF, Weihe E (1994) Functional identification of a vesicular acetylcholine transporter and its expression from a "cholinergic" gene locus. J Biol Chem 269:21929–21932

Giannone G, Mondin M, Grillo-Bosch D, Tessier B, Saint-Michel E, Czöndör K et al (2013) Neurexin-1β binding to neuroligin-1 triggers the preferential recruitment of PSD-95 versus gephyrin through tyrosine phosphorylation of neuroligin-1. Cell Rep 3:1996–2007

Gilbert MM, Auld VJ (2005) Evolution of CLAMS (cholinesterase-like adhesion molecules): structure and function during development. Front Biosci 10:2177–2192

Grifman M, Galyam N, Seidman S, Soreq H (1998) Functional redundancy of acetylcholinesterase and neuroligin in mammalian neuritogenesis. Proc Natl Acad Sci U S A 95:13935–13940

Holmes RS, Wright MW, Lauderkind SJ, Cox LA, Hosokawa M, Imai T et al (2010) Recommended nomenclature for five mammalian carboxylesterase gene families: human, mouse, and rat genes and proteins. Mamm Genome 21:427–441

Hrus A, Lau G, Hutter H, Schenk S, Ferralli J, Brown-Luedi M et al (2007) C. elegans agrin is expressed in pharynx, IL1 neurons and distal tip cells and does not genetically interact with genes involved in synaptogenesis or muscle function. PLoS One 2:e731

Hu Z, Hom S, Kudze T, Tong XJ, Choi S, Aramuni G et al (2012) Neurexin and neuroligin mediate retrograde synaptic inhibition in C. elegans. Science 337:980–984

Husain N, Pellikka M, Hong H, Klimentova T, Choe KM, Clandinin TR et al (2006) The agrin/perlecan-related protein eyes shut is essential for epithelial lumen formation in the drosophila retina. Dev Cell 11: 483–493

Irie M, Hata Y, Takeuchi M, Ichtchenko K, Toyoda A, Hirao K, Takai Y et al (1997) Binding of neuroligins to PSD-95. Science 277:1511–1515

Johnson G, Moore SW (2012a) The carboxylesterase/cholinesterase gene family in invertebrate deuterostomes. Comp Biochem Physiol D Genomics Proteomics 7:83–93

Johnson G, Moore SW (2012b) Why has butyrylcholinesterase been retained? Structural and functional diversification in a duplicated gene. Neurochem Int 61:783–797

Karmouch J, Dobbertin A, Sigoillot S, Legay C (2013) Developmental consequences of the ColQ/MuSK interactions. Chem Biol Interact 203:287–291

Kim YH, Lee SH (2013) Which acetylcholinesterase functions as the main catalytic enzyme in the Class Insecta? Insect Biochem Mol Biol 43:47–53

Kim E, Naisbitt S, Hsueh YP, Rao A, Rothschild A, Craig AM et al (1997) GKAP, a novel synaptic protein that interacts with the guanylate kinase-like domain of the PSD-95/SAP90 family of channel clustering molecules. J Cell Biol 136:669–678

Kim YH, Cha DJ, Jung JW, Kwon HW, Lee SH (2012) Molecular and kinetic properties of two acetylcholinesterases from the western honey bee, Apis mellifera. PLoS One 7:e48838

King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J et al (2008) The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. Nature 451:783–788

Knight D, Xie W, Boulianne GL (2011) Neurexins and neuroligins: recent insights from invertebrates. Mol Neurobiol 44:426–440

Krejci E, Duval N, Chatonnet A, Vincens P, Massoulie J (1991) Cholinesterase-like domains in enzymes and structural proteins: functional and evolutionary relationships and identification of a catalytically essential aspartic acid. Proc Natl Acad Sci U S A 88: 6647–6651

Kuhse J, Kalbouneh H, Schlicksupp A, Mükusch S, Nawrotzki R, Kirsch J (2012) Phosphorylation of gephyrin in hippocampal neurons by cyclin-dependent kinase CDK5 at Ser-270 is dependent on collybistin. J Biol Chem 287:30952–30966

Lenfant N, Hotelier T, Bourne Y, Marchot P, Chatonnet A (2013a) Proteins with an alpha/beta hydrolase fold: relationships between subfamilies in an ever-growing superfamily. Chem Biol Interact 203:266–268

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013b) ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins: tools to explore diversity of functions. Nucleic Acids Res 41:D423–D429

Leone P, Comoletti D, Taylor P, Bourne Y, Marchot P (2010) Structure-function relationships of the alpha/beta-hydrolase fold domain of neuroligin: a comparison with acetylcholinesterase. Chem Biol Interact 187:49–55

Marchot P, Chatonnet A (2012) Enzymatic activity and protein interactions in alpha/beta hydrolase fold proteins: moonlighting versus promiscuity. Protein Pept Lett 19:132–143

Miwa JM, Stevens TR, King SL, Caldarone BJ, Ibanez-Tallon I, Xiao C, Fitzsimonds RM et al (2006) The prototoxin lynx1 acts on nicotinic acetylcholine receptors to balance neuronal activity and survival in vivo. Neuron 51:587–600

Olson PF, Fessler LI, Nelson RE, Sterne RE, Campbell AG, Fessler JH (1990) Glutactin, a novel Drosophila basement membrane-related glycoprotein with sequence similarity to serine esterases. EMBO J 9: 1219–1227

Paz A, Zeev-Ben-Mordehai T, Lundqvist M, Sherman E, Mylonas E, Weiner L et al (2008) Biophysical characterization of the unstructured cytoplasmic domain of the human neuronal adhesion protein neuroligin 3. Biophys J 95:1928–1944

Pettem KL, Yokomaku D, Takahashi H, Ge Y, Craig AM (2013) Interaction between autism-linked MDGAs and neuroligins suppresses inhibitory synapse development. J Cell Biol 200:321–336

Pezzementi L, Chatonnet A (2010) Evolution of cholinesterases in the animal kingdom. Chem Biol Interact 187:27–33

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C et al (2012) The Pfam protein families database. Nucleic Acids Res 40: D290–D301

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Schulte J, Tepass U, Auld VJ (2003) Gliotactin, a novel marker of tricellular junctions, is necessary for septate junction development in Drosophila. J Cell Biol 161:991–1000

Sheng M, Kim E (2000) The Shank family of scaffold proteins. J Cell Sci 113:1851–1856

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W et al (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539

Silman I, Sussman JL (2005) Acetylcholinesterase: 'classical' and 'non-classical' functions and pharmacology. Curr Opin Pharmacol 5:293–302

Soreq H, Seidman S (2001) Acetylcholinesterase—new roles for an old actor. Nat Rev Neurosci 2:294–302

Sullivan JC, Finnerty JR (2007) A surprising abundance of human disease genes in a simple "basal" animal, the starlet sea anemone (Nematostella vectensis). Genome 50:689–692

Takács VT, Freund TF, Nyiri G (2013) Neuroligin 2 is expressed in synapses established by cholinergic cells in the mouse brain. PLoS One 8:e72450

Takahashi T, Hamaue N (2010) Molecular characterization of Hydra acetylcholinesterase and its catalytic activity. FEBS Lett 584:511–516

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Teese MG, Campbell PM, Scott C, Gordon KH, Southon A, Hovan D et al (2010) Gene identification and proteomic analysis of the esterases of the cotton bollworm, Helicoverpa armigera. Insect Biochem Mol Biol 40:1–16

Zhang W, Coldefy AS, Hubbard SR, Burden SJ (2011) Agrin binds to the amino-terminal region of Lrp4 protein and stimulates association between Lrp4 and the first immunoglobulin-like domain in muscle-specific kinase (MuSK). J Biol Chem 286:40624–40630

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97–166