A phylogenetic analysis of MCTP proteins: from amino acid sequence to function

José Luis Téllez Arreola¹, Argel Estrada-Mondragón², Ataúlfo Martínez Torres^{1*}.

¹ Departamento de Neurobiología Celular y Molecular, Instituto de Neurobiología, Universidad Nacional Autónoma de México, Campus Juriquilla, Boulevard Juriquilla 3001, C.P. 76215 Juriquilla, Querétaro, México.

² Department of Clinical and Experimental Medicine, Linköping University, SE-58185 Linköping, Sweden

* Corresponding author. E-mail address ataulfo@unam.mx (A. Martínez-Torres).

Acknowledgments

This work was supported by CONACYT A1-S-7659 to AMT. José Luis Téllez Arreola is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received fellowship 395834 from CONACYT and Fulbright-Garcia Robles (COMEXUS). Jessica González Norris kindly edited the manuscript.

* Authors have no conflict of interest to declare.

Abstract

MCTPs (Multiple C2 domain proteins with two transmembrane regions) are evolutionarily and structurally related to other C2 proteins which play fundamental roles in exocytosis and membrane trafficking, however their specific role has been little studied. This work points out possible functional implications of MCTPs by comparing their primary amino acid sequence and functional domains. MCTP amino acid sequences were identified in non-chordates and chordates. The primary sequences grouped in three classes: MCTP, MCTP-1 and MCTP-2. MCTP is present only in non-chordates, while MCTP-1 and MCTP-2 are present in chordates. MCTP genes emerged early in metazoan evolution and are well conserved across species including humans. Genomic analysis of diverse species of representative phyla showed that the three C2 domains (C2A-C2C) and transmembrane regions (TMR) are well conserved. The C2 domains have eight β strands as well as aspartate residues known to bind calcium. Interestingly, we identified a lysine-rich cluster, also known as polybasic cluster in C2A and C2B, which is known to bind lipids in other proteins. We also describe the phylogenetic distribution of MCTPs and analyze conserved domains and their predicted secondary structure in metazoans. We highlight important motifs that have not been previously described in MCTPs C2A and C2B domains that suggest MCTPs potentially bind phospholipids. Our observations show MCTPs are proteins widely distributed in eukaryotic organisms and may play an important role in membrane fusion or exocytosis.

Keywords: C2 domain; calcium signaling; lipid binding; membrane traffic; protein targeting; membrane tethering

Introduction

Cellular signaling through proteins that possess calcium-binding domains is fundamental for multiple processes such as exocytosis, vesicular trafficking, protein kinase and phospholipase activities, among many others. The C2 domain is present in many proteins involved in synaptic transmission such as synaptotagmins (Bagur and Hajnóczky, 2017; Evans et al., 2016; Nalefski and Falke, 1996; Shin et al., 2005). This domain was originally described as the second of four calcium-binding domains of the calcium-dependent protein kinase C (Nalefski and Falke, 1996); however, it is now well known that this domain is widespread in diverse eukaryotic proteins that could act or not as calcium sensors, triggering diverse cellular responses (Bagur and Hajnóczky, 2017). Molecular and functional studies have revealed the role of proteins that possess C2 domains in different signaling cascades; nevertheless, multiple questions remain. C2 proteins involved in membrane trafficking, such as synaptotagmins and ferlins, include a transmembrane region and at least two C2 domains whose calcium affinity is higher in the presence of phospholipids (Rizo and Sudhof, 1998). A novel family of C2 proteins (MCTPs) was discovered by *in silico* genomic sequence analyses (Shin et al., 2005). They present three conserved C2 domains (C2A, C2B and C2C) facing the cytoplasm, one- or two-transmembrane regions that anchor the protein to intracellular vesicles and one short C-terminus sequence. Vertebrates have two MCTP coding genes (*MCTP1* and *MCTP2*). While, invertebrates such as *Drosophila melanogaster* and *Caenorhabditis elegans* have only one MCTP gene (Shin et al., 2005). The cellular distribution of MCTPs remains little known. They appear to be located in intracellular vesicles and endoplasmic reticulum, which are a suitable positions for sensing changes in calcium concentration and to play a role in vesicular trafficking (Shin et al., 2005), yet their functions are not known.

Assessment with selective antibodies for each isoform proved that MCTPs are widely distributed in vertebrates, especially in striated muscle and heart (Qiu et al., 2015; Shin et al., 2005). This is interesting since a 2.2 Mb deletion in the 15q26.2 chromosome, in which the *MCTP2* gene is located, causes coarctation of the aorta and hypoplastic heart syndrome (Lalani et al., 2013). Furthermore, the

importance of the *MCTP2* gene for proper cardiogenesis in frogs, was proved by knocking down the expression of the gene with morpholinos, which yielded embryos with a heart condition resembling coarctation of the aorta (Lalani et al., 2013). In other studies, single nucleotide polymorphisms in the *MCTP2* gene were associated with schizophrenia and bipolar disorders, yet its role in evoking these psychiatric diseases has been elusive (Djurovic et al., 2009). In 2015, a study described the distribution of the MCTP-1 in rats. It showed that MCTP-1 is widely distributed in the brain and is found in secretory vesicles and endosomes of neurons (Qiu et al., 2015). More recently, it was found that human MCTP2 transiently expressed in COS7 cells is located in the endoplasmic reticulum and peroxisomes (Joshi et al., 2018).

Studies of the role of MCTPs have also been carried out in invertebrates. Only one gene has been identified in the genome of *D. melanogaster*, and a gene-trap study showed that it is expressed in the accessory cells of the olfactory organ. A P-transposon insertion at the 3' end of the non-coding region of the MCTP gene induces high lethality in the fly (Tunstall et al., 2012). In contrast, other allelic mutants of the MCTP gene are viable (Genç et al., 2017) and this evidence allowed the authors to suggest that this protein is necessary for presynaptic homeostatic plasticity and to set off the baseline for neurotransmitter release. This study also established that MCTP is located in the endoplasmic reticulum of motoneurons, where it senses changes in calcium concentrations, thus facilitating neurotransmitter release. But, further evidence is required to conciliate these contrasting evidences.

In *C. elegans* the study of MCTP is more limited. Two isoforms are predicted to be expressed by the use of two alternative promoters, but the gene expression pattern or cellular localization of the protein has not been determined. Nevertheless, a massive antisense RNAi screening disclosed that when expression of MCTP is knocked down (Maeda et al., 2001), larval lethality is high, consistent with the study of Tunstall et al. (2012) in *D. melanogaster* and suggesting the relevance of this gene for proper development.

Much has been learned about the structure and role of proteins that include the C2 domain, originally identified in the PKC proteins (Nishizuka, 1988). The structure of this domain was revealed by high-resolution X-ray diffraction of synaptotagmin crystals, a family of proteins involved in synaptic transmission (Alwarawrah et al., 2017; Corbalan-Garcia and Gómez-Fernández, 2014; Evans et al., 2016; Sutton et al., 1995). C2s are conformed by approximately 130 amino acids and include aspartate residues that coordinate calcium binding. MCTPs are structurally related to synaptotagmins, extended-synaptotagmins (Esys) and ferlins (Min et al., 2007), proteins that possess two or six C2 domains in tandem, respectively. Nevertheless, the function of the C2 domains in the MCTPs as well as the functional role of these proteins are not fully understood.

In this study, we present a phylogenetic analysis of the MCTP protein family, aiming to shed light on the structural characteristics of this emerging family of calcium sensors.

1 **Methods**

2 **Identification of the MCTP protein sequences.**

3 All the sequences for MCTP proteins were obtained by using SMART (Schultz et al., 1998) and
 4 applying the *C. elegans* MCTP protein sequence (**NP491908.2**) as template. The presence of three C2
 5 domains in their primary sequence and the predicted transmembrane regions were the main criteria to
 6 identify MCTP sequences in other species.

8 **Multiple sequence alignment**

9 Sequence alignments were performed in COBALT using default parameters. COBALT uses multiple
 10 alignment of sequences, a conserved domain database, protein motif databases and local sequence
 11 similarity using RPS-BLAST, BLASTP and PHI-BLAST (Papadopoulos and R., 2007). Jalview was
 12 used to color sequence alignments derived from COBALT results. Zappo color schemes were used to
 13 classify amino acids according to their physicochemical properties (Waterhouse et al., 2009).

15 **MCTP phylogenetic analysis**

16 We used multiple sequences obtained by COBALT as input to generate phylogenetic trees, to analyze
 17 the spread of MCTP proteins in all different species. The phylogenetic tree of experimental model
 18 organisms and C2 domains tree were generated by using muscle alignment program, maximum
 19 likelihood method and Bayesian inference (M.Bayes3.2) or bootstrapping (100 replicates) within
 20 phylogeny.fr platform as needed.

21

22

23

24

25

1 MCTP secondary structure prediction and structural modeling

2 We used PSIPRED software for secondary structure prediction, and RaptorX Structure Prediction
3 software for modelling the tertiary structure of MCTP from *C. elegans* (NP491908.2) using the default
4 parameters of the server (Jones, 1999; Källberg et al., 2012). The automatically generated pdb model
5 was manually cropped in Molsoft ICM-Pro v3.8-3 keeping the integrity of all the predicted domains in a
6 common configuration. One calcium molecule was loaded from ChemSpider into the ICM project
7 (CSID: 4573905), and fitted in the MCTP C2A domain using as guide the structure of the C2A domain
8 of Synaptotagmin-7 (PDB: 6ANJ, Voleti et al., 2017) using the function Pocket Finder of ICM (Dey and
9 Chen, 2011). The final model was submitted to Yasara energy minimization server (Krieger et al., 2009)
10 and the images generated in PyMOL v2.2.2 Schrodinger LLC (Schrödinger, San Diego, CA).

12 Results

13 MCTPs are found in metazoan but not in unicellular eukaryotes

14 MCTP protein sequences were found to be widespread in metazoans, from nematodes to humans. We
15 found no evidence of MCTPs in prokaryotic organisms or unicellular eukaryotes. We generated an
16 evolutionary tree of MCTPs from different organisms that are used as experimental biological models
17 including the sea squirt *Ciona intestinalis*, a tunicate that represents the base line of chordates (Fig. 1).
18 The tree analysis shows the divergence of the MCTP protein family into three major groups, MCTPs
19 class1 and class 2 are present in chordates such as bony fish, reptiles, amphibians, birds and mammals
20 (rodents, ungulates and primates including humans, Figure 1, and supplementary list). The clas MCTP
21 was found only in invertebrates (nematodes, flies and mollusks) including parasitic and free-living
22 nematodes.

1 Analysis and modeling of the conserved functional domains of the MCTPs.

2 As previously mentioned, MCTP proteins possess three cytosolic C2 domains in tandem and two
3 transmembrane regions (Fig. 3A). We aligned the amino acid sequences corresponding to individual C2
4 domains (C2A, C2B and C2C) from the different three classes of MCTPs as well as their C-terminus to
5 determine the degree of conservation of the amino acid sequence of the transmembrane regions (Fig 3B-
6 G). A comparison of the C2 domains revealed a high conservation in the primary sequence, although the
7 degree of conservation is different among them. In general, the C2C domain is the most conserved,
8 followed by C2A and finally C2B; however, calcium binding sites are more conserved in C2A than in
9 C2C (Fig. 3 B-D). The number and position of the aspartate residues that potentially bind calcium in the
10 MCTPs was analyzed and compared to the C2A of Synaptotagmin 1 (**P21579.1**), and we found that the
11 position of five of these residues are partially conserved in the three C2 domains of MCTP. For example,
12 the aspartate residues 1 to 4 are conserved in the C2C domain, whereas the fifth presents the substitution
13 D → E in *C. elegans*, *H. contortus*, *D. melanogaster* and *C. intestinalis* (Fig 3D); furthermore C2B
14 exhibits only four aspartate residues instead of five and all of them show diverse amino acid
15 substitutions, including D → A, E, H, N or T (Fig. 3C). Other MCTP2-C2Bs from *C. elegans*,
16 *haemonchus contortuos*, *Xenopus leavis*, *Xenopus tropicalis* and some MCTP2 isoforms in human and
17 mice exhibit only one or two aspartate residues, thus suggesting that they may not bind calcium (Fig.
18 3B-D). Amino acid residues predicted to bind calcium are well conserved in C2A and C2C in all
19 analyzed MCTPs analyzed but remarkably that is not the case for C2B.

20

21 In C2A and C2B domains of the MCTPs, a lysine-rich domain previously named the polybasic cluster
22 was detected (Corbalan-Garcia and Gómez-Fernández, 2014; Di Paolo and De Camilli, 2006). This
23 domain is predicted by the presence of positive charged (K, K, K) and aromatic (Y, W) residues (Figs. 3
24 F-G, 4). The polybasic cluster is responsible for binding phosphatidylserine and PtdIns(4,5)P₂ (PIP₂) in
25 calcium-binding domains of C2 proteins such as Syt1 (Corbalan-Garcia and Gómez-Fernández, 2014).

1 This presence of lysine-rich domain in the C2A and C2B of MCTP has opened the door for a working
2 hypothesis to determine whether phospholipids could be necessary for calcium binding, which contrasts
3 with the initial observations that suggested that MCTP function is independent of lipid interactions (Shin
4 et al., 2005). Altogether, we found at least two potential PIP₂ binding sites that resemble other C2
5 domains known to penetrate the plasma membrane, these PIP₂ binding sites, in coordination with the
6 aspartates present in C2A, C2B and C2C can potentially stabilize a calcium molecule (Fig. 3A-D). On
7 the other hand, the SMART software yielded a model that suggests that the C2 domains of MCTP spans
8 around 125 residues, whereas PSIPRED suggests that the secondary structural of the C2 domains are
9 eight β -strands, that is a characteristic of the C2 domain conformation Class 2 (topology-II) in which the
10 N and C terminus are near to the bottom (Figs. 3 A, 4).

11 A second criteria to establish that a protein belongs to MCTPs was the presence of two stretches of
12 hydrophobic residues that form the transmembrane region (TMR) that allows the protein to anchor to
13 cell vesicles and/or the plasma membrane. It has been shown that only one segment suffices to anchor
14 the protein to intracellular vesicles (Shin et al 2005), although is unknown whether missing one of the
15 TMRs misdirects the trafficking of the protein. The amino acid sequences of the TMRs of MCTPs were
16 analyzed by SMART. MCTP-2 includes two TMRs located towards the C-terminus, with exception of
17 the isoform 2 from *H. sapiens* MCTP-2, which lacks one TMR (Fig. 3E). The amino acid sequence of
18 these TMRs is not related to other known C2 proteins involved in membrane trafficking that also include
19 in their structure TMRs, such as synaptotagmins. In MCTPs, the second TMR is well conserved in
20 different phyla, whereas the sequence of the first TMR is dispersed. Functional analysis using GFP-
21 tagged versions of mammalian MCTPs previously revealed that both TMRs are functional in transfected
22 cells (Shin et al., 2005) and direct the protein to intracellular vesicles, suggesting that one TMR suffices
23 to anchor the protein in vesicles.

24

25

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Divergence of the C2 domain

We explored whether the amino acid sequence of C2 domains of MCTPs vary among phyla. Alignments of all three C2 domains from representative species of each clade were used to generate maximum likelihood trees. We observed three general clades: the first included the C2C domain from all species (*C. elegans*, *H. contortus*, *C. intestinalis*, *X. leavis*, *X. tropicalis*, *D. rerio*, *M. musculus* and *H. sapiens*). The second and major clade includes C2A and C2B, thus suggesting that they may have similar functions, at least as suggested by their primary amino acid sequence. Interestingly, both C2As and C2Bs possess the putative lipid binding residues included within the lysine-rich domain disclosed by our analysis, in sharp contrast, C2Cs do not have this domain, pointing out to a structural feature that suggests different functions (Fig. 2).

Discussion

In the present study we provide an evolutionary analysis of the MCTP proteins. By definition, MCTPs are composed of a variable N-terminal region, three C2 domains, one or two TMRs and one short C-terminal sequence (Shin et al., 2005). We used this approach to determine peculiarities in the amino acid sequence of the MCTPs to formulate a working hypothesis to try to understand their functional role. In plants sixteen like-MCTPs classes were identified, however most of them have more than three C2 domains and three to four TMRs (Liu et al., 2017). These features are more related to other C2 proteins such as E-syts and tricalbins (Fernandez-Busnadiego et al., 2015; Saheki and De Camilli, 2017), thus, these proteins were not included in our analysis. In this work, we suggest the existence of three classes of MCTPs: MCTP, MCTP-1 and MCTP-2, that emerged from a common ancestor. Their phylogenetic distribution is disperse but different. MCTP-1 and MCTP-2 are present only in vertebrates, (bony fish, snakes, birds, rodents, ungulates and primates), whereas MCTP was observed only in invertebrates (Fig. 1). The calcium sensor proteins synaptotagmins are central molecular components of the

1 neurotransmitter exocytotic machinery in the presynaptic terminal. They appeared early in the evolution
2 of eukaryotes, and evidence shows that they are present even in primitive forms of eukaryotic
3 multicellular organisms such as the placozoans (Barber et al., 2009; Lek et al., 2012; Washington and
4 Ward, 2006). In contrast, MCTPs were not identified using the structural criteria listed above.
5 Unicellular organisms present C2 proteins that have from one to seven C2 domains, including PKCs,
6 ferlins and tricalbins (Barber et al., 2009; Lek et al., 2010; Schulz and Creutz, 2004). Thus, it can be
7 suggested that MCTPs emerged as an adaptation in metazoans, but we do not know yet what role they
8 may play. To understand how the proteins function in nature and how they evolve, it is important to
9 explore shared similarities from amino acid sequences in different species to identify paralogs and
10 homologs (Bolker, 2014; Lek et al., 2010; Train et al., 2017). We also identified highly conserved
11 regions (C2 domains and TMRs) and found that the sequences of the C2B and TMR1 domains have
12 more amino acid substitutions compared to the C2A,C2C an TMR2.

13 X-ray diffraction analysis indicated that C2 domains are composed of eight-stranded β -sandwich with
14 calcium-binding sites coordinated by aspartate residues (Rizo and Sudhof, 1998; Sutton et al., 1995).
15 We analyzed the sequences of the three MCTP C2 domains to try to understand the structural array of
16 these domains comparing them with the C2A and C2B domains of Synaptotagmin-1. The secondary
17 structure is highly conserved in all sequences from selected organisms. Our results are similar to those
18 reported by Shin et al. (2005) who showed that several organisms lack the calcium-binding aspartates
19 in the C2B domain. Interestingly, we also identified that *C. elegans*, *H. contortus*, *X. laevis*, *X.*
20 *tropicalis*, *M. musculus* and *H. sapiens* lack two calcium-binding sites in the C2B, this suggests that
21 this domain does not bind calcium in all species. MCTPs have been described as a novel family of
22 proteins that can bind calcium independently of phospholipid complexes (Shin et al., 2005); however,
23 we identified that C2A and C2B have a lysine-rich domain (a polybasic cluster) or the β -groove motif
24 previously identified in the synaptotagmin-1 C2B domain (**P21579.1**) (Corbalan-Garcia and Gómez-
25 Fernández, 2014). Some studies have demonstrated that this motif is important for interactions with

inositol polyphosphates (IP₄, IP₅ and IP₆) (Kojima, 1995) and PIP₂, which are directly implicated in exocytosis, vesicle fusion, membrane trafficking, cytoskeleton dynamics and membrane permeability, as well as in the transport of membrane channels (Paolo and Camilli, 2006). It seems that this motif acts as a signal for C2 domain proteins to target membranes. It also suggests that MCTPs could bind different phospholipids from those reported by Shin et al. (2005). In the structural model here presented at least two potential binding sites for PIP₂ in the polybasic clusters are disclosed; this opens the possibility that upon PIP₂ binding these domains bring closer the C2 domains with the TMR. This requires further experimental evidence. Meanwhile the question remains as to whether MCTPs bind or not phospholipids.

Regarding the TMRs, we identified two of these domains in all analyzed sequences, except one human isoform that possess one single TMR. This isoform is probably the result of an alternative splicing event in exon 17 in the *MCTP-2* gene. The functional difference between MCTPs with one or two TMRs is currently not well understood; both have three C2 domains, thus they have the same potential to bind calcium and may play similar and/or complementary roles in the cell.

Thus, this differences could be essential for distinguishing different classes of MCTPs, and it might be related with the specialized function of each C2 domain such as in ferlins and Syts (Barber et al., 2009; Washington and Ward, 2006).

Conclusions and future perspectives

In summary, we described the distribution of MCTPs in different phyla, analyzed in detail highly conserved regions and described their secondary structure. Our studies show that these proteins are widespread in metazoans and, because of their selective location in the endoplasmic reticulum and endocytic/exocytic pathways, they may play important roles in processes such as vesicle fusion, exocytosis and membrane trafficking, such as those of other C2 domain proteins. Finally, our analysis

discloses unique characteristics of MCTPs that allows speculation on their function in phospholipid and calcium signaling and suggests a working hypothesis to test their function.

Bibliography

- Alwarawrah, M., Wereszczynski, J., Qiu, X., Ge, J., Gao, Y., Teng, M., Niu, L., 2017. Investigation of the Effect of Bilayer Composition on PKC α -C2 Domain Docking Using Molecular Dynamics Simulations. *J. Phys. Chem. B* 121, 78–88.
- Bagur, R., Hajnóczky, G., 2017. Intracellular Ca²⁺ Sensing: Its Role in Calcium Homeostasis and Signaling. *Mol. Cell* 66, 780–788.
- Barber, C.F., Jorquera, R. a, Melom, J.E., Littleton, J.T., 2009. Postsynaptic regulation of synaptic plasticity by synaptotagmin 4 requires both C2 domains. *J. Cell Biol.* 187, 295–310.
- Bolker, J. a, 2014. Model species in evo-devo: a philosophical perspective. *Evol. Dev.* 16, 49–56.
- Corbalan-Garcia, S., Gómez-Fernández, J.C., 2014. Signaling through C2 domains: More than one lipid target. *Biochim. Biophys. Acta - Biomembr.* 1838, 1536–1547.
- Desper, R., Gascuel, O., 2004. Theoretical Foundation of the Balanced Minimum Evolution Method of Phylogenetic Inference and Its Relationship to Weighted Least-Squares Tree Fitting. *Mol. Biol. Evol.* 21, 587–598.
- Dey, R., Chen, L., 2011. In search of allosteric modulators of $\alpha 7$ -nAChR by solvent density guided virtual screening. *J Biomol Struct Dyn* 28, 695–715.
- Di Paolo, G., De Camilli, P., 2006. Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443, 651–657.

1 Djurovic, S., Le Hellard, S., Kähler, A.K., Jönsson, E.G., Agartz, I., Steen, V.M., Hall, H., Wang, A.G.,
2 Rasmussen, H.B., Melle, I., Werge, T., Andreassen, O. a, 2009. Association of MCTP2 gene
3 variants with schizophrenia in three independent samples of Scandinavian origin (SCOPE).
4 Psychiatry Res. 168, 256–8.

5 Evans, C.S., He, Z., Bai, H., Lou, X., Jeggle, P., Sutton, R.B., Edwardson, J.M., Chapman, E.R., 2016.
6 Functional analysis of the interface between the tandem C2 domains of synaptotagmin-1. Mol.
7 Biol. Cell 27, 979–989.

8 Fernandez-Busnadiego, R., Saheki, Y., De Camilli, P., 2015. Three-dimensional architecture of
9 extended synaptotagmin-mediated endoplasmic reticulum-plasma membrane contact sites. Proc
10 Natl Acad Sci U S A 112.

11 Genç, Ö., Dickman, D.K., Ma, W., Tong, A., Fetter, R.D., Davis, G.W., 2017. MCTP is an ER-resident
12 calcium sensor that stabilizes synaptic transmission and homeostatic plasticity. Elife 6, 1–23.

13 Grishin, N. V, 1995. Estimation of the number of amino acid substitutions per site when the
14 substitution rate varies among sites. J. Mol. Evol. 41, 675–679.

15 Jones DT. (1999) Protein secondary structure prediction based on position-specific scoring matrices. J.
16 Mol. Biol. 292: 195-202.

17 Joshi, A.S., Nebenfuehr, B., Choudhary, V., Satpute-Krishnan, P., Levine, T.P., Golden, A., Prinz,
18 W.A., 2018. Lipid droplet and peroxisome biogenesis occur at the same ER subdomains. Nat.
19 Commun. 9, 2940.

20 Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., Xu, J., 2012. Template-based protein
21 structure modeling using the RaptorX web server. Nat. Protoc. 7, 1511–1522.

22 Kojima, T., 1995. Functional Diversity of C2 Domains of Synaptotagmin Family. J. Biol. Chem. 270,
23 26523–26527.

1 Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., Karplus, K., 2009.
2 Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling:
3 Four approaches that performed well in CASP8. *Proteins*. 77, 114-22

4 Lalani, S.R., Ware, S.M., Wang, X., Zapata, G., Tian, Q., Franco, L.M., Jiang, Z., Bucasas, K., Scott,
5 D. a, Campeau, P.M., Hanchard, N., Umaña, L., Cast, A., Patel, A., Cheung, S.W., McBride, K.L.,
6 Bray, M., Craig Chinault, A., Boggs, B. a, Huang, M., Baker, M.R., Hamilton, S., Towbin, J.,
7 Jefferies, J.L., Fernbach, S.D., Potocki, L., Belmont, J.W., 2013. MCTP2 is a dosage-sensitive
8 gene required for cardiac outflow tract development. *Hum. Mol. Genet.* 22, 4339–48.

9 Lek, A., Evesson, F.J., Sutton, R.B., North, K.N., Cooper, S.T., 2012. Ferlins: Regulators of Vesicle
10 Fusion for Auditory Neurotransmission, Receptor Trafficking and Membrane Repair. *Traffic* 13,
11 185–194.

12 Lek, A., Lek, M., North, K.N., Cooper, S.T., 2010. Phylogenetic analysis of ferlin genes reveals ancient
13 eukaryotic origins. *BMC Evol. Biol.* 10, 231.

14 Liu, L., Li, C., Liang, Z., Yu, H., 2017. Characterization of Multiple C2 Domain and Transmembrane
15 Region Proteins in Arabidopsis. *Plant Physiol.* pp.01144.2017.

16 Maeda, I., Kohara, Y., Yamamoto, M., Sugimoto, A., 2001. Large-scale analysis of gene function in
17 *Caenorhabditis elegans* by high-throughput RNAi. *Curr. Biol.* 11, 171–176.

18 Min, S.W., Chang, W.P., Sudhof, T.C., 2007. E-Syts, a family of membranous Ca²⁺-sensor proteins
19 with multiple C2 domains. *Proc Natl Acad Sci U S A* 104, 3823–3828.

20 Nalefski, E.A., Falke, J.J., 1996. The C2 domain calcium-binding motif: Structural and functional
21 diversity. *Protein Sci.* 5, 2375–2390.

22 Nishizuka, Y., 1988. The molecular heterogeneity of protein kinase C and its implications for cellular
23 regulation. *Nature* 334, 661–665.

1 Papadopoulos, J.S., R., A., 2007. Sequence analysis COBALT : constraint-based alignment tool for
2 multiple protein sequences. *Bioinformatics* 23, 1073–1079.

3 Qiu, L., Yu, H., Liang, F., 2015. Multiple C2 domains transmembrane protein 1 is expressed in CNS
4 neurons and possibly regulates cellular vesicle retrieval and oxidative stress. *J. Neurochem.* 135,
5 492–507.

6 Rizo, J., Sudhof, 1998. Ca²⁺ -binding Domain. *J. Biol. Chem.* 273, 15879–15882.

7 Saheki, Y., De Camilli, P., 2017. The Extended-Synaptotagmins. *Biochim. Biophys. Acta - Mol. Cell*
8 *Res.* 1864, 1490–1493.

9 Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing
10 phylogenetic trees. *Mol Biol Evo* 4, 406–425.

11 Schultz, J., Milpetz, F., Bork, P., Ponting, C.P., 1998. SMART, a simple modular architecture research
12 tool: Identification of signaling domains. *Proc. Natl. Acad. Sci.* 95, 5857–5864.

13 Shin, O.H., Hau, W., Wang, Y., Sudhof, T.C., 2005. Evolutionarily conserved multiple C2 domain
14 proteins with two transmembrane regions (MCTPs) and unusual Ca²⁺ binding properties. *J. Biol.*
15 *Chem.* 280, 1641–1651.

16 Sutton, R.B., Davletov, B.A., Berghuis, A.M., Thomas, C.S., Sprang, S.R., 1995. Structure of the First
17 C2 Domain of Synaptotagmin I : A Novel Ca²⁺ + / Phospholipid-Binding Fold 80, 929–938.

18 Train, C.M., Glover, N.M., Gonnet, G.H., Altenhoff, A.M., Dessimoz, C., 2017. Orthologous Matrix
19 (OMA) algorithm 2.0: More robust to asymmetric evolutionary rates and more scalable
20 hierarchical orthologous group inference. *Bioinformatics* 33, i75–i82.

- 1 Tunstall, N.E., Herr, A., de Bruyne, M., Warr, C.G., 2012. A screen for genes expressed in the
2 olfactory organs of *Drosophila melanogaster* identifies genes involved in olfactory behaviour.
3 PLoS One 7.
- 4 Voleti, R., Tomchick, D.R., Südhof T.C., Rizo, J., 2017. Exceptionally tight membrane-binding may
5 explain the key role of the synaptotagmin-7 C₂A domain in asynchronous neurotransmitter
6 release. Proc Natl Acad Sci U S A. 114,40: E8518-E8527.
- 7 Washington, N.L., Ward, S., 2006. FER-1 regulates Ca²⁺ -mediated membrane fusion during *C.*
8 *elegans* spermatogenesis. J. Cell Sci. 119, 2552–2562.
- 9 Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview Version 2-A
10 multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191.

11

12

1 Figures

Fig. 1.- Phylogenetic analysis of the MCTPs proteins sequences. All sequences selected from the NCBI appear with their accession number. The tree was generated using maximum likelihood method and Bayesian inference using 100 bootstrap replicates, only the model species tree is shown because the patten is the same than the General analysis. Sequence of *C. intestinalis* was used as outgroup. The scale bar represents the length distance units of the tree branches.

Fig. 2.- MCTP C2 domains phylogeny. A Maximum likelihood tree of C2 Domains sequences from different animal models is shown. Bootstrap inference was used making 100 replicates, values along branches are indicated in red. Putative calcium binding and Lipid binding domains were colored in the figure and indicated respectively. The scale bar represents the length of the branches of the tree.

Fig. 3.- Functional Motifs in MCTPs Proteins. A) diagram of MCTP using consensus sequence generated by alignment of model species. Two transmembrane helix at the C- terminus anchored protein to the membrane, followed by three C2 domains. Putative calcium and lipid binding sites are shown in the scheme in C2A to C2C and C2A and C2B respectively. To determine the conservation of calcium sites B-D) and lipid binding pockets F,G) sequences from C2A and C2B from the human synaptotagmin-1 were used as a template respectively. The secondary structure was calculated using PSIPRED software. Columns marked with black asterisks are calcium binding sites in all figures, colored rectangles at the bottom of each alignment highlighted the β -groove motif. E) Alignment from

1 TMR-1 and TMR-2 from representative species. Alignment were made using muscle and Jalview for
2 colored them. Color scheme was zappo style. The conserved regions were calculated using SMART.
3 Columns into gray rectangles represent the TMR regions 1 and 2, respectively.

4

1

2 **Fig 4.- Structural model of C2 domains on MCTP.** Tridimensional representation of C2A (brown),
3 C2B (green), and C2C (blue), domains of MCTP from *C. elegans*. Calcium molecules (grey spheres)
4 interacting with Calcium-binding domain-related residues (red) are indicated. The predicted polybasic
5 cluster-related residues (orange) are also illustrated.

6