Modularity in protein evolution: modular organization and de novo domain evolution in mollusc metallothioneins.

Sara Calatayud, Mario García-Risco, Veronika Pedrini-Martha, Douglas J. Eernisse, Reinhard Dallinger, Óscar Palacios, Mercè Capdevila and Ricard Albalat

1Departament de Genètica, Microbiologia i Estadística and Institut de Recerca de la Biodiversitat (IRBio), Facultat de Biologia, Universitat de Barcelona. Av. Diagonal 643, E-08028, Barcelona, Catalonia, Spain.

2Departament de Química, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Cerdanyola del Vallès, Spain.

3Department of Zoology and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Austria.

4Department of Biological Science, California State University Fullerton, 800 N. State College Blvd, Fullerton, CA, 92831-3599, USA.

*Corresponding author. E-mail: ralbalat@ub.edu

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
ABSTRACT

Metallothioneins (MTs) are proteins devoted to the control of metal homeostasis and detoxification, and therefore, MTs have been crucial for the adaptation of the living beings to variable situations of metal bioavailability. The evolution of MTs is, however, not yet fully understood, and to provide new insights into it we have investigated the MTs in the diverse classes of Molluscs. We have shown that most molluscan MTs are bi-modular proteins that combine six domains –α, β1, β2, β3, γ and δ– in a lineage specific manner. We have functionally characterized the Neritimorpha β3β1 and the Patellogastropoda γβ1 MTs, demonstrating the metal-binding capacity of the new γ domain. Our results have revealed a modular organization of mollusc MT, whose evolution has been impacted by duplication, loss and de novo emergence of domains. MTs represent a paradigmatic example of modular evolution probably driven by the structural and functional requirements of metal binding.
1. INTRODUCTION

Metallothioneins (MTs) are a superfamily of intracellular, cysteine-rich (≈15-30%) and mostly low molecular weight (<100 amino acids) proteins present across eukaryotes, from various protists to plants, fungi and animals (Capdevila and Atrian 2011; Blindauer 2014). Their cysteine (Cys, C) residues are arranged in distinctive motifs (i.e. CxC, CC and CCC), whose number and distribution led to the original definition of the two functional domains in vertebrate MTs (Braun, et al. 1986) designated as α domain (with 11-12 cysteines at the C-terminal region) and β domain (with 9 cysteines at the N-terminal region), joined by a linker sequence. In this bi-modular structure made of two domains, the cysteine motifs of each domain are able to form metallic clusters, thus conferring the capacity of binding both essential and non-essential metals (Nielson and Winge 1985). Notice that the α and β nomenclature has been also used for referring to N-terminal (α) or C-terminal (β) domains in some gastropod MTs (Baumann, et al. 2017; Niederwanger, et al. 2017a; Palacios, et al. 2017; Schmielau, et al. 2018). In order to avoid confusion, however, we will use here the α/β nomenclature in its primary meaning for classifying domains based on the number and distribution of cysteine motifs (Jenny, et al. 2016; Nam and Kim 2017), and not on their N- or C-terminal position.

Thanks to their metal-binding capacity, MTs are directly involved in metal homeostasis and detoxification, but also in radical scavenging, oxidative stress protection and antiapoptotic defense (Capdevila, et al. 2012), leading to their role as a model system for the investigation of the genetic mechanisms by which organisms adapt to diverse metal bioavailabilities or stress. For instance, duplications of MT genes (Maroni, et al. 1987; Adamo, et al. 2012) or expansions of MT domains (Tanguy and Moraga 2001; Jenny, et al. 2016; Pedrini-Martha, et al. 2020), along with elevated levels of MT expression (Timmermans, et al. 2005; Janssens, et al. 2008; Janssens, et al. 2009; Costa, et al. 2012; Catalan, et al. 2016; de Francisco, et al. 2018) or changed metal specificity (Tio, et al. 2004; Palacios, et al. 2011; de Francisco, et al. 2017) have been considered adaptive events contributing to increase the metal and stress tolerance of the organisms in different environments. In particular, a previous thorough analysis investigating the role of cadmium (Cd) on the evolution of gastropod MTs has led us to suggest that lineage-specific changes of metal-selectivity features might have been important during the recurrent colonization of marine gastropods to terrestrial and freshwater habitats, where they had to face the challenge of adapting to ecosystems with different levels of metal bioavailability (Dallinger, et al. 2020). In this context,
studies on MT evolution have been of interest to evolutionary ecologists (Janssens, et al. 2009; Faddeeva-Vakhrusheva, et al. 2016; Zhang, et al. 2018; Purac, et al. 2019; Zhang, et al. 2019) who have associated environmental factors – i.e. concentrations of heavy metals – with the evolution of diverse MTs.

Gastropoda is comprised of five distinctive lineages ranked as subclasses, whose inter-relationships are still controversial: Patellogastropoda, Vetricastropoda, Neritimorpha, Caenogastropoda, and Heterobranchia (Zapata, et al. 2014; Cunha and Giribet 2019). Although several gastropod MTs have been previously investigated, these studies have an uneven phylogenetic distribution. MT identification and metal-binding selectivity have been determined for 15 MTs from seven heterobranch species (with the most extensive studies), for four MTs from three caenogastropod species, and for a single vetigastropod MT. In order to fully understand the evolution of gastropod MTs and their metal-binding specificities, it was therefore necessary to extend studies to include particular species of Patellogastropoda (true limpets) and Neritimorpha (snails known as nerites and their allies). To add both patellogastropods and neritimorphs here, we have conducted an exhaustive survey of gastropod MTs in public databases, including raw sequencing data from transcriptomic and genomic high-throughput sequencing projects. We have been able to reconstruct and identify new MTs from limpet and neritid species. We have characterized the metal-binding properties of new MTs for each, selecting Lottia gigantea and Nerita peloronta MTs as representatives of each taxon. Our data set extends the knowledge of metal-binding preferences to MTs of previously unstudied gastropod taxa, broadening the scope of known diversification of metal-binding selectivity in MTs for gastropods, the most species-rich, and arguably the most morphologically and ecologically diverse, class of Mollusca.

By extending what we have learned with new genomic searches, we have added novel MTs from the other three gastropod clades (Vetricastropoda, Caenogastropoda and Heterobranchia) and from all mollusc classes, identifying and classifying new MTs from other conchiferan (i.e. Gastropoda, Bivalvia, Cephalopoda, Monoplacophora and Scaphopoda) and aculiferan (i.e. Polyplacophora, Solenogastres and Caudofoveata classes) molluscs. In summary, by collecting 272 MTs from 189 different species we have created the most comprehensive catalog of mollusc MTs compiled so far, and we have exposed patterns of modular organization across molluscan MTs.

2. RESULTS
2.1. Identification of Patellogastropoda MTs.

We surveyed the genome project of the *Lottia gigantea* (Patellogastropoda: Lottiidae). This search yielded an automatically predicted MT (Gene ID: 20249168; hypothetical protein: XP_009056965.1), whose sequence, structure and size suspiciously differed from other Gastropoda MT genes. The low-quality sequence of the genomic region containing the putative MT gene (NW_008709190.1) prompted us to re-analyze this region by PCR amplifying, cloning and re-sequencing >12 Kb of *L. gigantea* genome (Figure S1A). Sequence analysis identified two MT genes tandem repeated in this region (accession number MK795721), which we named *LgiMT1* and *LgiMT2*. Our predictions were further supported by PCR amplification of the corresponding cDNAs (MK770430 and MK770431) and by nine *L. gigantea* expressed sequence tags (EST) (Table S1), corroborating the misassembly of the MT genomic region in the genome project, and revealing some degree of polymorphism at the amino acid level (S/T and S/A at amino acid positions 3 and 68, respectively, Figure S1B). Comparison of the coding region (CDS) of *LgiMT1* and *LgiMT2* genes revealed that they were 97.78% identical at the nucleotide level (220 out of 225 nt.), and both encoded for a 75 amino acid protein with 19 cysteines (25.7%). *LgiMT1* and *LgiMT2* proteins only differed at positions 6, 20 and 49: P, L and P residues in *LgiMT1*, and A, S and S in *LgiMT2* (Figure S1B).

We extended the limpet MT listing to three families and eight additional species of Patellogastropoda (Figure S2A, Table S2 and (Dallinger, et al. 2020)). These include five species of Lottiidae: *Lottia digitalis* (*LdiMT1* partial), *Lottia kogamogai* (*LkoMT1*), *Lottia scutum* (*LscMT1* partial), *Nipponacmea fuscoviridis* (*NfuMT1*), and *Patelloida pygmaea* (*PpyMT1*); two species of Nacellidae: *Cellana rota* (*CroMT1*) and *Nacella concinna* (*NcoMT1*); and one species of Patellidae: *Patella vulgata* (*PvuMT1* and *PvuMT2*). Amino acid comparison of the 11 Patellogastropoda MTs revealed high sequence similarity (from 69% to 97% identity) with the distinctive 19 cysteines fully conserved. Interestingly, the 19 cysteines were organized into two domains: a novel MT domain at the N-terminal region with 10 cysteines arranged in five CC pairs (CCx3CCx4CCx6CCx7CC) that we named γ domain (notice that this γ domain is not related with the N-terminal, 6 Cys domain of plant E-1 MTs, also named γ domain (Loebus, et al. 2011)), which was connected by a linker of four to six amino acids to an archetypal 9-Cys β1 domain ([CxC]x2[CxC]x2[CxS][CxC]x3[CxC]) (Jenny, et al. 2016; Nam and Kim 2017) at the C-terminal region (Table 1 and Figure S2A). We concluded therefore that the structure of Patellogastropoda MTs was of γ/β1 domains (Figure 1).
2.2. Metal-binging capacity of Patellogastropoda MTs

In order to demonstrate the metallothionein nature and to explore the metal selectivity features of the two *L. gigantea* MTs, we studied the formation of metal–LgiMT1 and metal–LgiMT2 complexes by the proteins heterologously expressed in *E. coli* and grown in media supplemented with copper (Cu), cadmium (Cd) or zinc (Zn) salts by inductively-coupled plasma atomic emission spectrometer (ICP-AES) and electrospray ionization mass spectrometry (ESI-MS) analyses. ICP-AES is an analytical technique that allows for protein quantification and metal-to-protein stoichiometry determination through the measurement of element composition of the samples (S, Zn, Cd and Cu) (Bongers, et al. 1988), and ESI-MS is used to determine the molecular mass of the species formed, i.e. the speciation of the samples (Capdevila, et al. 2012). The ICP-AES and native ESI-MS (recorded at neutral pH in order to allow the observation of unaltered species) analyses (data not shown) of the recovered samples showed that LgiMT1 and LgiMT2 rendered a mixture of metallated species both in Zn(II)- as well as in Cu(II)-supplemented cultures, thus indicating the absence of preference for either of those metal ions (Zn(II) and Cu(II) stand for divalent Zn$^{2+}$ and Cu$^{2+}$ ions). On the contrary, unique Cd$_7$-LgiMT1 and Cd$_7$-LgiMT2 species, i.e. MTs loaded with 7 Cd(II) ions, were recovered from Cd-enriched culture media (Figure 2A and 2B). The identification of single species after recombinant metal-supplemented productions normally reveals the existence of a preferred thermodynamically favored metal cluster, while in the absence of any specific species, the formation of a variety of species with similar but not identical metal content is observed (Palacios et al., 2014). The observation of single Cd$_7$-LgiMT species, together with the observation by acid ESI-MS (recorded at pH 2.4 that lead to partial protonation of thiol Cys groups, which normally results on the release of Zn and Cd while bound Cu is maintained) that both proteins are reluctant to release Cd(II) by acidification at pH 2.4 giving rise to Cd$_4$-LgiMT species, demonstrated the Cd-thionein character of both Patellogastropoda MTs.

2.3. Metal-binding functionality of the new $\gamma$ domain

In order to determine the independent metal-binding capacity of the new $\gamma$ domain, we analyzed its ability to form metal complexes when expressed alone. Thus, we heterologously expressed the 10-Cys $\gamma$ domain of LgiMT2 (from Met$_1$ to Gln$_{45}$), hereafter as $\gamma$LgiMT2, in *E. coli* grown in medium supplemented with Cu, Cd or Zn salts. The election of the $\gamma$ domain of LgiMT2 relied on the fact that it was more conserved than that of LgiMT1 when compared across the $\gamma$ domains of
other Patellogastropoda MTs (Figure S2A). The distinct metal–γLgiMT2 preparations obtained after purification were characterized by ICP-AES and ESI-MS. The results showed a clear preference of γLgiMT2 for divalent Zn(II) and Cd(II) ions, as major M₄-γLgiMT2 species were identified for both metal ions by ESI-MS at pH 7. Interestingly, formation of Cd₄-γLgiMT2 complexes was more favored than Zn₄-γLgiMT2 complexes, as suggested by the fact that the former one was obtained as a single species (Figure 2C) while the latter one coexisted with other minor species (data not shown). This Cd₄-γLgiMT2 cluster also exhibited a high resistance against demetalation at acidic pH levels, as this required an acidification down to pH 1 to obtain the apo-protein. Conversely to the Zn- and Cd-cultures, the preparations obtained from Cu(II)-enriched media yielded mixtures of multiple Cuₓ-γLgiMT2 complexes (x ranging from 5 to 10), confirming the poor preference of this domain for Cu(I).

2.4. Identification of Neritimorpha MTs

We collected 12 MTs from two families and six species of Neritimorpha, including five species of Neritidae: Clithon retropictum (CretrMT1 and CretrMT2), Nerita albicilla (NalMT1 and NalMT2), Nerita melanotragus (NmeMT1 and NmeMT2), Nerita peloronta (NpeMT1 and NpeMT2), Neritina pulligera (NpuMT1 and NpuMT2); and one species of Neritopsidae: Titiscania limacina (TliMT1 and TliMT2 partial) (Figure S2B and Table S2 and (Dallinger, et al. 2020)). All Neritimorpha species possessed two MTs, and this appears to be a lineage-specific gene duplication shared by at least two of the four extant superfamilies within Neritimorpha. As for patellogastropods, amino acid comparison of the 12 Neritimorpha MTs showed high sequence conservation (from 65% to 96% identity) with 19-20 cysteines organized in two MT domains: at the N-terminal region, a 9-Cys β3 domain ([CxC]₃-[CxC]ₓ₃-[CxC]ₓ₃-[CxC])x₂C, formerly known as α1/2 domain in non-Neritimorpha species (Baumann, et al. 2017; Niederwanger, et al. 2017a; Palacios, et al. 2017; Schmielau, et al. 2018) with an extra cysteine in Neritimorpha MT1s), connected by a linker of three to four amino acids to a 9-Cys β1 domain at the C-terminal region ([CxC]ₓ₃-[CxC]ₓ₃-[CxC]ₓ₃-[CxC]ₓ₃-[CxC]) (Table 1 and Figure S2B). The structure of Neritimorpha MTs was therefore of β3/β1 domains (Figure 1).

2.5. Metal-binding capacity of Neritimorpha MTs

In order to analyze the metal-binding abilities of Neritimorpha MTs, we studied the formation of metal–MT complexes of NpeMT1 and NpeMT2 heterologously expressed in E. coli grown in medium supplemented with Cu, Cd or Zn salts. Metal–NpeMT complexes were purified and analyzed by ICP-AES and
ESI-MS. Our data showed that both Neritimorpha MTs presented similar specificities for divalent metal ions, Zn(II) and Cd(II). However, their binding preference was not exactly the same, since NpeMT1 was more specific for Zn(II) than NpeMT2, while NpeMT2 was more specific for Cd(II) than NpeMT1. Consequently, Zn₂⁻-NpeMT1 and Cd₂⁺-NpeMT2 complexes could be recovered as single species (Figure 3). The Zn-thionein character of NpeMT1 and the Cd-thionein nature of NepMT2 were also supported by the recombinant productions of this protein in Cu-enriched media, which rendered mixtures of heterometallic Zn,Cu-MT complexes for NpeMT1, but mixtures of homometallic Cu-MT species for NepMT2 (data not shown).

2.6. MTs from other Gastropoda clades

We extended the surveys of MTs to gastropod species belonging to Caenogastropoda, Heterobranchia and Vetigastropoda, considered separately below. We collected a total of 163 MT sequences, including newly identified MTs and previously reported ones (Table S2).

2.6.1. Caenogastropoda MTs: We collected 55 MT sequences (43 new) of 43 Caenogastropoda species (Figure S2C and Table S2). MT multiplicity was observed in 11 species (10 species with two, and Pomacea canaliculata with three MTs). Size diversity within Caenogastropoda was high, with MTs ranging from 66 amino acids (BacMT1) to 251 amino acids (McornMT2), with a cysteine content of 17 (25.8%) and 72 cysteines (28.7%), respectively. Overall, Caenogastropoda MTs were organized in an N-terminal 9-Cys β3 domain (CxCxCx₅[CxC]ₓ₃) linked by two to three residues to a C-terminal 9-Cys β1 domain ([CxC]ₓ₃-[CxC][CxC][CxC][CxC][CxC][CxC][CxC]) (Table 1 and Figure S2C). The archetypal structure of Caenogastropoda MTs was therefore of β3/β1 domains (Figure 1). Remarkably, β3 domain duplications were observed in a number of Caenogastropoda MTs: one single duplication (i.e. β3.1/β3.2) in 13 MTs, two duplications (i.e. β3.1/β3.2/β3.3) in JjaMT1, PbrMT2 and PcanMT2 sequences, three duplications (from β3.1 to β3.4) in partial EheMT2 and PcanMT1, five duplications (from β3.1 to β3.6) in AplaMT2, and six duplications (from β3.1 to β3.7) in McornMT2 (Figure 4A and Figure S3A). Since these multi-β3 MTs were unevenly distributed among the Littorinimorpha, Neogastropoda and "architaenioglossan" representatives, independent events of internal domain duplications were the most plausible origin of such MTs (Schmielau, et al. 2018). Also noteworthy, 10 MTs (most of them duplicated copies) of the Architaenioglossa order showed an additional H₃₋₄C₄ motif (HxHHHx₂Cx₅Cx₆:x₆Cₓ₀₋₄) at the N-terminal region (Figure S2C). Since these MTs belonged to the same taxonomic
group, a lineage-specific event that added the H$_{3}C_{4}$ motif to the N-terminus of an ancestral Architaenioglossa MT duplicate would be the most parsimonious explanation.

2.6.2. Heterobranchia: From an MT perspective, this was the most studied Gastropoda clade. We collected 86 MT sequences (61 new) of 55 Heterobranchia species. MT multiplicity was observed in 22 species with two (14 species), three (seven species) or four (one species) MTs (Figure S2D and Table S2). Size diversity of Heterobranchia MTs was the highest within the Gastropoda class, with MTs ranging from 58 (AbaMT1, GscMT1, GcuMT1, GtrMT1 and LstMT1) to 319 (AbiMT4) amino acids, with a cysteine content of 15 (25%) and 85 cysteines (26.2%), respectively. Overall, Heterobranchia MTs were organized in a N-terminal 9-Cys β3 domain (Cx$_{3}$Cx$_{3}$[CxC]$_{3}$[CxC]$_{3}$[CxC]$_{3}$[Cx]) linked by two residues to a C-terminal 9-Cys β1 domain ([CxC]$_{5}$[CxC]$_{3}$[CxC]$_{3}$[CxC]$_{3}$[CxC]) (Table 1 and Figure S2D). The structure of Heterobranchia MTs was therefore of β3/β1 domains (Figure 1). There were, however, several exceptions to this structure. Species of Lymnaeidae, for instance, lacked the last [CxC] motif of the β1 domain, MTs of Bradybaena similis and Fiona pinnata had an additional β3 domain (i.e. β3.1/β3.2) (Figure S2D), and Alinda biplicata AbiMT3 and AbiMT4 had eight and nine β3 domains, respectively (Pedrini-Martha, et al. 2020). But the most deviant Heterobranchia MTs were those of the Biomphalaria species, B. glabrata and B. pfeifferi. These MTs lacked the β1 domain and had three β3 domains (β3.1/β3.2/β3.3) followed by a C-terminal tail with 5 cysteines (Cx$_{5}$Cx$_{5}$Cx$_{3}$CC) (Figure 4B and Figure S3B) (Niederwanger, et al. 2017a). The fact that both species shared the same structure suggested that it could be a general MT feature for Biomphalaria.

2.6.3. Vetigastropoda: We collected 22 MT sequences (18 new) of 18 Vetigastropoda species (Figure S2E and Table S2). MT multiplicity was found in four species: Clanculus pharaonius (CphMT1 and CphMT2 partial), Perotrochus lucaya (PluMT1 and PluMT2), Phasianella ventricosa (PveMT1 and PveMT2), and Prothalotia lehmanni (PleMT1 and PleMT2). Overall, Vetigastropoda MTs were 65-71 amino acid long with 18 cysteines (25.4%-27.7%) organized in a N-terminal 9-Cys β3 domain (Cx$_{3}$Cx$_{4}$[CxC]$_{3}$[CxC]$_{3}$[CxC]$_{3}$[CxC]$_{3}$[CxC]) linked by 3-6 residues to a C-terminal 9-Cys β1 domain ([CxC]$_{3}$[CxC]$_{5}$[CxC]$_{3}$[CxC]$_{3}$[CxC]) (Table 1 and Figure S2E). The last common ancestor for Vetigastropoda thus likely had MTs with β3/β1 domains (Figure 1).

Comparisons across all of the phylogenetically diverse Gastropoda have suggested that the ancestral Gastropoda already had a β3/β1 MT (Figure 1), which
underwent diverse lineage-specific modifications during evolution: (i) gene duplications leading to parallel MT multiplicity in divergent gastropod lineages, (ii) internal domain duplications and losses in the Caenogastropoda and Heterobranchia taxa, and (iii) acquisitions of novel modules such as the H$_3$-r-C$_4$ motif in certain Caenogastropoda, and the new γ domain in Patello gastropoda.

2.7. New MTs in other Mollusc classes

Recent phylogenomic analyses have split Molluscs into two major clades (see Figure 1) (Kocot, et al. 2011; Smith, et al. 2011; Kocot, et al. 2020): Conchifera (including Gastropoda, Bivalvia, Cephalopoda, Monoplacophora and Scaphopoda), and Aculifera (including Polyplacophora, Solenogastres and Caudofoveata). We used publicly available molluscan SRA projects to assemble new MTs that, together with previously reported ones, represented an extensive evolutionary list of molluscan MTs. Thus, our analysis has led us to identify for the first time MTs in Scaphopoda, Cephalopoda and Monoplacophora classes, and the first MTs for chiton, solenogastres and caudofoveatan representatives of Aculifera. Conchifera clade:

2.7.1. Scaphopoda. This molluscan class has been considered the sister group of Gastropods (Smith, et al. 2011), though this phylogenetic relationship is still under debate. We identified two Scaphopoda MTs from Antalis entalis (AenMT1) and Graptaeae eborea (GebMT1) species (Figure S2F and Table S2). Scaphopoda MTs were 68-72 amino acid long with 18-19 cysteines (26.4%) organized in a N-terminal 9-Cys β domain ((CxC)x$_5$[CxC]$_x$[CxC]$_{x2-4}$[CxC]$_x$C) linked by 8-10 residues to a C-terminal 9-Cys β1 domain ((CxC)$_x$[CxC]$_{x3}$[CxC]$_x$[CxC]$_x$C) (Table 1 and Figure S2F). Interestingly, the organization of the [CxC] motifs in the Scaphopoda N-terminal β domain, [CxC]$_x$-C, resembled that of β2 domain defined in the unconventional bivalve MTIIIs (Jenny, et al. 2016), while it differed from that of Gastropoda β1 and β3 domains –(CxC)$_x$-C-(CxC)$_x$ and C-C-(CxC)$^2$C, respectively – (Table 1 and Figure 5). The structure of Scaphopoda MTs was therefore of β2/β1 domains (Figure 1).

2.7.2. Bivalvia. In order to have a broad perspective of the MTs within Bivalvia, we analyzed 62 MTs from 36 species (Figure S2G and Table S2). For clarity, and since the structural diversity of MTs in oysters and mussels has been extensively described elsewhere (Mackay, et al. 1993; Jenny, et al. 2004; Aceto, et al. 2011; Jenny, et al. 2016), we focused our analysis on the MTI and MTIV within available species of Ostreidae, and on MT10, MT10B and MT20 within Mytilidae, which have been considered to have the plesiomorphic condition for bivalve MTs (Nam and Kim 2017). Most Bivalvia MTs were 66-78 amino acid long with 21-25...
cysteines (26.9%-32%) organized in an N-terminal 12-Cys α domain (((CxC)_{x_5}[CxC]_{x_3}[CxC]_{x_4.6}[CxC]_{x_3}[CxC]_{x_3-4}[CxC]_{x_2}C) linked by three residues to a C-terminal 9-Cys β1 domain (((CxC)_{x_3-6}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3-5}[CxC])) (Table 1 and Figure S2G). The likely ancestral structure of these Bivalvia MTs was therefore of α/β1 domains (Figure 1). Other Bivalvia MTs (e.g. Crassostrea MTIIIs in the order Ostreoidea and PmarMT2 of Pinctada martensii in the order Pterioida) were organized in two β2 domains. These MTs probably derived from an ancestral β2/β1 form (Figure 1). Deviations from these structures were found, ranging from loss or gain of some cysteines or small protein fragments (e.g. MquMT1 and CgiMTIV), to significant structural modifications, such as those previously described in a sphaeriid clam (Pisidium coreanum) PcorMT1 and the oyster (Alectryonella plicatula) ApliMT1: α/β1/β1; in Crassostrea MTIIIs: α, or α/β1/β1; or in the scallop (Argopecten irradians) AirMT1 and AirMT2: β2-like/β2-like/β2-like/β1 and β2-like/β2-like/β1, respectively (notice that organization of the [CxC] motifs in the β2-like domain was more similar to β3 than to β2 domain) (Tanguy and Moraga 2001; Baek, et al. 2009; Jenny, et al. 2016; Nam and Kim 2017).

2.7.3. Cephalopoda: We identified the MTs of four Cephalopoda species: Nautilus pompilius (NpoMT1), Octopus vulgaris (OvuMT1), Octopus bimaculoides (ObiMT1), and Sepia esculenta (SesMT1) (Figure S2H and Table S2). Cephalopoda MTs were 69-73 amino acid long with 19-21 cysteines (27.5%-28.7%) organized in α/β1 domains more similar to those of other molluscan classes in the Nautilus MT than in the other species. The N-terminal 12-Cys α domain (((CxC)_{x_5-6}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_2}C) in the Nautilus MT, but slightly divergent in the coleoid cephalopods ((CxC)_{x_5})

ε_{x_6}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_2}C in Sepia and Octopus MTs), was linked by three or four residues to a C-terminal 9-Cys β1 domain (((CxC)_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]) in Nautilus and Sepia MTs, but somewhat modified in Octopus MTs; Table 1 and Figure S2H). Based on Nautilus and Sepia MTs, we concluded that the prototypical structure of Cephalopoda MTs was of α/β1 domains (Figure 1).

2.7.4. Monoplacophora: We identified the first MT in a Monoplacophora species, Laeviplina hyalina (LhyMT1) (Figure S2I and Table S2). LhyMT1 was 63 amino acid long with 19 cysteines (30.2%) organized in an N-terminal β2 domain (((CxC)_{x_6}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_2}C) linked by three residues to a C-terminal 9-Cys β1 domain (((CxC)_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]_{x_3}[CxC]) (Table 1 and Figure S2I). The structure of the Monoplacophora MT was therefore of β2/β1 domains (Figure 1).

Aculifera clade:
2.7.5. Polyplacophora: We identified five MTs in Polyplacophora species: *Acanthochitona crinita* (AcriMT1), *Chaeopleura apiculata* (CapMT1), *Chiton olivaceus* (ColMT1), *Tonicella lineata* (TilMT1), and *Leptochiton rugatus* (LruMT1) (Figure S2J and Table S2). Polyplacophora MTs were 70-73 amino acid long with 21 cysteines (28.8%-30.0%) organized in a N-terminal 12-Cys α domain ([CxC]x₅[CxC]x₃[CxC]x₄[CxC]x₅[CxC]x₃[CxC]x₂[C]) linked by three residues to a C-terminal 9-Cys β1 domain ([CxC]x₅[CxC]x₃[CxC]x₃[CxC]x₃[CxC]) (Table 1 and Figure S2J). The structure of Polyplacophora MTs was therefore of α/β1 domains (Figure 1).

2.7.6. Solenogastres: We identified eight MTs of Solenogastres species: *Alexandromenia crassa* (AcraMT1 partial), *Amphimeniidae sp.* (AspMT1 partial), *Micronium fodiens* (MfoMT1 partial), *Neomenia carinata* (NcaMT1), *Neomenia megatrapezata* (NmegMT1 and NmegMT2) and *Neomeniomorpha sp.* (NspMT1 and NspMT2 partial) (Figure S2K and Table S2). Full-length Solenogastres MTs (i.e. NcaMT1, NmegMT1 and NspMT1) were 65-66 amino acid long with 18-19 cysteines (27.3-29.3%) organized in an N-terminal 9-Cys β2 domain ([CxC]x₅[CxC]x₃[CxC]x₄[CxC]x₂[C]) linked by six residues to a C-terminal 9-Cys β1 domain ([CxC]x₅[CxC]x₃[CxC]x₃[CxC]x₃[CxC]) with an additional cysteine in NmegMT1 and NspMT1 sequences (Table 1 and Figure S2K). The structure of these MTs was therefore of β2/β1 domains (Figure 1). Interestingly, *N. megatrapezata* species had a second MT, NmegMT2, whose N-terminal domain was a 12-Cys α domain ([CxC]x₅[CxC]x₃[CxC]x₃[CxC]x₃[CxC]x₂[C]) linked by seven residues to a C-terminal 9-Cys β1 domain ([CxC]x₅[CxC]x₃[CxC]x₃[CxC]x₃[CxC]) (Table 1 and Figure S2L). The NmegMT2 structure was, therefore, of the α/β1 type, meaning that two structurally different MTs coexisted in the same Solenogastres species (Figure 1).

2.7.7. Caudofoveata: We identified four MTs in Caudofoveata species: *Chaetodermia nitidulum* (CniMT1), *Falcidens caudatus* (FcaMT1), *Falcidens sagittiferus* (FsaMT1), and *Scutopus ventrolineatus* (SveMT1) (Figure S2M and Table S2). The Caudofoveata MTs were 73-78 amino acid long with 23 cysteines (29.5%-31.5%) organized in a novel N-terminal 14-Cys domain ([CxC]x₅[CxC]CCx₇[CxC]x₅[CxC]x₃[CxC]C) that we named δ domain, linked by three residues to a C-terminal 9-Cys β1 domain ([CxC]x₅[CxC]x₃[CxC]x₃[CxC]x₃[CxC]) (Table 1 and Figure S2M). The structure of Caudofoveata MTs was therefore of δ/β1 domains (Figure 1).

In summary, we have identified more than 270 MTs in 189 different species distributed across the eight molluscan classes, ≈64% as single copy sequences
Table S2; the existence of additional MTs in some species cannot be excluded because some sequence databases are still in progress). Our data showed that a single MT with a two-domain (i.e. bi-modular) structure was the predominant type of MTs in most Mollusca species (Figure 1 and Table 1). Our results revealed, however, many exceptions to this situation, with MT multiplicity and/or multi-modular MTs in many Mollusca species, which denoted an intricate and dynamic evolutionary history of mollusc MTs.

3. DISCUSSION

3.1. Functional evolution of mollusc MTs

3.1.1. Evolution of the metal binding capacity

Most mollusc MTs are bi-modular proteins with 18 (in β2-3/β1 MTs), 19 (in γ/β1), 21 (in α/β1) or 23 (in δ/β1) cysteines. Neritimorpha β3/β1 MTs render homometallic complexes with six divalent ions, either Zn(II) or Cd(II) (Figure 3), meaning that each 9-Cys domain is designed to allocate three divalent metal ions. This capacity is similar to those reported for other bi-modular gastropod MTs (Perez-Rafael, et al. 2012; Palacios, et al. 2014; Perez-Rafael, et al. 2014; Dvorak, et al. 2018), and agrees with three-dimensional (3D) structural analysis of Littorina littorea (Baumann, et al. 2017) and Helix pomatia MTs (Beil, et al. 2019). The metal binding capacity of Patellogastropoda γ/β1 LgiMTs, both with 19 cysteines, is slightly higher since they bind seven divalent metal ions (Figure 2A and B), similar to the 21-Cys α/β1 MTs of Bivalvia (Munoz, et al. 2002; Orihuela, et al. 2008). The successful synthesis of the γ domain of LgiMT2, with 10 Cys residues, and the characterization of the species produced in Cd(II)-enriched media indicate that the LgiMTs render metal-aggregates containing seven Cd(II) ions because these MTs are capable of binding three Cd(II) ions in their β1 domain while their γ domain allocates four Cd(II) ions (Figure 2C). The metal binding capacity of MTs appears, therefore, to rely on the number and position of the cysteines in the different domains and, as expected, the higher the cysteine content, the higher the metal-binding capacity.

In that sense, an effective evolutionary strategy for increasing the metal binding capacity of MTs has been the design of multi-modular forms with high cysteine content and a high capacity of metal binding (Niederwanger, et al. 2017b; Palacios, et al. 2017; Calatayud, et al. 2018). Multi-modular MTs had been identified in a few Bivalve and Gastropoda species (Tanguy and Moraga 2001; Jenny, et al. 2004; Jenny, et al. 2016; Baumann, et al. 2017; Nam and Kim 2017; Niederwanger, et al. 2017a; Palacios, et al. 2017; Schmielau, et al. 2018; Pedrini-
Martha, et al. 2020), and our results have increased this list with 16 new proteins containing a variable number of repeated domains. Sequence comparisons (Figure 4 and Figure S3) and structural analysis (Baumann, et al. 2017) indicate that these multi-modular MTs originated by N-terminal duplications of the β3 domain, in agreement with the idea that proteins tend to increase in length mainly by the gain of sequences at the 5’-end of their genes (Toll-Riera and Alba 2013). The evolution of such multi-modular MTs in some mollusc species suggest that diverse lineages have had to adapt to different conditions of metal bioavailability and stress, despite the ecophysiological determinants that have favored them in only some species remain, however, unknown.

3.1.2. Evolution of the metal binding preference

The evolution of MTs with different metal preferences has usually been associated to scenarios of MT multiplicity, in which neofunctionalization processes yielded MT duplicates with new metal binding selectivities. Sixty-seven species patchily distributed across the mollusc phylogeny possess at least two MTs (Table S2), and their sequence (Figure S2) and domain conservation (Figure 1 and Table 1) suggest that most of them originated after the splitting of the main Mollusc groups by lineage-specific duplications. For instance, a gene duplication in the ancestor of the Neritimorpha class resulted in two MTs that according to the analysis of N. peloronta NpeMT1 and NpeMT2 diverged in their metal preferences (i.e. a Zn-thionein character for NpeMT1 and a Cd-thionein nature for NpeMT2; Figure 3). In contrast, the duplication found in L. gigantea appears more recent since both duplicates are 96% identical (Figure S1) and still share a Cd binding preference (Figure 2). Our results support the idea that metal-binding preference and hence, functional specificity, does not mainly rely on the number and position of the cysteines (>94% identical between NpeMT1 and NpeMT2) but on the nature of the non-coordinating amino acids (=70% different between NpeMT1 and MpeMT2) distributed along the protein sequence (Palacios, et al. 2011; Perez-Rafael, et al. 2014; Dallinger, et al. 2020), although we are still far from being able to predict metal preference based on the analysis of the non-coordinating amino acids.

We do not know the biological determinants that favored the evolution of MTs with different metal preferences in certain mollusc lineages, but the colonization of new habitats (Dallinger, et al. 2020) and the emergence of physiological novelties (Dallinger, et al. 2005; Höckner, et al. 2011) have been proposed as significant evolutionary factors. We do not know either the binding selectivity of the ancestral mollusc MT, but the widespread cadmium-binding
capacity of many MTs—not only in mollusc but in diverse marine animals (Narula, et al. 1995; Riek, et al. 1999; Valls, et al. 2001; Guirola, et al. 2012; Calatayud, et al. 2018)—along with the ancient origin α and β domains (see below), lead us to speculate that ancestral MTs might have been a detoxification system that was later co-opted for homeostatic functions for essential metals. Other scenarios are possible but since Cd is a highly toxic metal because it competes for Zn-dependent cellular processes, and Cd is frequently found with Zn in ore deposits of the earth crust, an early evolution of Cd-detoxifying MTs could have conferred a significant physiological advantage to marine organisms. This advantage would be especially important after increased Cd levels during Paleozoic era (Dallinger, et al. 2020), concomitantly with the Cambrian explosion and the emergence of most animal phyla. From the ancestral MT, different metal-selective MTs would have independently evolved in different molluscan lineages: Cu-selective MTs in Heterobranchia species (Höckner, et al. 2011; Perez-Rafael, et al. 2011; Palacios, et al. 2014); Zn-selective forms in Neritimorpha gastropods (this work) and bivalves (Orihuela, et al. 2008); and metal-unselective MTs in Vetigastropoda and Heterobranchia lineages (Höckner, et al. 2011; Perez-Rafael, et al. 2011; Perez-Rafael, et al. 2012; Perez-Rafael, et al. 2014; Niederwanger, et al. 2017a). This functional diversification might be related to the extraordinary evolutionary success of the phylum, with species that have colonized and adapted to very diverse habitats around the world.

3.2. Structural evolution of Mollusca MTs.

3.2.1. Bi-modular structure of Mollusca MTs.

Vertebrate MTs have a bi-modular structure made of two independent functional domains—an 11-12-Cys α domain and a 9-Cys β domain—each one capable to bind metal ions (Braun, et al. 1986; Capdevila, et al. 1997; Cols, et al. 1999). By comparison with vertebrates, the bi-modular structure has been extended to gastropod (Palacios, et al. 2011; Perez-Rafael, et al. 2012; Perez-Rafael, et al. 2014; Dvorak, et al. 2018; Beil, et al. 2019) and bivalve MTs (Jenny, et al. 2004; Jenny, et al. 2016; Nam and Kim 2017; Yingprasertchait, et al. 2019). Our results spread the bi-modular structure to the MTs of all molluscan classes since 90% of their MTs are two-domain proteins. In these bi-modular MTs, an N-terminal domain that it is variable depending on the taxon (an α domain in Bivalvia, Cephalopoda and Polyplacophora MTs; a β2 domain in Scaphopoda, Monoplacophora and Solenogastres MTs; a δ domain in Caudofoveata MTs; and a β3 domain in Gastropoda MTs, with the exception of the γ domain in Patellogastropoda MTs) is linked to a conserved β1 domain at the C-terminal
region (Figure 1 and Table 1). The pervasiveness of the C-terminal β1 domain suggests a conserved role for this domain, probably related to the stabilization of the three-dimensional cluster structure of the entire protein (Dallinger, et al. 2020).

3.2.2. Modular evolution of MTs: ancient and recent domains

The origin and the evolutionary relationships of the distinct α, β1, β2, β3, γ and δ domains were intriguing. The finding of α, β1 and β2 domains in diverse molluscan classes together with the presence of α/β1 and β2/β1 MTs in species of both Conchifera and Aculifera (Figure 1) supports that the origin of the domains predated the diversification of the phylum, which has estimated at more than 545 millions years ago (MYA) (Kocot, et al. 2020), and suggests that an ancient MT multiplicity was subsequently maintained or lost in a lineage-specific manner. In addition, a possible connection between α and β domains might be envisaged based on the number and configuration of the cysteine motifs (Figure 5). Thus, the three β domains might have derived from an ancestral α domain trimmed at the C-terminal end (β1), at the middle (β2), and at the N-terminal end (β3), respectively. This possibility challenges the classical view that β domains represent ancestral forms, and that α domains evolved later to provide detoxification capacity in front of toxic metals such as cadmium (Cols, et al. 1999).

In contrast to the ancient origin of α, β1 and β2 domains, the restricted distribution of β3, γ and δ domains in some mollusc lineages suggests a more recent origin, probably concomitant with the appearance of these taxonomic groups (Figure 1). Novel N-terminal domains would have replaced former and older ones specifically in some mollusc lineages: the β3 domain in Gastropoda, the γ domain in Patello gastropoda, and the δ domain in Caudofoveata. Other scenarios of domain evolution cannot be ruled out, but they would require assuming complex processes of parallel or convergent evolution to justify the current distribution of the different domains throughout Mollusca.

3.2.3. De novo domain evolution

Whereas there might be an evolutionary relationship between α and β domains, any link among these domains and the novel γ and δ domains is obscured by their differences. Although it cannot be ruled out that γ and δ domains derived from α or β domains that have diverged too much for homology to be recognized, the sequence and cysteine motifs in the γ and δ domains are so divergent from those in α and β domains that they seem to have evolved de novo. Recent analyses have shown that, indeed, de novo evolution is more frequent than previously thought (Neme and Tautz 2013; Toll-Riera and Alba 2013; Weisman and Eddy 2017; Levy 2019). Studies about the emergence of novel domains in
human proteins, for instance, have revealed more than 400 'young' domains, 164 of which are found combined with older ones and preferentially located at the N-terminus of the proteins (Toll-Riera and Alba 2013). These new domains are rich in low-complexity sequences (Toll-Riera, et al. 2012) and tend to be structurally disordered (Moore and Bornberg-Bauer 2012). Such structural features match well those of MTs, which are considered as low complexity and intrinsically disordered proteins. Thus, de novo emergence of MT domains might be relatively easy under an evolutionary perspective because the only requirement for a peptide to function as a metal ion chelator would be a high content of coordinating residues (e.g. cysteines) and a relative small length that favored the polypeptide folding (Capdevila and Atrian 2011). De novo evolution of MTs has been, indeed, implicitly stated from diverse evolutionary studies concluding that MTs likely evolved more than once in different animal phyla (Capdevila and Atrian 2011; Blindauer 2014; Isani and Carpene 2014; Ziller and Fraissinet-Tachet 2018).

In summary, the evolution of the mollusc MTs is intriguing. At the short term (at low taxonomic ranks), it appears to have followed the habitual evolutionary patterns based on progressive accumulation of changes in the sequence, and on duplications or losses of internal domains or genes. At the long term (at high taxonomic ranks), in contrast, MT evolution seems to have been mainly impacted by emergence of new structural domains. The modular structure of mollusc MTs makes the analyses of their domain organization and cysteine motifs more informative for inferring their evolution during the diversification of the phylum than the classic comparisons of sequences, which may be biased towards a general cysteine-richness due to the structural and functional requirements of metal binding.

4. MATERIALS & METHODS
4.1. Database searches and MT identification

Molluscan MT sequences were identified from public databases by Entrez searches using ‘metallothionein’ and ‘Mollusca’ as queries. Retrieved MT sequences were then used as queries in tblastn searches in EST and genomic NCBI and eSnail (http://soft.bioinfo-minzhao.org/esnail/index.html) databases. In addition, RNA Sequence Read Archives (RNA-SRA) for each mollusc species deposited in NCBI were blast searched using as queries MT sequences from the nearest phylogenetically species as well as from different mollusc species covering all the major clades. Raw sequence data was retrieved from the SRA and assembled using SeqMan 8.0.2 (Pro Assembler) software from the DNASTAR
Lasergene package, and manually inspected in order to reconstruct new MT sequences. The MT nature of each new identified sequence was evaluated by blastx searches against metazoan NCBI non-redundant protein sequence database. The amino acid sequences and the accession numbers of the retrieved MTs are provided in Table S2.

4.2. Characterization of *L. gigantea* MT genes.

Collection of three *L. gigantea* specimens, dissection of their hepatopancreas, RNA extraction and storage in RNA later (Thermo Fisher Scientific, Waltham, CA, USA) were performed by one of us (DJE). Genomic DNA was obtained from hepatopancreatic tissue disrupted with the TissueLyser II (Qiagen, Hilden, Germany) procedure, and following the manufacturer’s instructions of DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) for DNA extraction. The DNA concentration, purity and integrity was checked using Tecan Infinite M200 (Tecan Group Ltd, Switzerland) measuring the absorbance at 260 and 280 nm. For the RNA extraction, tissue of the midgut gland was homogenized with glass beads using the Precellys homogenizer (Bertin Instruments, Montigny-le-Bretonneux, France). Total RNA was isolated applying the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) including on-column DNase I digestion (Qiagen, Hilden, Germany) according to manufacturer’s constructions. RNA integrity was checked by visualization on a 1.5 % agarose gel (Biozym, Hessisch Oldendorf, Germany). For cDNA synthesis 450 ng total RNA was used applying the RevertAid Reverse Transcriptase (Fermentas by Thermo Fisher Scientific, Waltham, MA, USA).

The genomic region containing the putative *MT* genes of *L. gigantea* was PCR amplified, cloned PCR primers (Table S3) were designed based on the *L. gigantea* genome project (Scaffold and re-sequenced. 35, NW_008709190.1) in order to amplify overlapping fragments of different sizes (from 500 bp to 4 kb), covering the entire putative MT-containing genomic region. For each PCR reaction, 1 ng of genomic DNA were amplified using selected pairs of primers and the Phusion High-Fidelity DNA Polymerase (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) in a final 25 µl reaction. PCR conditions were 98°C 30 seconds (s); 35 cycles of 98°C 10 s, 58°C 30 s and 72°C 3 minutes (min); and 72°C 10 min. PCR products were visualized in 0.7% agarose gels, isolated with the GelElute Plasmid Miniprep Kit (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany), and cloned with TOPO® TA Cloning® Kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Plasmid DNA was purified from bacteria using the GeneElutet Plasmid Miniprep Kit (Sigma-Aldrich), screened for insert presence by
digestion with EcoRI (EcoRI Fast Digest Restriction Enzyme, Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), and sequenced at the Scientific and Technological Centers of the University of Barcelona using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) in an automatic sequencer (ABI PRISM 310, Applied Biosystems). Several internal primers (Table S3) were used to fully sequence the overlapping PCR fragments, which were manually assembled in order to reconstruct the entire genomic region.

The prediction of the L. gigantea MT genes was corroborated by the PCR amplification of the corresponding cDNAs using gene specific primers (Table S3) designed with CLC main workbench (Version 6.9), and Advantage 2 polymerase (Clontech, Takara Bio Europe, Saint-Germain-en-Laye, France) in a final 50 µl reaction. Cycling conditions were 95°C 1 min; 30 cycles 95°C 30 s, 53°C (for LgiMT2) / 55.5°C (for LgiMT1) 30 s, 68°C 40 s; and 68°C 5 min. PCR products were visualized on a 1.5 % agarose gels, purified using the QIAquick gel extraction kit (Qiagen, Hilden, Germany) and cloned with the TOPO® TA Cloning® Kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Insert containing plasmids were isolated using the QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany) and sent for sequencing to Microsynth (Balgach, Switzerland). Sequences were analyzed via CLC main workbench (Version 6.9).

4.3. Production and purification of recombinant of metal-MT complexes.

Synthetic cDNAs codifying the selected MTs (i.e. LgiMT1, LgiMT2, LgiMT2-γ domain, NpeMT1 and NpeMT2) were provided by Synbiotech (Monmouth Junction NJ, USA) cloned in the pGEX-4T-1 expression vector (GE Healthcare). Recombinant plasmids were transformed in E. coli BL21 strain, a protease deficient strain used for heterologous protein expression. For heterologous protein production, 500 mL of LB medium with 100 µg/mL ampicillin were inoculated with E. coli BL21 cells transformed with the corresponding recombinant plasmid. After overnight growth at 37°C/250 rpm, the cultures were used to inoculate 5 L of fresh LB-100 µg/mL ampicillin medium. Gene expression was induced with 100 µM isopropyl-β-D-thiogalactopyranoside (IPTG) for 3 hours (h). After the first 30 min of induction, cultures were supplemented with ZnCl₂ (300 µM), CdCl₂ (300 µM) or CuSO₄ (500 µM) in order to generate metal-MT complexes. Cells were harvested by centrifugation for 5 min at 9100 g (7700 rpm), and bacterial pellets were suspended in 125 mL of ice-cold PBS (1.4 M NaCl, 27 mM KCl, 101 mM Na₂HPO₄, 18 mM KH₂PO₄ and 0.5% v/v β-mercaptoethanol). Resuspended cells were sonicated (Sonifier Ultrasonic Cell Disruptor) 8 min at voltage 6 with pulses of 0.6 sec, and then centrifuged for 40 min at 17200 g (12000 rpm) and 4°C. Soluble
protein extracts containing GST-MT fusion proteins were incubated with glutathione sepharose beads (GE Healthcare) for 1 h at room temperature with gentle rotation. GST-MT fusion proteins bound to the sepharose beads were washed with 30 mL of cold 1xPBS bubbled with argon to prevent oxidation. After three washes, GST-MT fusion proteins were digested with thrombin (GE Healthcare, 25 U/L of culture or SERVA, 25U/L of culture) overnight at 17°C, thus enabling separation of the metal-MT complexes from the GST that remained bound to the sepharose matrix. The eluted metal-MT complexes were concentrated with a 3 kDa Centriprep Low Concentrator (Amicon, Merck), and fractionated on a Superdex-75 FPLC column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 7.0. The protein-containing fractions, identified by their absorbance at 254 nm, were pooled and stored at -80°C until use.

4.4. Analysis of metal-MT complexes.

Protein quantification and element composition of all the samples was achieved by Inductively-Coupled Plasma Atomic Emission Spectrometer (ICP-AES) measurements performed in a Optima 4300DV (Perkin-Elmer, MA, USA) apparatus (S, 182.040 nm; Zn, 213.856 nm; Cd, 228.802 nm; Cu, 324.803 nm) under conventional conditions following an already established method (Bongers, et al. 1988).

Molecular weights were determined by ESI-MS, in a MicroTof-Q instrument (Bruker Daltonics GmbH, Bremen, Germany) connected to a Series 1100 HPLC pump (Agilent Technologies) controlled by the Compass Software. The instrument was calibrated with ESI-L Low Concentration Turning Mix (Agilent Technologies, Santa Clara, CA, USA). Metallated forms were detected under native conditions: 20 µL of sample injected through a PEEK tube at 30-50 µL·min⁻¹ in a 3.5-5.0 kV capillary-counter voltage, at 90-110°C of desolvation temperature, and with dry gas at 6 L·min⁻¹. Spectra were recorded between a m/z range from 800 to 3000. The liquid carrier was a 90:10 mixture of 15 mM ammonium acetate and acetonitrile at pH 7.0. All molecular masses were calculated according to the bibliography (Fabris, et al. 1996).

5. DATA AVAILABILITY

The data underlying this article are available in NCBI (https://www.ncbi.nlm.nih.gov/) and eSnail (http://soft.bioinformaticsminzhao.org/esnail/index.html) databases. The amino acid sequences and the accession numbers of the mollusc MTs are available in Table S2 in online supplementary material.
6. ACKNOWLEDGEMENTS

Researchers from Barcelona want to acknowledge the Spanish Ministerio de Ciencia e Innovación and FEDER for the projects BIO2015-67358-C2-2-P (M.C. and O.P.) and BIO2015-67358-C2-1-P (R.A.). M.C. and O.P. are members of the “Grup de Recerca de la Generalitat de Catalunya”, ref. 2017SGR-864, and R.A. of ref. 2017SGR-1665. M.G.-R. acknowledges to the UAB the PIF grant. D.J.E. acknowledges support from the United States National Science Foundation grant DEB-1355230. Parts of this study were co-financed by a cooperation grant to R.D. from the Austrian Science Fund and the Swiss National Science Foundation (DACH grant No I 1482-N28). We also thank the Servei d’Anàlisi Química (SAQ) at the Universitat Autònoma de Barcelona (ICP-AES, ESI-MS) for allocating instrument time, the Centres Científics i Tecnològics (CCIT) at the Universitat de Barcelona for DNA sequencing, and to Sebastian Artme for the experimental support.

7. REFERENCES

Aceto S, Formisano G, Carella F, De Vico G, Gaudio L 2011. The metallothionein genes of Mytilus galloprovincialis: genomic organization, tissue expression and evolution. Mar Genomics 4: 61-68. doi: 10.1016/j.margen.2011.01.001

Adamo GM, Lotti M, Tamas MJ, Brocca S 2012. Amplification of the CUP1 gene is associated with evolution of copper tolerance in Saccharomyces cerevisiae. Microbiology 158: 2325-2335. doi: 10.1099/mic.0.058024-0

Baek M-K, Lee J-S, Kang S-W, Lee J-B, Kang H-J, Jo YH, Noh M-Y, Han Ys, Choi S-H, Chae S-H, et al. 2009. Phylogenetic Analysis based on Metallothionein Gene Sequence of an Indigenous Species Pisidium (Neopisidium) coreanum in Korea. The Korean Journal of Malacology 25.

Baumann C, Beil A, Jurt S, Niederwanger M, Palacios O, Capdevila M, Atrian S, Dallinger R, Zerbe O 2017. Structural Adaptation of a Protein to Increased Metal Stress: NMR Structure of a Marine Snail Metallothionein with an Additional Domain. Angew Chem Int Ed Engl 56: 4617-4622. doi: 10.1002/anie.201611873

Beil A, Jurt S, Walser R, Schonnut T, Guntert P, Palacios O, Atrian S, Capdevila M, Dallinger R, Zerbe O 2019. The Solution Structure and Dynamics of Cd-Metallothionein from Helix pomatia Reveal Optimization for Binding Cd over Zn. Biochemistry 58: 4570-4581. doi: 10.1021/acs.biochem.9b00830

Blindauer CA. 2014. Metallothioneins. In. Binding, Transport and Storage of Metal Ions in Biological Cells: The Royal Society of Chemistry. p. 606-665.

Bongers J, Walton CD, Richardson DE, Bell JU 1988. Micromolar protein concentrations and metalloprotein stoichiometries obtained by inductively coupled plasma atomic emission spectrometric determination of sulfur. Anal Chem 60: 2683-2686.

Braun W, Wagner G, Worgotter E, Vasak M, Kagi JH, Wuthrich K 1986. Polypeptide fold in the two metal clusters of metallothionein-2 by nuclear magnetic resonance in solution. J Mol Biol 187: 125-129. doi: 10.1016/0022-2836(86)90412-2
Calatayud S, Garcia-Risco M, Rojas NS, Espinosa-Sanchez L, Artime S, Palacios O, Cañestro C, Albalat R 2018. Metallothioneins of the urochordate Oikopleura dioica have Cys-rich tandem repeats, large size and cadmium-binding preference. Metallomics 10: 1585-1594. doi: 10.1039/c8mt00177d

Capdevila M, Atrian S 2011. Metallothionein protein evolution: a miniassay. J Biol Inorg Chem 16: 977-989. doi: 10.1007/s00775-011-0798-3

Capdevila M, Bofill R, Palacios Ò, Atrian S 2012. State-of-the-art of metallothioneins at the beginning of the 21st century. Coordination Chemistry Reviews 256: 46-62. doi: http://dx.doi.org/10.1016/j.ccr.2011.07.006

Capdevila M, Cols N, Romero-Isart N, Gonzalez-Duarte R, Atrian S, Gonzalez-Duarte P 1997. Recombinant synthesis of mouse Zn3-beta and Zn4-alpha metallothionein 1 domains and characterization of their cadmium(II) binding capacity. Cell Mol Life Sci 53: 681-688. doi: 10.1007/s000180050088

Catalan A, Glaser-Schmitt A, Argyridou E, Duchen P, Parsch J 2016. An Indel Polymorphism in the MtnA 3' Untranslated Region Is Associated with Gene Expression Variation and Local Adaptation in Drosophila melanogaster. PLoS Genet 12: e1005987. doi: 10.1371/journal.pgen.1005987

Cols N, Romero-Isart N, Bofill R, Capdevila M, Gonzalez-Duarte P, Gonzalez-Duarte R, Atrian S 1999. In vivo copper- and cadmium-binding ability of mammalian metallothionein beta domain. Protein Eng 12: 265-269. doi: 10.1093/protein/12.3.265

Costa D, Marien J, Janssens TK, van Gestel CA, Driessen G, Sousa JP, van Straalen NM, Roelofs D 2012. Influence of adaptive evolution of cadmium tolerance on neutral and functional genetic variation in Orchesella cincta. Ecotoxicology 21: 2078-2087. doi: 10.1007/s10646-012-0961-9

Cunha TJ, Giribet G 2019. A congruent topology for deep gastropod relationships. Proc Biol Sci 286: 20182776. doi: 10.1098/rspb.2018.2776

Dallinger R, Chabicovsky M, Hödl E, Prem C, Hunziker P, Manzl C 2005. Copper in Helix pomatia (Gastropoda) is regulated by one single cell type: differently responsive metal pools in rhogocytes. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 289: R1185-R1195. doi: 10.1152/ajpregu.00052.2005

Dallinger R, Zerbe O, Baumann C, Egger B, Capdevila M, Palacios O, Albalat R, Calatayud S, Ladurner P, Schlick-Steinbr, et al. 2020. Metallomics reveal a persisting impact of cadmium on the evolution of metal-selective snail metallothioneins. Metallomics 12: 702-720. doi: 10.1039/c9mt00259f

de Francisco P, Martin-Gonzalez A, Turkewitz AP, Gutierrez JC 2017. Extreme metal adapted, knockout and knockdown strains reveal a coordinated gene expression among different Tetrahymena thermophila metallothionein isoforms. PLoS ONE 12: e0189076. doi: 10.1371/journal.pone.0189076

de Francisco P, Martin-Gonzalez A, Turkewitz AP, Gutierrez JC 2018. Genome plasticity in response to stress in Tetrahymena thermophila: selective and reversible chromosome amplification and paralogous expansion of metallothionein genes. Environ Microbiol 20: 2410-2421. doi: 10.1111/1462-2920.14251

Dvorak M, Lackner R, Niederwanger M, Rotondo C, Schnegg R, Ladurner P, Pedrini-Martha V, Salvenmoser W, Kremser L, Lindner H, et al. 2018. Metal binding functions of metallothioneins in the slug Arion vulgaris differ from metal-specific isoforms of terrestrial snails. Metallomics. doi: 10.1039/c8mt00215k

Fabris D, Zaia J, Hathout Y, Fenselau C 1996. Retention of Thiol Protons in Two Classes of Protein Zinc Ion Coordination Centers. Journal of the American Chemical Society 118: 12242-12243. doi: 10.1021/ja962863r

Faddeeva-Vakhрушева A, Derks MF, Anvar SY, Agamennone V, Suring W, Smit S, van Straalen NM, Roelofs D 2016. Gene Family Evolution Reflects
Adaptation to Soil Environmental Stressors in the Genome of the Collembolan Orchesella cincta. Genome Biol Evol 8: 2106-2117. doi: 10.1093/gbe/evw134

Guirola M, Perez-Rafael S, Capdevila M, Palacios O, Atrian S 2012. Metal dealing at the origin of the Chordata phylum: the metallothionein system and metal overload response in amphioxus. PLoS ONE 7: e43299. doi: 10.1371/journal.pone.0043299

Höckner M, Stefanon K, de Vaufleury A, Monteiro F, Pérez-Rafael S, Palacios O, Capdevila M, Atrian S, Dallinger R 2011. Physiological relevance and contribution to metal balance of specific and non-specific Metallothionein isoforms in the garden snail, Cantareus aspersus. Biometals 24: 1079-1092. doi: 10.1007/s10534-011-9466-x

Isani G, Carpene E 2014. Metallothioneins, unconventional proteins from unconventional animals: a long journey from nematodes to mammals. Biomolecules 4: 435-457. doi: 10.3390/biom4020435

Janssens TK, Lopez Rdel R, Marien J, Timmermans MJ, Montagne-Wajer K, van Straalen NM, Roelofs D 2008. Comparative population analysis of metallothionein promoter alleles suggests stress-induced microevolution in the field. Environ Sci Technol 42: 3873-3878. doi: 10.1021/es702618s

Janssens TKS, Roelofs D, Van Straalen NM 2009. Molecular mechanisms of heavy metal tolerance and evolution in invertebrates. Insect Science 16: 3-18. doi: 10.1111/j.1744-7917.2009.00249.x

Jenny MJ, Payton SL, Baltzegar DA, Lozier JD 2016. Phylogenetic Analysis of Molluscan Metallothioneins: Evolutionary Insight from Crassostrea virginica. J Mol Evol 83: 110-125. doi: 10.1007/s00239-016-9758-4

Jenny MJ, Ringwood AH, Schey K, Warr GW, Chapman RW 2004. Diversity of metallothioneins in the American oyster, Crassostrea virginica, revealed by transcriptomic and proteomic approaches. Eur J Biochem 271: 1702-1712. doi: 10.1111/j.1432-1033.2004.04071.x

Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, Santos SR, Schander C, Moroz LL, Lieb B, et al. 2011. Phylogenomics reveals deep molluscan relationships. Nature 477: 452-456. doi: 10.1038/nature10382

Kocot KM, Poustka AJ, Stoger I, Halanych KM, Schrodli M 2020. New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. Sci Rep 10: 101. doi: 10.1038/s41598-019-56728-w

Levy A 2019. How evolution builds genes from scratch. Nature 574: 314-316. doi: 10.1038/d41586-019-03061-x

Loebus J, Peroza EA, Bluthgen N, Fox T, Meyer-Klaucke W, Zerbe O, Freisinger E 2011. Protein and metal cluster structure of the wheat metallothionein domain gamma-E(c)-1: the second part of the puzzle. J Biol Inorg Chem 16: 683-694. doi: 10.1007/s00775-011-0770-2

Mackay EA, Overnell J, Dunbar B, Davidson I, Hunziker PE, Kagi JH, Forthergill JE 1993. Complete amino acid sequences of five dimeric and four monomeric forms of metallothionein from the edible mussel Mytilus edulis. Eur J Biochem 218: 183-194. doi: 10.1111/j.1432-1033.1993.tb18364.x

Maroni G, Wise J, Young JE, Otto E 1987. Metallothionein gene duplications and metal tolerance in natural populations of Drosophila melanogaster. Genetics 117: 739-744.

Moore AD, Bornberg-Bauer E 2012. The dynamics and evolutionary potential of domain loss and emergence. Mol Biol Evol 29: 787-796. doi: 10.1093/molbev/msr250

Munoz A, Forsterling FH, Shaw CF, 3rd, Petering DH 2002. Structure of the (113)Cd(3)beta domains from Homarus americanus metallothionein-1: hydrogen bonding and solvent accessibility of sulfur atoms. J Biol Inorg Chem 7: 713-724. doi: 10.1007/s00775-002-0345-3
Nam YK, Kim EJ 2017. Diversification and domain evolution of molluskan metallothioneins: a mini review. Fisheries and Aquatic Sciences 20: 8. doi: 10.1186/s41240-017-0054-z

Narula SS, Brouwer M, Hua Y, Armitage IM 1995. Three-dimensional solution structure of Callinectes sapidus metallothionein-1 determined by homonuclear and heteronuclear magnetic resonance spectroscopy. Biochemistry 34: 620-631. doi: 10.1021/bi00002a029

Neme R, Tautz D 2013. Phylogenetic patterns of emergence of new genes support a model of frequent de novo evolution. BMC Genomics 14: 117. doi: 10.1186/1471-2164-14-117

Niederwanger M, Calatayud S, Zerbe O, Atrian S, Albalat R, Capdevila M, Palacios O, Dallinger R 2017a. Biomphalaria glabrata Metallothionein: Lacking Metal Specificity of the Protein and Missing Gene Upregulation Suggest Metal Sequestration by Exchange Instead of through Selective Binding. Int J Mol Sci 18. doi: 10.3390/ijms18071457

Niederwanger M, Dvorak M, Schnegg R, Pedrini-Martha V, Bacher K, Bidoli M, Dallinger R 2017b. Challenging the Metallothionein (MT) Gene of Biomphalaria glabrata: Unexpected Response Patterns Due to Cadmium Exposure and Temperature Stress. Int J Mol Sci 18. doi: 10.3390/ijms18081747

Nielson KB, Winge DR 1985. Independence of the domains of metallothionein in metal binding. J Biol Chem 260: 8698-8701.

Orihuela R, Domenech J, Bofill R, You C, Mackay EA, Kagi JH, Capdevila M, Atrian S 2008. The metal-binding features of the recombinant mussel Mytilus edulis MT-10-IV metallothionein. J Biol Inorg Chem 13: 801-812. doi: 10.1007/s00775-008-0367-6

Palacios O, Jimenez-Marti E, Niederwanger M, Gil-Moreno S, Zerbe O, Atrian S, Dallinger R, Capdevila M 2017. Analysis of Metal-Binding Features of the Wild Type and Two Domain-Truncated Mutant Variants of Littorina littorea Metallothionein Reveals Its Cd-Specific Character. Int J Mol Sci 18. doi: 10.3390/ijms18071452

Palacios O, Pagani A, Pérez-Rafael S, Egg M, Höckner M, Brandstätter A, Capdevila M, Atrian S, Dallinger R 2011. Shaping mechanisms of metal specificity in a family of metazoan metallothioneins: evolutionary differentiation of mollusc metallothioneins. BMC Biol 9: 4. doi: 10.1186/1741-7007-9-4

Palacios O, Perez-Rafael S, Pagani A, Dallinger R, Atrian S, Capdevila M 2014. Cognate and noncognate metal ion coordination in metal-specific metallothioneins: the Helix pomatia system as a model. J Biol Inorg Chem 19: 923-935. doi: 10.1007/s00775-014-1127-4

Pedrini-Martha V, Koll S, Dvorak M, Dallinger R 2020. Cadmium Uptake, MT Gene Activation and Structure of Large-Sized Multi-Domain Metallothioneins in the Terrestrial Door Snail Alinda biplicata (Gastropoda, Clausiliidae). Int J Mol Sci 21. doi: 10.3390/ijms21051631

Perez-Rafael S, Atrian S, Capdevila M, Palacios O 2011. Differential ESI-MS behaviour of highly similar metallothioneins. Talanta 83: 1057-1061. doi: 10.1016/j.talanta.2010.10.060

Perez-Rafael S, Mezger A, Lieb B, Dallinger R, Capdevila M, Palacios O, Atrian S 2012. The metal binding abilities of Megathura crenulata metallothionein (McMT) in the frame of gastropoda MTs. J Inorg Biochem 108: 84-90. doi: 10.1016/j.jinorgbio.2011.11.025

Perez-Rafael S, Monteiro F, Dallinger R, Atrian S, Palacios O, Capdevila M 2014. Cantareus aspersus metallothionein metal binding abilities: the unspecific CaCd/CuMT isoform provides hints about the metal preference determinants in metallothioneins. Biochim Biophys Acta 1844: 1694-1707. doi: 10.1016/j.bbabap.2014.06.018
Purac J, Nikolic TV, Kojic D, Celic AS, Plavska JJ, Blagojevic DP, Petri ET 2019. Identification of a metallothionein gene in honey bee Apis mellifera and its expression profile in response to Cd, Cu and Pb exposure. Mol Ecol 28: 731-745. doi: 10.1111/mec.14984

Riek R, Precheur B, Wang Y, Mackay EA, Wider G, Guntert P, Liu A, Kagi JH, Wuthrich K 1999. NMR structure of the sea urchin (Strongylocentrotus purpuratus) metallothionein MTA. J Mol Biol 291: 417-428. doi: 10.1006/jmbi.1999.2967

Schmielau L, Dvorak M, Niederwanger M, Dobieszewski N, Pedrini-Martha V, Ladurner P, Pedregal JR, Marechal JD, Dallinger R 2018. Differential response to Cadmium exposure by expression of a two and a three-domain metallothionein isoform in the land winkle Pomatias elegans: Valuating the marine heritage of a land snail. Sci Total Environ 648: 561-571. doi: 10.1016/j.scitotenv.2018.07.426

Smith SA, Wilson NG, Goetz FE, Feehery C, Andrade SC, Rouse GW, Giribet G, Dunn CW 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480: 364-367. doi: 10.1038/nature10526

Tanguy A, Moraga D 2001. Cloning and characterization of a gene coding for a novel metallothionein in the Pacific oyster Crassostrea gigas (CgMT2): a case of adaptive response to metal-induced stress? Gene 273: 123-130. doi: 10.1016/s0378-1119(01)00577-7

Timmermans MJ, Ellers J, Roelofs D, van Straalen NM 2005. Metallothionein mRNA expression and cadmium tolerance in metal-stressed and reference populations of the springtail Orchesella cincta. Ecotoxicology 14: 727-739. doi: 10.1007/s10646-005-0020-x

Tio L, Villarreal L, Atrian S, Capdevila M 2004. Functional differentiation in the mammalian metallothionein gene family: metal binding features of mouse MT4 and comparison with its paralog MT1. J Biol Chem 279: 24403-24413. doi: 10.1074/jbc.M401346200

Toll-Riera M, Alba MM 2013. Emergence of novel domains in proteins. BMC Evol Biol 13: 47. doi: 10.1186/1471-2148-13-47

Toll-Riera M, Rado-Trilla N, Martys F, Alba MM 2012. Role of low-complexity sequences in the formation of novel protein coding sequences. Mol Biol Evol 29: 883-886. doi: 10.1093/molbev/msr263

Valls M, Bofill R, Gonzalez-Duarte R, Gonzalez-Duarte P, Capdevila M, Atrian S 2001. A new insight into metallothionein (MT) classification and evolution. The in vivo and in vitro metal binding features of Homarus americanus recombinant MT. J Biol Chem 276: 32835-32843. doi: 10.1074/jbc.M102151200

Weisman CM, Eddy SR 2017. Gene Evolution: Getting Something from Nothing. Curr Biol 27: R661-R663. doi: 10.1016/j.cub.2017.05.056

Yingprasertchai T, Yu RMK, Tran TKA, Chong Kong RY, O'Connor WA, MacFarlane GR 2019. Characterisation of the metallothionein gene in the Sydney rock oyster and its expression upon metal exposure in oysters with different prior metal exposure histories. Mar Environ Res 151: 104775. doi: 10.1016/j.marenvres.2019.104775

Zapata F, Wilson NG, Howison M, Andrade SC, Jorger KM, Schrodl M, Goetz FE, Giribet G, Dunn CW 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proc Biol Sci 281: 20141739. doi: 10.1098/rspb.2014.1739

Zhang C, Jansen M, De Meester L, Stoks R 2019. Rapid evolution in response to warming does not affect the toxicity of a pollutant: Insights from experimental evolution in heated mesocosms. Evol Appl 12: 977-988. doi: 10.1111/eva.12772
Zhang C, Jansen M, De Meester L, Stoks R 2018. Thermal evolution offsets the elevated toxicity of a contaminant under warming: A resurrection study in Daphnia magna. Evol Appl 11: 1425-1436. doi: 10.1111/eva.12637

Ziller A, Fraissinet-Tachet L 2018. Metallothionein diversity and distribution in the tree of life: a multifunctional protein. Metallomics 10: 1549-1559. doi: 10.1039/c8mt00165k
A.

**CAENOGASTROPODA**
- 36 MTs: \( \beta 3.1 \)
- 13 MTs: \( \beta 3.1, \beta 3.2 \)
- 4 MTs: \( \beta 3.1, \beta 3.2, \beta 3.3 \)
- 2 MTs: \( \beta 3.1, \beta 3.2, \beta 3.3, \beta 3.4 \)

**AbiMT2**
- \( \beta 3.1, \beta 3.2, \beta 3.3, \beta 3.4, \beta 3.5, \beta 3.6 \)

**McoMT2**
- \( \beta 3.1, \beta 3.2, \beta 3.3, \beta 3.4, \beta 3.5, \beta 3.6, \beta 3.7 \)

**β1**

B.

**HETEROBRANCHIA**
- 76 MTs: \( \beta 3.1 \)
- 2 MTs: \( \beta 3.1, \beta 3.2 \)

**Biomphalaria MTs**
- \( \beta 3.1, \beta 3.2, \beta 3.3 \)

**Abi MT3**
- \( \beta 3.1, \beta 3.2, \beta 3.3, \beta 3.4, \beta 3.5, \beta 3.6, \beta 3.7, \beta 3.8 \)

**Abi MT4**
- \( \beta 3.1, \beta 3.2, \beta 3.3, \beta 3.4, \beta 3.5, \beta 3.6, \beta 3.7, \beta 3.8, \beta 3.9 \)

**C-term**

**β1**
1. **β1 domain** $[\text{CxCl}]_2$-$\text{C}$-$[\text{CxCl}]_2$
2. **β2 domain** $[\text{CxCl}]_4$-$\text{C}$
3. **β3 domain** $\text{C}$-$[\text{CxCl}]_5$-$\text{C}$
4. **α domain** $[\text{CxCl}]_2$-$\text{C}$-$[\text{CxCl}]_2$-$\text{C}$