



Article

The Modular Architecture of Metallothioneins Facilitates Domain Rearrangements and Contributes to Their Evolvability in Metal-Accumulating Mollusks

Sara Calatayud ¹, Mario Garcia-Risco ², Veronika Pedrini-Martha ³, Michael Niederwanger ³, Reinhard Dallinger ³, Òscar Palacios ², Mercè Capdevila ² and Ricard Albalat ^{1,4,*}

¹ Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, E-08028 Barcelona, Spain

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

³ Department of Zoology and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria

⁴ Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, E-08028 Barcelona, Spain

* Correspondence: ralbalat@ub.edu



Citation: Calatayud, S.; Garcia-Risco, M.; Pedrini-Martha, V.; Niederwanger, M.; Dallinger, R.; Palacios, Ò.; Capdevila, M.; Albalat, R. The Modular Architecture of Metallothioneins Facilitates Domain Rearrangements and Contributes to Their Evolvability in Metal-Accumulating Mollusks. *Int. J. Mol. Sci.* **2022**, *23*, 15824. <https://doi.org/10.3390/ijms232415824>

Academic Editor: Andrzej Kloczkowski

Received: 10 November 2022

Accepted: 10 December 2022

Published: 13 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Protein domains are independent structural and functional modules that can rearrange to create new proteins. While the evolution of multidomain proteins through the shuffling of different preexisting domains has been well documented, the evolution of domain repeat proteins and the origin of new domains are less understood. Metallothioneins (MTs) provide a good case study considering that they consist of metal-binding domain repeats, some of them with a likely de novo origin. In mollusks, for instance, most MTs are bidomain proteins that arose by lineage-specific rearrangements between six putative domains: α , $\beta 1$, $\beta 2$, $\beta 3$, γ and δ . Some domains have been characterized in bivalves and gastropods, but nothing is known about the MTs and their domains of other Mollusca classes. To fill this gap, we investigated the metal-binding features of NpoMT1 of *Nautilus pompilius* (Cephalopoda class) and FcaMT1 of *Falcidens caudatus* (Caudofoveata class). Interestingly, whereas NpoMT1 consists of α and $\beta 1$ domains and has a prototypical Cd^{2+} preference, FcaMT1 has a singular preference for Zn^{2+} ions and a distinct domain composition, including a new Caudofoveata-specific δ domain. Overall, our results suggest that the modular architecture of MTs has contributed to MT evolution during mollusk diversification, and exemplify how modularity increases MT evolvability.

Keywords: protein modularity; domain repeat proteins; de novo domains; sORFs and lncRNAs; Cephalopoda and Caudofoveata MTs

1. Introduction

A protein domain is a well-defined region of a protein that constitutes a stable, independently folding, and compact structural unit that might perform a specific function [1–3]. Protein domains may occur alone, but are more frequently found in combination with other domains in multidomain proteins [4–6]. Novel multidomain proteins might arise by rearranging preexisting domains or by merging de novo emerged domains with old domains. While domain rearrangements have been analyzed extensively ([4,6–11]; reviewed in [12,13]), the rise and evolution of new domains are still poorly understood [14–16], and examples in which the new domains have been functionally characterized are very scarce [17]. In this scenario, metallothioneins (MTs), a heterogeneous family of proteins made of metal-binding domains, are becoming a valuable case study to infer events of the de novo emergence of domains and to understand the impact of a modular architecture on the evolution of novel proteins.

MTs have classically been defined as low molecular weight (<100 amino acids) and cysteine-rich ($\approx 15\text{--}30\%$) proteins that, because they bind transition metal ions and lack enzymatic functions, have been related with metal homeostasis and detoxification in living beings ([18,19]; reviewed in [20]). Their cysteine residues (Cys, C) are arranged in CC, CxC or CCC motifs, whose number and distribution define the structural domains that form metal–thiolate clusters. In mollusks, for instance, MT domains have been classified into six types: α domains (with 11/12 Cys), $\beta 1$, $\beta 2$ and $\beta 3$ domains (with 9 Cys, each one with different Cys motif arrangements), γ domains (with 10 Cys) and δ domains (with 14 Cys) [21]. Sequence analyses have led to suggest that most mollusk MTs would be bidomain proteins that combine an invariable carboxy-terminal $\beta 1$ domain with one of the other five domains (α , $\beta 2$, $\beta 3$, γ or δ) in a lineage-specific manner: an α domain in some Bivalvia, Cephalopoda, Polyplacophora and Solenogastres (also known as Neomeniamorpha) MTs, a $\beta 2$ domain in several Bivalvia, Monoplacophora, Scaphopoda and Solenogastres MTs, a $\beta 3$ domain in all Gastropoda MTs except in Patellogastropoda MTs that contain a γ domain, and a δ domain only found in Caudofoveata, (also known as Chaetodermomorpha) MTs [21–23].

Mollusk β and γ domains have been functionally characterized, and their metal preferences as well as their binding capacities have been determined [21,24]. In contrast, functional information on the α and δ domains is scarce or totally absent. To fill this gap, we have now analysed the MTs of two singular mollusk species, the “living fossil” *Nautilus pompilius* of the Cephalopoda class, and *Falcidens caudatus*, a Caudofoveata species of the Aculifera group that includes worm-like mollusk species that have no shell. Based on their sequences, the MTs of both species were predicted to be bidomain MTs, with α and $\beta 1$ domains for the *N. pompilius* MT (i.e., $\alpha\beta 1$ -NpoMT1), and δ and $\beta 1$ domains for the *F. caudatus* MT (i.e., $\delta\beta 1$ -FcaMT1) [21]. In this study, we determined the metal-binding abilities of both $\alpha\beta 1$ - and $\delta\beta 1$ -MTs characterizing by means of spectrometric (ESI-MS) and spectroscopic (ICP-AES) techniques the metal–protein complexes when these MTs are heterologously expressed in *E. coli*. We also analysed the features of α and δ domains when expressed alone, demonstrating their functional and structural independence to form metal–protein complexes. Our analyses highlight the relevance of the modular organization in MT evolvability by showing how functionally autonomous domains might arise and recombine into new MTs that have been adaptively selected according to the metal requirements and bioavailability in different mollusk lineages.

2. Results

2.1. Characterization of Cephalopoda and Caudofoveata MTs: NpoMT1 and FcaMT1

NpoMT1 from *N. pompilius* and FcaMT1 of *F. caudatus* were selected as representative Cephalopoda and Caudofoveata MTs, respectively. NpoMT1 was a 72 amino acid protein with 21 Cys (29%) organized in two putative domains: an amino-terminal α domain with 12 Cys arranged in a pattern of [CxC] $_5$ [CxC] $_3$ Cx $_4$ [CxC] $_3$ [CxC] $_3$ [CxC] $_2$ C motifs, linked by three residues to a carboxyl-terminal $\beta 1$ domain with its 9 Cys arranged in a pattern of [CxC] $_3$ [CxC] $_3$ Cx $_5$ [CxC] $_3$ [CxC] motifs (Figure 1). The patterns of Cys motifs in the α and $\beta 1$ domains of *Nautilus* NpoMT1 were more similar to those in other mollusk α and $\beta 1$ domains than those of other cephalopod species such as *Sepia* and *Octopus*, suggesting that they represented the ancestral forms [21]. FcaMT1 was a 74 amino acid protein with 23 Cys (31%) organized in two predicted domains: an amino-terminal δ domain with 14 Cys arranged in a pattern of [CxC] $_3$ [CxC] $_3$ Cx $_4$ [CxC] $_3$ CCx $_4$ [CxCC] motifs, linked by three residues to a carboxyl-terminal archetypal $\beta 1$ domain (Figure 1). Importantly, the δ domain has so far only been found in Caudofoveata MTs, suggesting a lineage-specific innovation [21].

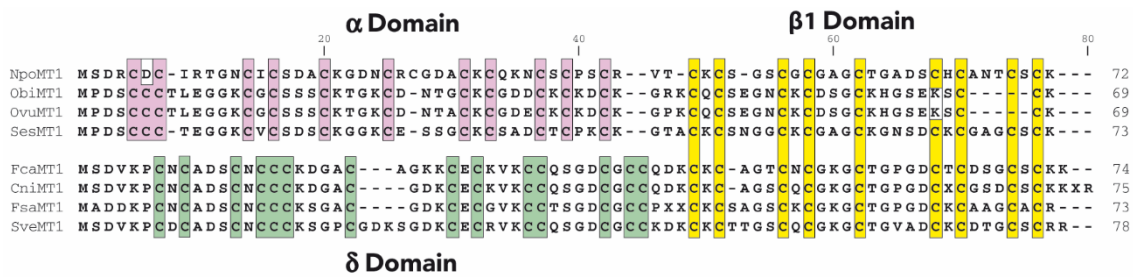


Figure 1. Amino acid alignments of Cephalopoda and Caudofoveata MTs. At the top of the figure, alignment of *Nautilus pompilius* (NpoMT1), *Octopus bimaculoides* (ObiMT1), *Octopus vulgaris* (OvuMT1) and *Sepia esculenta* (SesMT1) MTs showing the bidomain $\alpha\beta 1$ structure of Cephalopoda MTs. At the bottom, alignment of *Falcidens caudatus* (FcaMT1), *Chaetoderma nitidulum* (CniMT1), *Falcidens sagittiferus* (FsaMT1) and *Scutopus ventrolineatus* (SveMT1) MTs showing the $\delta\beta 1$ structure of Caudofoveata MTs. Cysteines are highlighted in pink in Cephalopoda α domains, in green in Caudofoveata δ domains and in yellow in the $\beta 1$ domains shared by both groups of species. Sequences are from [21].

In order to characterize the biochemical properties and metal-binding features of NpoMT1 and FcaMT1, we took advantage of a well-established procedure for MT analysis, which determines the metal preference and binding capacity of any MT (or MT domain) based on the study of metalated species in recombinantly produced metal–protein complexes [20]. Thus, both proteins were expressed as GST–fusion proteins in *E. coli* BL21 cultures supplemented with $ZnCl_2$, $CdCl_2$ or $CuSO_4$. Metal–protein complexes were purified from *E. coli* total protein extracts by a GST–affinity system, followed by the cleavage with thrombin of the GST tag, and FPLC chromatography. The FPLC fractions containing the metal–MT complexes were characterized by ESI–MS analyses. The experimental masses corresponding to apo-NpoMT1 and apo-FcaMT1 (7471.8 Da and 7593.6Da, respectively) after demetallation by acidification of the NpoMT1 and FcaMT1 productions were consistent with the theoretical masses (7472.52 Da and 7594.79 Da, respectively) (Figure 2A,C), verifying the identity of the recombinantly expressed MTs.

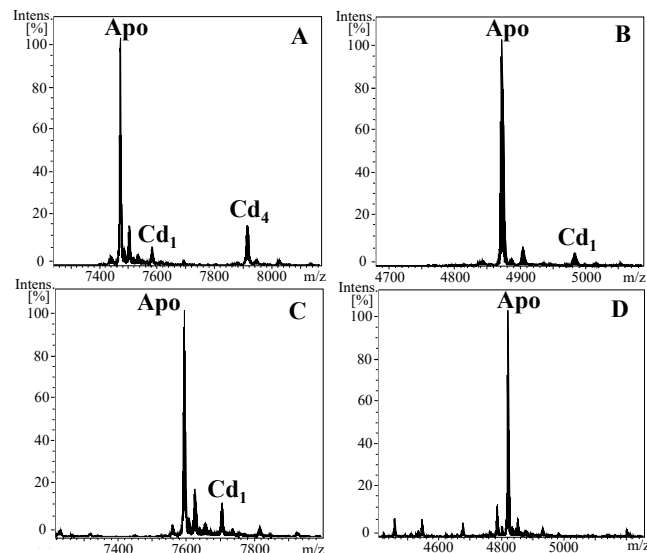


Figure 2. Deconvoluted ESI–MS spectra of the apo–MTs obtained after the demetallation by acidification (pH 2.4) of recombinant Cd–protein complexes. Notice that in the production process, the digestion with thrombin of the GST–MT fusion proteins resulted in the addition of two extra residues (glycine and serine) at the N–terminal end of the purified MTs that, as shown in previous studies, did not interfere with the metal–binding features of recombinant MTs [25,26]. (A) apo–NpoMT1, (B) apo– α NpoMT1, (C) apo–FcaMT1 and (D) apo– δ FcaMT1.

The ICP-AES and ESI-MS analyses of the NpoMT1 produced in cadmium (Cd)-enriched culture medium yielded a unique metalated species of NpoMT1 coordinating seven Cd²⁺ ions (Cd₇-NpoMT1, Figure 3A). This result indicated an efficient and thermodynamically favored binding of NpoMT1 with Cd²⁺ ions. This fact was corroborated by the reluctance of some metal clusters to exchange their metals by protons in acidic conditions (see Cd₄-NpoMT1 in Figure 2A). In contrast, when NpoMT1 was produced in zinc (Zn) surplus conditions, it rendered a mixture of Zn₆- and Zn₇-NpoMT1 as major species (Figure 3B), suggesting that although NpoMT1 binds both divalent metal ions, Cd²⁺ and Zn²⁺, it had a higher preference for Cd²⁺ than for Zn²⁺ ions. These results were additionally supported by the presence of small but perceptible amounts of glycosylated MT species in the productions under Zn²⁺ surplus conditions (peak denoted by an asterisk in Figure 3B), which is characteristic of partly structured Cd-thioneins when produced in absence of their preferred metal [27]. Finally, the productions of NpoMT1 in copper (Cu) surplus conditions also sustained its Cd-specificity. Cd-thioneins typically yield a great variety of homometallic Cu-MT complexes when expressed in Cu-enriched medium, whereas Zn-thioneins produce heterometallic Zn/Cu-MT complexes [24,28]. Our results clearly showed many homometallic Cu-NpoMT1 complexes, confirming its Cd-thionein character (Figure 3C). The Cd preference of NpoMT1 agrees with the association between MT expression and the elevated levels of Cd (but not Zn) reported for *Octopus vulgaris*, another cephalopod species [29]. Based on all this evidence, we classified NpoMT1 as a genuine Cd-thionein.

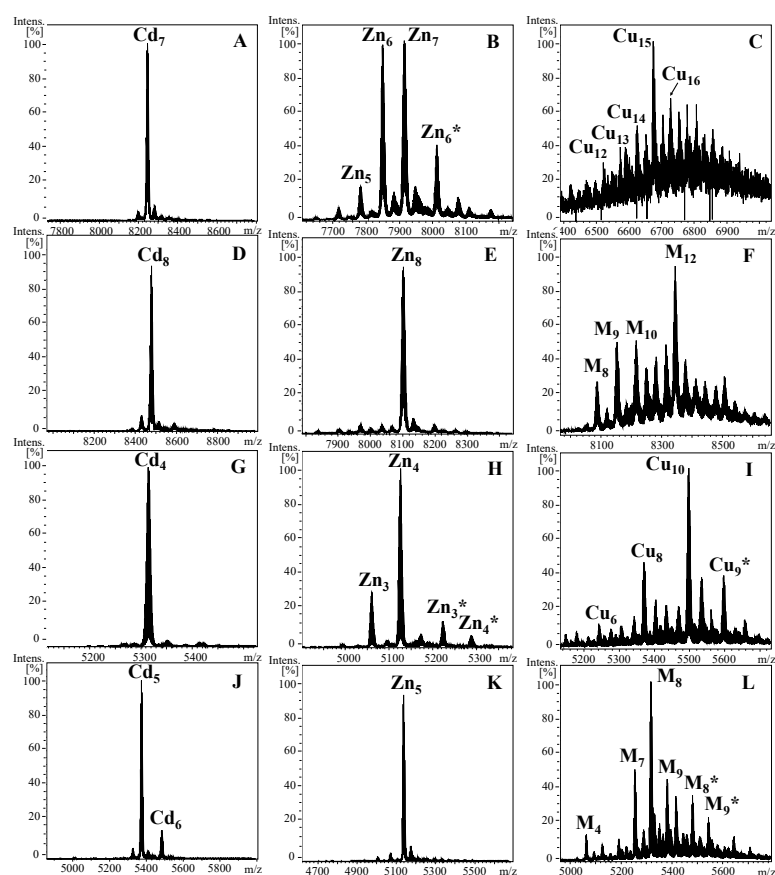


Figure 3. Deconvoluted ESI-MS spectra of metal–protein complexes recombinantly produced in *E. coli* growth in metal-enriched culture medium. (A) Cd-NpoMT1, (B) Zn-NpoMT1, (C) Cu-NpoMT1, (D) Cd-FcaMT1, (E) Zn-FcaMT1, (F) Cu-FcaMT1, (G) Cd- α NpoMT1, (H) Zn- α NpoMT1, (I) Cu- α NpoMT1, (J) Cd- δ FcaMT1, (K) Zn- δ FcaMT1 and (L) Cu- δ FcaMT1. M in panels F and L stands for heterometallic (Zn + Cu)-MT complexes. Glycosylated metal-MT species are depicted with an asterisk (*).

Regarding FcaMT1, the analysis of metal-FcaMT1 complexes revealed some interesting differences in comparison with NpoMT1. A unique metal-MT species (Zn_8 - and Cd_8 -FcaMT) was observed when FcaMT1 was synthesized in Zn^{2+} - or Cd^{2+} -enriched medium (Figure 3D,E). This result contrasted with the mixture of Zn^{2+} -MT species observed for NpoMT1, and suggested that FcaMT1 was able to bind both divalent metal ions with a similar efficiency. To clarify whether FcaMT1 had a more Zn-thionein or Cd-thionein character, we took advantage of additional information provided by the results of the glycosylation analysis and the productions under Cu surplus conditions. The observation of heterometallic Zn,Cu-FcaMT1 complexes when expressed in Cu-enriched medium (Figure 3F) indicated a strong Zn preference for FcaMT1, which was corroborated by the absence of glycosylated forms when produced under Zn surplus conditions due to a high structuration level of the FcaMT1 coordinating Zn^{2+} ions (Figure 3E) [27]. Based on these results, we concluded that FcaMT1 showed a clear Zn-thionein behavior.

2.2. Characterization of the α and δ Domains of Cephalopoda and Caudofoveata MTs

In order to characterize the biochemical properties and metal-binding features of the predicted α (from Met₁ to Thr₄₄) and δ (from Met₁ to Lys₄₅) domains of NpoMT1 and FcaMT1, respectively, we followed the same strategy as for the full-length proteins. In brief, both putative domains were expressed as GST-fusion proteins in *E. coli* BL21 cultures supplemented with $ZnCl_2$, $CdCl_2$ or $CuSO_4$, and metal-protein complexes were purified and characterized by ICP-AES and ESI-MS analyses. The experimental masses corresponding to the apo- α and apo- δ domains (4872.0 Da and 4821.0 Da, respectively) were consistent with their theoretical masses (4872.57 Da and 4821.59 Da, respectively) (Figure 2B,D), confirming the identity of both recombinantly produced domains.

The α domain of NpoMT1 produced in Cd-enriched culture medium yielded a unique species with four Cd^{2+} ions (Cd_4 -NpoMT1, Figure 3G), whereas when it was produced in Zn surplus conditions, it rendered a mixture of Zn_3 - and Zn_4 - α complexes (Figure 3H), suggesting a higher preference for Cd^{2+} than for Zn^{2+} ions. The presence of glycosylated MT species under Zn surplus conditions (peaks denoted by an asterisk in Figure 3H), and the yield of homometallic Cu-MT complexes when expressed in Cu-enriched medium (Figure 3I) supported the Cd-specificity of the α domain when expressed alone. In contrast, for the δ domain of FcaMT1, a unique metal-domain species (Zn_5 - δ FcaMT1) was observed when it was synthesized in Zn-enriched medium (Figure 3K), and two species, a major Cd_5 - δ FcaMT1 and a minor Cd_6 - δ FcaMT1, when it was produced in Cd surplus conditions (Figure 3E). These results, together with the absence of glycosylated forms and with the formation of heterometallic Zn/Cu- δ domain complexes in Cu-enriched preparations (Figure 3L) sustained a Zn preference of δ FcaMT1 when produced alone. In summary, our results showed that α and δ domains yield unique Cd^{2+} - and Zn^{2+} -complexes, respectively (Figure 3G,K), indicating that they form single and structurally well-defined metal clusters, and supporting that α and δ domains are functionally autonomous MT domains.

3. Discussion

3.1. Cephalopoda NpoMT1 and Caudofoveata FcaMT1, and the α and δ Domains

Metallothioneins are becoming a valuable model system to investigate the origin and evolution of protein domains for two main reasons. First, based on their patterns of Cys motifs, MT domains can be predicted in broad phylogenetic contexts, and thereby, the changes in their domain organization can be visualized along the evolutionary ladder. Second, standardized protocols of analytical techniques (ESI-MS and ICP-AES) allow the functional characterization of the predicted MT domains based on their metal-binding features, that is, their metal-binding preferences and capacities, exposing the functional consequences of the innovations in domain rearrangements. Several mollusk MTs and some domains had previously been functionally characterized in the Bivalvia and Gastropoda classes, but nothing was known about the MTs from the Cephalopoda class of the Conchifera clade, nor from any Caudofoveata MTs of the Aculifera clade. We have now characterized the

putative MTs from Cephalopoda *N. pompilius* (NpoMT1) and Caudofoveata *F. caudatus* (FcaMT1) species, and their distinctive α and δ domains. We determined the MT features of both MTs, establishing that NpoMT1 and its α domain render homometallic complexes with seven and four Cd^{2+} ions, respectively (Figure 3A,G), whereas FcaMT1 and its δ domain coordinate eight and five zinc ions, respectively (Figure 3E,K). Considering that β 1 domains bind three divalent metal ions [21,24], the metal-to-protein stoichiometries of the α and δ domains fully agree with those of the full-length MTs: $4(\alpha) + 3(\beta 1) = 7\text{Cd}^{2+}$ ions for NpoMT1, and $5(\delta) + 3(\beta 1) = 8\text{Zn}^{2+}$ ions for FcaMT1. These stoichiometries show a higher metal-binding capacity of FcaMT1 in comparison with NpoMT1 due to the two extra Cys at its N-terminal domain: 14 Cys in the δ domain versus 12 Cys in the α domain. Our results also indicate that the α and δ domains are functionally autonomous, and experimentally confirm the predicted bidomain $\alpha\beta$ 1 and $\delta\beta$ 1 structures of Cephalopoda and Caudofoveata MTs, respectively.

Importantly, our results revealed a different metal preference for Cephalopoda and Caudofoveata MTs, Cd specificity for NpoMT1 and Zn specificity for FcaMT1. These metal preferences coincide with those of their N-terminal domains (Cd^{2+} ions for the α domain, and Zn^{2+} ions for the δ domain) indicating that the distinctive N-terminal domain is fundamental for the metal preference of the full-length proteins. From an evolutionary perspective, considering the ancient origin of the α domain and the pervasiveness of Cd-specific MTs in mollusks [21,30], our results suggest that the Zn-specific FcaMT1 was a functional innovation in the *F. caudatus* lineage, and thereby, likely associated to the appearance of the δ domain in the Caudofoveata class. Because marine mollusk species typically have Cd-thioneins [30], we investigated the possibility that Caudofoveata species might have an additional “conventional” $\alpha\beta$ 1- or β 2 β 1-MT with Cd preference. Database surveys did not retrieve any putative Caudofoveata MT supporting this possibility, although we cannot definitively rule it out because the available Caudofoveata databases are still incomplete. Interestingly, however, we identified a second putative MT in *F. caudatus*, FcaMT2, made up of a “degenerated” δ -like domain (48% identical to the δ domain of FcaMT1) maintaining only 9 of the 14 Cys, fused to a conserved β 1 domain (76% identical to the β 1 domain of FcaMT1) (Supplementary Figure S1). Whether FcaMT2 is a Zn-specific MT or it may be a Cd-thionein cannot be inferred from sequence analysis, and it requires further investigation.

3.2. Biological Implications of the Cephalopoda and Caudofoveata MTs

In some cases, it has been possible to relate the chemical findings on the metal-binding properties of some MTs with their biological functions. This has been achieved, for instance, in the case of the Cd- and Cu-specific MT isoforms of the terrestrial snails *Helix pomatia* and *Cornu aspersum*. The chemical properties of these MTs agree with the biological findings on their organ- and cell-specific accumulation and distribution as well as with their presumed role in the metal metabolism of these two snail species [31]. In the case of the MTs analysed in this work, although there are no data on the metal metabolism of *N. pompilius*, results from other cephalopod species consistently indicate that cephalopods preferentially accumulate Cd in their midgut gland, where they can detoxify it by binding to “metallothionein-like proteins”, including metallothioneins [29,32,33]. The closely related species *Nautilus macromphalus*, for example, has been reported to accumulate huge amounts of many different metals in its digestive gland, which may serve as a storage and detoxification organ for most of them too, probably by binding to MTs and by intracellular compartmentalization [32]. Interestingly, another study showed that some cephalopod species collected near the submarine volcano Tagoro (Canary Islands) show highly elevated concentrations for nine different metals, including Cd, apparently due to volcanic emissions of metal-containing compounds [34]. Against this background, it is therefore highly comprehensible that cephalopods have Cd-specific MTs. In contrast, biological data that could explain the analytically proven zinc preference of the metallothionein of *F. caudatus*, are completely lacking. This is mainly because there are hardly any studies

on metal accumulation in Caudofoveates. One reason for this is probably that the class Caudofoveata is small and most representatives of this class live burrowing in the seabed and in deep-sea sediments [35], which makes it very difficult to perform, for example, metal exposure experiments with these animals in the laboratory. However, based on our results, it is not unreasonable to expect that zinc plays an important role in the physiology and ecology of this peculiar group of animals.

3.3. Origin and Evolution of MT Domains

Based on the phylogenetic distribution of the different MT domains, it is thought that the origin of α and $\beta 1$ domains predated the diversification of the Mollusca phylum, while the δ domain would have evolved de novo, probably concurring with the appearance of the Caudofoveata class [21]. De novo evolution events in the MT world seem to have been more frequent than previously thought. The lineage-specific γ and δ domains in mollusks ([21] and this work), the evolution of a new 12C domain in appendicularian tunicates [36,37] or the recent evolutionary origin proposed for the yeast metallothionein CUP1 [38] sustain this idea, and support the polyphyly of MTs implicitly suggested in diverse evolutionary studies [19,20,39,40]. This polyphyly implies that although MTs form a seemingly homogenous group of proteins, they are not, and the domains or proteins grouped in the MT family would be, in fact, the result of the convergent evolution of independent sequences driven by the chemical requirements of metal coordination such as a high content of coordinating residues (i.e., cysteines) and a relatively small length that would favor its proper folding [20,21,30].

We do not know the source of the new MT domains or proteins, but in the last years data have accumulated showing that short open reading frames (sORFs) in long non-coding RNAs (lncRNAs) might be an important source of de novo peptides [41–43]. These sORFs have a median size of around 43 amino acids (although some of them might encode peptides longer than 100 amino acids) [44], and because they are lineage-specific, they are poorly conserved across species [41,42]. Both features, size and conservation, match those of the de novo MT domains well. The δ domain of FcaMT1 or the γ domain of LgiMT2, for instance, are 45 amino acid peptides that are specific to Caudofoveata and Patellogastropoda lineages, respectively [21,30]; the 12C domain of OdiMT1 is an appendicularian-specific peptide of 42 amino acids, different from the 12C domains of other chordate MTs [36,37]; and yeast CUP1 is a genus-specific MT [38] with its cysteines distributed throughout 44 of its 61 amino acids [45]. It is not too daring to hypothesize, therefore, that the sORFs of lncRNAs might have been an important source of novel MT domains or proteins.

3.4. MT Modularity and Evolvability

Protein modularity accelerates the evolutionary process of generating genetic diversity because it favors the domain shuffling, which is an effective means of exploring new viable sequences and creating new proteins as it has a greater impact on the functional properties of proteins than the progressive accumulation of point mutations in their sequences [11,46]. Modularity therefore increases protein evolvability by favouring the emergence of new protein variants that can be adaptively selected ([47] and references therein). Our analysis of NpoMT1 and FcaMT1 supports that the modular architecture of MTs contributed to their evolvability because it provides experimental evidence that the diverse mollusk MTs came from the rearrangements of structurally intact and functionally autonomous domains. Our results also show that domain rearrangements might involve not only preexisting domains (e.g., α and $\beta 1$ domain) but also de novo domains (e.g., δ domain), raising the number of distinct domains available to create MTs with new functional abilities (Figure 4).

Altogether, our results substantiate that the modular structure of MTs would increase their evolvability by facilitating the appearance of new forms (e.g., $\delta\beta 1$ -FcaMT1) with different metal-binding properties (i.e., a higher metal-binding capacity and a new preference for Zn^{2+} ions) that would have been selected depending on the species-specific adaptive requirements in the Caudofoveata lineage (e.g., new physiological or ecological roles of

zinc in the Caudofoveata lineage). Furthermore, the modular architecture of MTs would also have facilitated the duplications and losses of complete domains in a lineage-specific manner. For instance, losses of α and β domains occurred during the evolution of some mollusk lineages, including the Caudofoveata species, while these domains have been tandem duplicated in several gastropod and bivalve species, giving rise to large multidomain forms (Figure 4) [21,24,37,48,49]. In fact, duplications and losses of complete domains have also been described in MTs from other organisms, including species of the Chordata phylum [36,37,50]. In summary, it can be concluded that working as structurally and functionally independent modules, MT domains have been recurrently gained, lost, modified or recombined in mollusks and in other animal phyla, and therefore it is not unreasonable to assume that by favoring structural and functional innovation, the modular architecture of MTs has played a key role in the origin and evolution of most metazoan MTs.

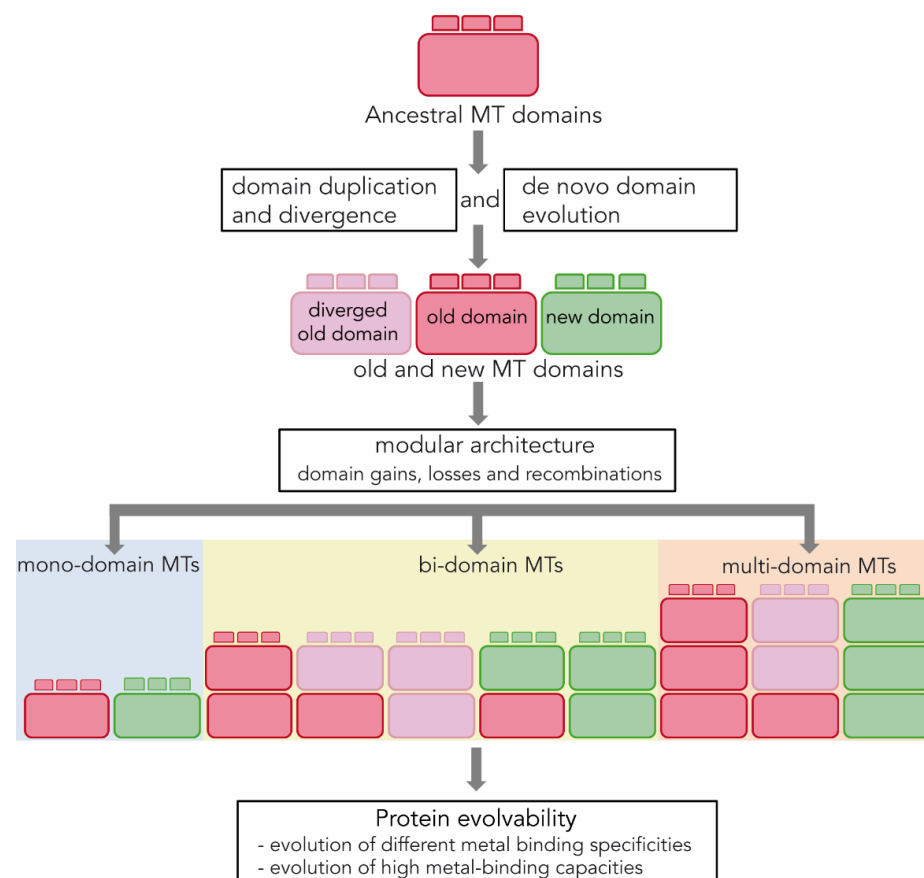


Figure 4. Modular architecture and MT evolvability. MTs are proteins made up of self-stabilizing, independent folding and structurally compact modules called “domains”. Some MT domains are old and date back to the metazoan origins (red), but others are more recent and would have arisen by duplication and divergence mechanisms (pink) or by de novo evolution events (green) during the diversification of different animal lineages. Most animal MTs, including those of mollusks and chordates, are bidomain proteins. This modular architecture favors the rearrangements of old and new domains leading to novel domain combinations in bidomain forms, or to domain gains and losses in multidomain and monodomain MTs, respectively. Several MTs have been described that exemplify the different rearrangements depicted in this figure [21,24,36,37,49,50]. Overall, our results suggest that the modular architecture of the MTs increases their evolvability by facilitating the appearance of new forms with different metal binding specificities (i.e., Cd-, Zn- or Cu-thioneins) or with augmented metal-binding capacities (i.e., from monodomain to bidomain and multidomain forms) that can be adaptively selected according to the metal requirements and bioavailability of the different animal lineages.

4. Materials and Methods

4.1. Production and Purification of Recombinant Metal-MT Complexes

Production and purification of recombinant metal-MT complexes of NpoMT1, FcaMT1, α NpoMT1 and δ FcaMT1 domains were performed as described elsewhere [37,50]. In brief, synthetic cDNAs codifying the different constructs were provided by Synbio Technologies (Monmouth Junction, NJ, USA), cloned in the pGEX-4T-1 expression vector (GE Healthcare, Chicago, IL, USA) and transformed in protease-deficient *E. coli* BL21 strain. Metal-MT complexes were produced in *E. coli* BL21 cultures expressing the recombinant plasmids, after induction with isopropyl- β -D-thiogalactopyranoside (100 μ M) and supplementation with ZnCl₂ (300 μ M), CdCl₂ (300 μ M) or CuSO₄ (500 μ M). Metal-MT complexes were purified from the soluble protein fraction of sonicated bacteria by affinity purification of the GST-tagged proteins, and digestion with thrombin. The 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 and run at 0.8 mL min⁻¹. The protein-containing fractions, identified by their absorbance at 254 nm, were pooled and stored at -80 °C until use.

4.2. Analysis of Metal-MT Complexes

All designed constructs, NpoMT1, FcaMT1, α NpoMT1 (from Met₁ to T₄₄) and δ FcaMT1 (from Met₁ to K₄₅), were characterized by means of mass spectrometry (ESI-MS) and spectroscopy (ICP-AES). An electrospray ionization mass spectrometry (ESI-MS) MicroTOF-Q Instrument (Bruker Daltonics GmbH, Bremen, Germany) interfaced with a Series 1100 HPLC pump (Agilent Technologies, Santa Clara, CA, USA), was used to determine the molecular mass of the recombinant proteins. The instrument was calibrated with ESI-L Low Concentration Tuning Mix (Agilent Technologies) and the experimental conditions were set up as follows: injection of 10–20 μ L of sample through a PEEK long tube (1–1.5 m \times 0.18 mm i.d.) at 30–50 μ L min⁻¹; capillary-counter-electrode voltage, 3.5–5.0 kV; desolvation temperature, 90–110 °C; dry gas, 6 L min⁻¹. Data were acquired over an m/z range of 800 to 3000. The liquid carriers were a 90:10 mixture of 15 mM ammonium acetate and acetonitrile at pH 7.0 and a 95:5 mixture of formic acid and acetonitrile at pH 2.4.

Element concentrations of S, Zn, Cd and Cu in the sample were determined by Inductively Coupled Plasma Atomic Emission Spectroscopy (ICP-AES) by means of a PerkinElmer Optima 4300DV (Waltham, MA, USA) at the correct wavelength (S, 182.04 nm; Zn, 213.86 nm; Cd, 228.80 nm; Cu, 324.80 nm) under conventional conditions [51]. MTs' concentration was calculated based on the S concentration obtained by ICP-AES, assuming that all the sulfur measured comes from peptides' Cys and Met residues.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232415824/s1>, Figure S1: Amino acid alignment of *Falciidens caudatus* MTs.

Author Contributions: Conceptualization, S.C., R.D., M.C., Ò.P. and R.A.; methodology, S.C., M.G.-R., V.P.-M., M.N. and R.A.; formal analysis, R.A. and S.C.; investigation, S.C., M.G.-R., V.P.-M., M.N. and R.A.; resources, R.D., M.C., Ò.P. and R.A.; writing—original draft preparation, R.A. and S.C.; writing—review and editing, S.C., M.G.-R., V.P.-M., M.N., R.D., Ò.P., M.C. and R.A.; supervision, R.D., M.C., Ò.P. and R.A.; project administration, R.D., M.C., Ò.P. and R.A.; funding acquisition, R.D., M.C., Ò.P. and R.A.; All authors have read and agreed to the published version of the manuscript.

Funding: R.A. was supported by PID2021-123258NB-I00, and M.C. and Ò.P. were supported by BIO2015-67358-C2-2-P from the Spanish Ministerio de Ciencia e Innovación. R.A. was a member of the “Grup de Recerca de la Generalitat de Catalunya,” ref. 2017SGR-1665, and M.C. and Ò.P. of the ref. 2017SGR-864. M.G.-R. acknowledges the UAB for the PIF grant. Parts of this study were co-financed by a cooperation grant to R.D. from the Austrian Science Fund (DACH grant No I 3032-B21).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We acknowledge the Centres Científics i Tecnològics (CCiT) de la Universitat de Barcelona (DNA sequencing) and the Servei d'Anàlisi Química (SAQ) de la Universitat Autònoma de Barcelona (ICP-AES, ESI-MS) for allocating instrument time. We also thank to Sebastian Artime for experimental support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wetlaufer, D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 697–701. [[CrossRef](#)] [[PubMed](#)]
2. Janin, J.; Chothia, C. Domains in proteins: Definitions, location, and structural principles. *Methods Enzym.* **1985**, *115*, 420–430. [[CrossRef](#)]
3. Heringa, J. Protein domains. In *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2005.
4. Apic, G.; Gough, J.; Teichmann, S.A. Domain combinations in archaeal, eubacterial and eukaryotic proteomes. *J. Mol. Biol.* **2001**, *310*, 311–325. [[CrossRef](#)] [[PubMed](#)]
5. Levitt, M. Nature of the protein universe. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11079–11084. [[CrossRef](#)]
6. Wang, M.; Caetano-Anolles, G. The evolutionary mechanics of domain organization in proteomes and the rise of modularity in the protein world. *Structure* **2009**, *17*, 66–78. [[CrossRef](#)]
7. Björklund, A.K.; Ekman, D.; Light, S.; Frey-Skott, J.; Elofsson, A. Domain rearrangements in protein evolution. *J. Mol. Biol.* **2005**, *353*, 911–923. [[CrossRef](#)]
8. Ekman, D.; Björklund, A.K.; Frey-Skott, J.; Elofsson, A. Multi-domain proteins in the three kingdoms of life: Orphan domains and other unassigned regions. *J. Mol. Biol.* **2005**, *348*, 231–243. [[CrossRef](#)]
9. Kummerfeld, S.K.; Teichmann, S.A. Relative rates of gene fusion and fission in multi-domain proteins. *Trends Genet.* **2005**, *21*, 25–30. [[CrossRef](#)]
10. Fong, J.H.; Geer, L.Y.; Panchenko, A.R.; Bryant, S.H. Modeling the evolution of protein domain architectures using maximum parsimony. *J. Mol. Biol.* **2007**, *366*, 307–315. [[CrossRef](#)]
11. Dohmen, E.; Klasberg, S.; Bornberg-Bauer, E.; Perrey, S.; Kemena, C. The modular nature of protein evolution: Domain rearrangement rates across eukaryotic life. *BMC Evol. Biol.* **2020**, *20*, 30. [[CrossRef](#)]
12. Moore, A.D.; Björklund, A.K.; Ekman, D.; Bornberg-Bauer, E.; Elofsson, A. Arrangements in the modular evolution of proteins. *Trends Biochem. Sci.* **2008**, *33*, 444–451. [[CrossRef](#)] [[PubMed](#)]
13. Forslund, S.K.; Kaduk, M.; Sonnhammer, E.L.L. Evolution of Protein Domain Architectures. *Methods Mol. Biol.* **2019**, *1910*, 469–504. [[CrossRef](#)] [[PubMed](#)]
14. Kersting, A.R.; Bornberg-Bauer, E.; Moore, A.D.; Grath, S. Dynamics and adaptive benefits of protein domain emergence and arrangements during plant genome evolution. *Genome Biol. Evol.* **2012**, *4*, 316–329. [[CrossRef](#)] [[PubMed](#)]
15. Moore, A.D.; Bornberg-Bauer, E. The dynamics and evolutionary potential of domain loss and emergence. *Mol. Biol. Evol.* **2012**, *29*, 787–796. [[CrossRef](#)]
16. Toll-Riera, M.; Alba, M.M. Emergence of novel domains in proteins. *BMC Evol. Biol.* **2013**, *13*, 47. [[CrossRef](#)]
17. Margolin, J.F.; Friedman, J.R.; Meyer, W.K.; Vissing, H.; Thiesen, H.J.; Rauscher, F.J., 3rd. Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 4509–4513. [[CrossRef](#)]
18. Vasak, M.; Meloni, G. Chemistry and biology of mammalian metallothioneins. *J. Biol. Inorg. Chem.* **2011**, *16*, 1067–1078. [[CrossRef](#)]
19. Ziller, A.; Fraissinet-Tachet, L. Metallothionein diversity and distribution in the tree of life: A multifunctional protein. *Metallomics* **2018**, *10*, 1549–1559. [[CrossRef](#)]
20. Capdevila, M.; Atrian, S. Metallothionein protein evolution: A miniassay. *J. Biol. Inorg. Chem.* **2011**, *16*, 977–989. [[CrossRef](#)]
21. Calatayud, S.; Garcia-Risco, M.; Pedrini-Martha, V.; Eernisse, D.J.; Dallinger, R.; Palacios, O.; Capdevila, M.; Albalat, R. Modularity in Protein Evolution: Modular Organization and De Novo Domain Evolution in Mollusk Metallothioneins. *Mol. Biol. Evol.* **2021**, *38*, 424–436. [[CrossRef](#)]
22. Jenny, M.J.; Payton, S.L.; Baltzegar, D.A.; Lozier, J.D. Phylogenetic Analysis of Molluscan Metallothioneins: Evolutionary Insight from *Crassostrea virginica*. *J. Mol. Evol.* **2016**, *83*, 110–125. [[CrossRef](#)] [[PubMed](#)]
23. Nam, Y.K.; Kim, E.J. Diversification and domain evolution of molluscan metallothioneins: A mini review. *Fish. Aquat. Sci.* **2017**, *20*, 1–18. [[CrossRef](#)]
24. Garcia-Risco, M.; Calatayud, S.; Pedrini-Martha, V.; Albalat, R.; Dallinger, R.; Palacios, O.; Capdevila, M. Metal-Specificity Divergence between Metallothioneins of *Nerita peloronta* (Neritimorpha, Gastropoda) Sets the Starting Point for a Novel Chemical MT Classification Proposal. *Int. J. Mol. Sci.* **2021**, *22*, 13114. [[CrossRef](#)]
25. Cols, N.; Romero-Isart, N.; Capdevila, M.; Oliva, B.; Gonzalez-Duarte, P.; Gonzalez-Duarte, R.; Atrian, S. Binding of excess cadmium(II) to Cd7-metallothionein from recombinant mouse Zn7-metallothionein 1. UV-VIS absorption and circular dichroism studies and theoretical location approach by surface accessibility analysis. *J. Inorg. Biochem.* **1997**, *68*, 157–166. [[CrossRef](#)]

26. Espart, A.; Marin, M.; Gil-Moreno, S.; Palacios, O.; Amaro, F.; Martin-Gonzalez, A.; Gutierrez, J.C.; Capdevila, M.; Atrian, S. Hints for metal-preference protein sequence determinants: Different metal binding features of the five tetrahymena thermophila metallothioneins. *Int. J. Biol. Sci.* **2015**, *11*, 456–471. [[CrossRef](#)] [[PubMed](#)]
27. Garcia-Risco, M.; Gonzalez, A.; Calatayud, S.; Lopez-Jaramillo, F.; Pedrini-Martha, V.; Albalat, R.; Dallinger, R.; Dominguez-Vera, J.M.; Palacios, O.; Capdevila, M. Glycosilation in Recombinant Metallothioneins. *Chem. Commun.* **2022**. [[CrossRef](#)] [[PubMed](#)]
28. Palacios, O.; Atrian, S.; Capdevila, M. Zn- and Cu-thioneins: A functional classification for metallothioneins? *J. Biol. Inorg. Chem.* **2011**, *16*, 991–1009. [[CrossRef](#)]
29. Raimundo, J.; Costa, P.M.; Vale, C.; Costa, M.H.; Moura, I. Metallothioneins and trace elements in digestive gland, gills, kidney and gonads of *Octopus vulgaris*. *Comp. Biochem. Physiol. C Toxicol. Pharm.* **2010**, *152*, 139–146. [[CrossRef](#)]
30. Dallinger, R.; Zerbe, O.; Baumann, C.; Egger, B.; Capdevila, M.; Palacios, O.; Albalat, R.; Calatayud, S.; Ladurner, P.; Schlick-Steiner, B.C.; et al. Metallomics reveals a persisting impact of cadmium on the evolution of metal-selective snail metallothioneins. *Metallomics* **2020**, *12*, 702–720. [[CrossRef](#)]
31. Palacios, O.; Pagani, A.; Perez-Rafael, S.; Egg, M.; Hockner, M.; Brandstatter, A.; Capdevila, M.; Atrian, S.; Dallinger, R. Shaping mechanisms of metal specificity in a family of metazoan metallothioneins: Evolutionary differentiation of mollusc metallothioneins. *BMC Biol.* **2011**, *9*, 4. [[CrossRef](#)]
32. Bustamante, P.; Cosson, R.P.; Gallien, I.; Caurant, F.; Miramand, P. Cadmium detoxification processes in the digestive gland of cephalopods in relation to accumulated cadmium concentrations. *Mar. Environ. Res.* **2002**, *53*, 227–241. [[CrossRef](#)] [[PubMed](#)]
33. Penicaud, V.; Lacoue-Labarthe, T.; Bustamante, P. Metal bioaccumulation and detoxification processes in cephalopods: A review. *Environ. Res* **2017**, *155*, 123–133. [[CrossRef](#)] [[PubMed](#)]
34. Lozano-Bilbao, E.; Lozano, G.; Gutierrez, A.J.; Hardisson, A.; Rubio, C.; Paz, S.; Weller, D.G. The influence of the degassing phase of the Tagoro submarine volcano (Canary Islands) on the metal content of three species of cephalopods. *Mar. Pollut. Bull.* **2022**, *182*, 113964. [[CrossRef](#)] [[PubMed](#)]
35. Passos, F.D.; Fassina, P.V.; Corrêa, P.V.F.; Miranda, M.S. Separated from the cradle: A new species of *Falcidens* (Mollusca, Aplacophora, Caudofoveata) reveals weird patterns of distribution in the deep-sea. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **2022**, *186*, 103825. [[CrossRef](#)]
36. Calatayud, S.; Garcia-Risco, M.; Capdevila, M.; Canestro, C.; Palacios, O.; Albalat, R. Modular Evolution and Population Variability of *Oikopleura dioica* Metallothioneins. *Front. Cell Dev. Biol.* **2021**, *9*, 702688. [[CrossRef](#)] [[PubMed](#)]
37. Calatayud, S.; Garcia-Risco, M.; Palacios, O.; Capdevila, M.; Canestro, C.; Albalat, R. Tunicates Illuminate the Enigmatic Evolution of Chordate Metallothioneins by Gene Gains and Losses, Independent Modular Expansions, and Functional Convergences. *Mol. Biol. Evol.* **2021**, *38*, 4435–4448. [[CrossRef](#)] [[PubMed](#)]
38. Blevins, W.R.; Ruiz-Orera, J.; Messeguer, X.; Blasco-Moreno, B.; Villanueva-Canas, J.L.; Espinar, L.; Diez, J.; Carey, L.B.; Alba, M.M. Uncovering de novo gene birth in yeast using deep transcriptomics. *Nat. Commun.* **2021**, *12*, 604. [[CrossRef](#)]
39. Blindauer, C.A. Metallothioneins. In *Binding, Transport and Storage of Metal Ions in Biological Cells*; Royal Society of Chemistry: Cambridge, UK, 2014; Volume 2, pp. 606–665.
40. Isani, G.; Carpenne, E. Metallothioneins, unconventional proteins from unconventional animals: A long journey from nematodes to mammals. *Biomolecules* **2014**, *4*, 435–457. [[CrossRef](#)]
41. Ruiz-Orera, J.; Messeguer, X.; Subirana, J.A.; Alba, M.M. Long non-coding RNAs as a source of new peptides. *eLife* **2014**, *3*, e03523. [[CrossRef](#)]
42. Ruiz-Orera, J.; Villanueva-Canas, J.L.; Alba, M.M. Evolution of new proteins from translated sORFs in long non-coding RNAs. *Exp. Cell Res.* **2020**, *391*, 111940. [[CrossRef](#)]
43. Reinhardt, J.A.; Wanjiru, B.M.; Brant, A.T.; Saelao, P.; Begun, D.J.; Jones, C.D. De Novo ORFs in *Drosophila* Are Important to Organismal Fitness and Evolved Rapidly from Previously Non-coding Sequences. *PLoS Genet.* **2013**, *9*, e1003860. [[CrossRef](#)] [[PubMed](#)]
44. Ji, Z.; Song, R.; Regev, A.; Struhl, K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* **2015**, *4*, e08890. [[CrossRef](#)] [[PubMed](#)]
45. Karin, M.; Najarian, R.; Haslinger, A.; Valenzuela, P.; Welch, J.; Fogel, S. Primary structure and transcription of an amplified genetic locus: The CUP1 locus of yeast. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 337–341. [[CrossRef](#)] [[PubMed](#)]
46. Cui, Y.; Wong, W.H.; Bornberg-Bauer, E.; Chan, H.S. Recombinatoric exploration of novel folded structures: A heteropolymer-based model of protein evolutionary landscapes. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 809–814. [[CrossRef](#)] [[PubMed](#)]
47. Rorick, M.M.; Wagner, G.P. Protein structural modularity and robustness are associated with evolvability. *Genome Biol. Evol.* **2011**, *3*, 456–475. [[CrossRef](#)]
48. Niederwanger, M.; Calatayud, S.; Zerbe, O.; Atrian, S.; Albalat, R.; Capdevila, M.; Palacios, O.; Dallinger, R. *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.* **2017**, *18*, 1457. [[CrossRef](#)]
49. Pedrini-Martha, V.; Köll, S.; Dvorak, M.; Dallinger, R. 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.* **2020**, *21*, 1631. [[CrossRef](#)]

-
50. Calatayud, S.; Garcia-Risco, M.; Rojas, N.S.; Espinosa-Sanchez, L.; Artime, S.; Palacios, O.; Canestro, C.; Albalat, R. Metallothioneins of the urochordate *Oikopleura dioica* have Cys-rich tandem repeats, large size and cadmium-binding preference. *Metallomics* **2018**, *10*, 1585–1594. [[CrossRef](#)]
 51. Bongers, J.; Walton, C.D.; Richardson, D.E.; Bell, J.U. Micromolar protein concentrations and metalloprotein stoichiometries obtained by inductively coupled plasma atomic emission spectrometric determination of sulfur. *Anal. Chem.* **1988**, *60*, 2683–2686. [[CrossRef](#)]