Received: 5 September 2024 | Revised: 4 November 2024 | Accepted: 23 November 2024

DOI: 10.1002/pro.5247

RESEARCH ARTICLE

THE PROTEIN SOCIETY WILEY

Experimental recombining of repetitive motifs leads to large functional metallothioneins and demonstrates their modular evolvability potential

Reinhard Dallinger[1] | **Veronika Pedrini-Martha**[1] | **Maria Lucia Burdisso**[2,3] | **Mercè Capdevila**[4] | **Oscar Palacios**[4] | **Ricard Albalat**[2,5]

[1]Institute of Zoology and Center of Molecular Biosciences Innsbruck, University of Innsbruck, Innsbruck, Austria

[2]Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

[3]Centro de Estudios Fotosintéticos y Bioquímicos (CEFOBI-CONICET), Universidad Nacional de Rosario, Rosario, Argentina

[4]Departament de Química, Facultat de Ciències, Universitat Autònoma de Barcelona (UAB), Cerdanyola del Vallès, Spain

[5]Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona (UB), Barcelona, Spain

Correspondence
Ricard Albalat, Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Av. Diagonal 645, Barcelona 08028, Spain.
Email: ralbalat@ub.edu

Reinhard Dallinger, Institute of Zoology, University of Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria.
Email: reinhard.dallinger@uibk.ac.at

Funding information
MICIU/AEI/10.13039/501100011033, Grant/Award Number: PID2021-123258NB-I00; European Regional Development Fund/European Union; Spanish Ministerio de Ciencia e Innovación, Grant/Award Number: BIO2015-67358-C2-2-P; Austrian Science Fund, Grant/Award Number: P33973-B

Review Editor: Aitziber L. Cortajarena

Abstract

Protein modularity is acknowledged for promoting the emergence of new protein variants via domain rearrangements. Metallothioneins (MTs) offer an excellent model system for experimentally examining the consequences of domain rearrangements due to the possibility to assess the functional properties of native and artificially created variants using spectroscopic methods and metal tolerance assays. In this study, we have investigated the functional properties of AbiMT4 from the snail *Alinda biplicata* (Gastropoda, Mollusca), a large MT comprising 10 putative β domains ($\beta 3_9 \beta 1$), alongside four artificially designed variants differing in domain number, type, or order. Our findings reveal that AbiMT4 is a cadmium-selective protein with a high metal-binding capacity, characterized by structurally and functionally independent domains repeated in tandem along the protein. Our results indicate that due to its modular organization, AbiMT4 remains functional even when the number, type, and order of the domains are significantly altered. Furthermore, we demonstrate that the metal-binding properties of AbiMT4 are not dictated by the overall architecture of the protein but primarily arise from the properties of each individual domain. Using MTs as example, this work provides empirical evidence that domain rearrangements are an effective strategy for exploring new viable sequences and creating novel protein variants subject to adaptive selection. Thus, our study highlights the importance of the modular structure of proteins, as increasing their functional flexibility enhances their evolvability.

Additionally, our work demonstrates a simple way to design and model new proteins for predefined functions.

KEYWORDS

artificial variants, domain repeat proteins, evolvability, gastropoda/mollusca MTs, protein modularity, snail *Alinda biplicata*

1 | INTRODUCTION

Protein modularity refers to the fact that many proteins are composed of discrete structural and functional units called domains that can be multiplied, rearranged, combined and modified, to generate proteins with novel functions. This concept is crucial for understanding protein evolution, structure–function relationships, and engineering of proteins for diverse applications. Studying protein modularity is essential for deciphering the complexity of biological systems, driving advancements in biotechnology and medicine, and tackling challenges in protein engineering and drug discovery. Most research on protein modularity has been conducted through in silico analyses of the domain composition and arrangement in thousands of proteins sourced from proteome databases spanning phylogenetically diverse species (Apic et al. 2001; Björklund et al. 2005; Björklund et al. 2006; Coban et al. 2022; Cohen-Gihon et al. 2011; Culbertson and Levin 2023; Dohmen et al. 2020; Ekman et al. 2005; Fong et al. 2007; Forslund et al. 2019; Kersting et al. 2012; Kummerfeld and Teichmann 2005; Levitt 2009; Tordai et al. 2005; Wang and Caetano-Anolles 2009). These computational approaches have revealed that the modular architecture of the proteins enhances their evolvability by facilitating the emergence of new protein variants through domain rearrangements (Cui et al. 2002; Rorick and Wagner 2011). While the evolution of multidomain proteins composed of different domain types—i.e., heteromultimeric proteins—has been extensively analyzed (Apic et al. 2001; Björklund et al. 2005; Dohmen et al. 2020), the origin and evolution of multidomain proteins formed by tandem repeats—i.e., homomultimeric proteins—and the structural and functional principles governing domain expansions in these proteins are less understood (Björklund et al. 2006). To address this gap, we have chosen metallothioneins as a model system to experimentally investigate the impact of repeat domain rearrangements on protein function.

Metallothioneins (MTs) have traditionally been considered small proteins (<100 amino acids) that bind metal ions through the thiol groups of their cysteines (C, Cys). Cysteines are arranged in CxC, CC, and CCC motifs, and depending on their number and distribution within the peptide chain, they form one or more metal clusters. For example, vertebrate MTs form two metal clusters with divalent metal ions (M_3Cys_9 and M_4Cys_{11}) (Otvos and Armitage 1980) thanks to two functionally independent domains (Braun et al. 1986), while *Drosophila* MTs have a single domain forming a highly compact metal cluster (Egli et al. 2006; Atrian 2009). Standard MTs are therefore monodomain or bidomain proteins. In recent years, however, large MTs (>100 amino acids) made of repeated sequences have been identified in certain chordate (Calatayud et al. 2018; Calatayud et al. 2021a), mollusk (Baumann et al. 2017; Calatayud et al. 2021b; Pedrini-Martha et al. 2020), and fungal (Iturbe-Espinoza et al. 2016; Palacios et al. 2014) species. This discovery has led to the proposal that these MTs are multidomain proteins, with each repeat functioning as an autonomous protein domain. Unfortunately, experimental evidence for this modular architecture has been limited to a few examples such as *Oikopleura dioica* OdiMT2 and *Cryptococcus neoformans* CnMT1 and CnMT2 (Calatayud et al. 2021c; Espart et al. 2015).

In gastropods, as in other animal lineages, most MTs are small bidomain proteins with a N-terminal β3 domain and a C-terminal β1 domain (Beil et al. 2019; Garcia-Risco et al. 2021). It came therefore as a surprise that the marine periwinkle *Littorina littorea* and its relative, the land winkle *Pomatias elegans*, both possess Cd-selective MTs with three functional domains (Baumann et al. 2017; Schmielau et al. 2019). In the meantime, it was found that gastropoda represent the mollusk lineage with the highest number of predicted multidomain MTs (Calatayud et al. 2021b; Dallinger 2024), many of them with more than three metal-binding repeats. The MT of the terrestrial snail *Alinda biplicata*, for example, expresses two native MTs: the small 2d-AbiMT1 and the large 10d-AbiMT4 (Pedrini-Martha et al. 2020). This latter protein comprises nine β3 repeats and one C-terminal β1 domain ($\beta3_9\beta1$), but it remained unclear whether its repeat sequences correspond to a true modular structure and, if so, what impact this modular organization has on its chemical and functional properties. To address this issue, we aimed to analyze the metal-binding properties of the full-length *A. biplicata* 10d-AbiMT4 and several artificial variants thereof. We took advantage of the fact

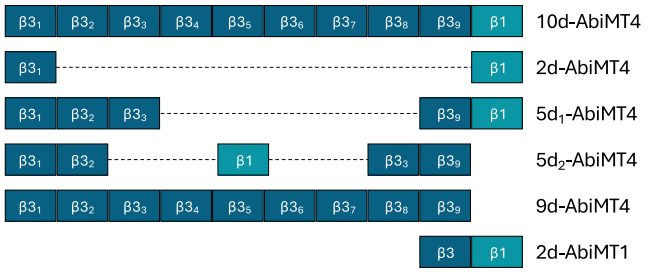

FIGURE 1 Schematic representation of AbiMT constructs. The organization of the native10d-AbiMT4 (top; accession number: MK648140.1) and 2d-AbiMT1 (bottom; accession number: MK648138.1) of *Alinda biplicata*, and of the four artificial AbiMT4 variants (2d-AbiMT4, 5d1-AbiMT4, 5d2-AbiMT4, and 9d-AbiMT4) analyzed in this work are shown. Putative β3 (dark blue) and β1 (light blue) domains are depicted.

that the functional properties of MTs can be determined experimentally using analytical chemistry methods and metal tolerance assays with recombinant proteins produced in heterologous cell systems. Chemical analysis of recombinant MTs has long been established as a reliable means of determining metal selectivity and binding capacity, offering advantages over the isolation and purification of MTs from native sources and enabling the study of artificially designed mutants, while minimizing methodological artifacts (Bofill et al. 2001; Cols et al. 1997; Orihuela et al. 2010; Palacios et al. 2011a; for a critical review, see Blindauer and Leszczyszyn 2010). Complementarily, metal tolerance assays provide a means to evaluate the functionality of any MT within a cellular environment. These assays assess the ability of cells expressing a given MT to withstand exposure to different essential and toxic metals and to varying concentrations of them (Han et al. 1992; Li et al. 2021; Liu et al. 2014; Ma et al. 2011; Ma et al. 2019; Palacios et al. 2011a; Pan et al. 1993; Thiele et al. 1986). Thus, by combining analytical chemistry methods and tolerance assays, we have characterized and compared the functional properties of 10d-AbiMT4 and several artificial constructs of it with variations in the number, type or arrangement of its putative domains, using the native 2d-AbiMT1 from *A. biplicata* as a native two-domain reference.

2 | RESULTS

2.1 | The multidomain MT of *Alinda biplicata* as a model system

We focused our study on the large 10d-AbiMT4 of *A. biplicata*, a MT of more than 300 amino acids composed of nine β3 repeats and one C-terminal β1 domain, thus organized into 10 putative β domains (10d: $\beta3_9\beta1$) (Calatayud et al. 2021b; Pedrini-Martha et al. 2020) (Figure 1). We characterized the metal-binding features of the 10d-AbiMT4 and several artificially created variants with different numbers (two, five, or nine), types (with or without β1) or orders ($5d_1$-MT4 or $5d_2$-MT4) of possible domains (Figure 1). As a reference, we also analyzed the properties of the native 2d-AbiMT1, a typical bidomain (2d: β3β1) gastropod MT, from the same species. We employed two complementary experimental methodologies: (1) chemical characterization of the metal–protein complexes of the native 10d-AbiMT4 and 2d-AbiMT1, as well as four artificial AbiMT4 variants using mass spectrometry and ICP-AES; and (2) evaluation of the ability of the native and artificial variants to enhance metal resistance in metal-tolerance assays.

2.2 | Chemical characterization of the metal–protein complexes

To investigate the formation of metal–MT complexes, we heterologously expressed the different AbiMT constructs (Figure 1) in *Escherichia coli* cultured in media supplemented with copper (Cu), cadmium (Cd), or zinc (Zn) salts. We purified the complexes and characterized them using inductively coupled plasma atomic emission spectrometry (ICP-AES) and electrospray ionization mass spectrometry (ESI-MS) analyses. The ICP-AES (Table 1) and ESI-MS (Figures 2 and S1, Supporting Information) analyses of the recovered samples showed that all the constructs exhibit metal binding abilities, not only the native MTs but also the artificially designed forms. When considering the species formed on each production, it was observed that all the constructs analyzed exhibited a Cd-thionein character, as they all rendered a single species in the Cd-productions (Figure 2, middle panels), while the Zn- and Cu-enriched productions rendered mostly a mixture of species (Figure 2, left and right panels, respectively). Interestingly, the number of Cd(II) ions detected per protein was directly related to the number of repeating tandem domains, so that each domain binds three Cd(II) ions. Thus, those with two repeats (2d-AbiMT4 and 2d-AbiMT1) rendered a single Cd_6-species, those with five domains ($5d_1$- and $5d_2$-AbiMT4) showed a Cd_{15}-species and the ones with 10 and 9 repeats formed Cd_{30}- and Cd_{27}-species, respectively.

Zn-enriched productions predominantly yielded a mixture of metallated species, with the maximum Zn(II) loading consistently correlating with the amount

TABLE 1 ICP-AES values measured for the different MT preparations.

Protein	Metal sup	Cys + Met	[MT] × 10^4 (M)	[Zn]/[MT]	[Cd]/[MT]	[Cu]/[MT]	Zn/Cu
10d-AbiMT4	Zn	91	0.75	25.1	—	—	—
	Cd		0.62	—	28.2	—	—
	Cu		0.24	10.7	—	11.6	0.92
2d-AbiMT4	Zn	19	1.36	5.7	—	—	—
	Cd		0.87	—	5.9	—	—
	Cu		1.28	2.2	—	4.9	0.44
$5d_1$-AbiMT4	Zn	46	1.01	12.7	—	—	—
	Cd		0.61	—	13.3	—	—
	Cu		0.34	7.2	—	3.9	1.84
$5d_2$-AbiMT4	Zn	46	1.42	13.2	—	—	—
	Cd		1.13	—	14.1	—	—
	Cu		0.64	5.0	—	8.6	0.58
9d-AbiMT4	Zn	82	0.84	22.1	—	—	—
	Cd		0.57	—	24.3	—	—
	Cu		0.27	9.5	—	7.5	1.26
2d-AbiMT1	Zn	19	2.00	5.0	—	—	—
	Cd		1.28	—	5.2	—	—
	Cu		0.67	3.9	—	7.5	0.52

Note: The concentration of the protein was calculated based on S concentration and taking into account the number of Cys and Met residues of each protein. The error associated with the quantification of S is lower than 2%.

of Cd(II) detected in the Cd-productions (Figure 2, left panels), indicating that no additional Zn-binding occurred in the recombinant proteins. When produced in Cu-enriched media, all the constructs rendered heteronuclear Zn, Cu-species (Figure 2, right panels), indicating a low Cu-preference for the proteins. Notably, ICP-AES analysis showed that the Zn/Cu molar ratio in the Cu-productions was highly dependent on protein size (Table 1). For almost all cases of the larger constructs, including the native 10d structure, the relative amount of Zn required to form the species in Cu-enriched medium was higher than that required for the smaller proteins. The fact that all constructs render mixtures of heterometallic species and this patent need of high amounts of Zn, already highlights the difficulties of the larger proteins to properly fold in the absence of Zn indicating that Cu(I) is not their cognate metal. Additionally, these Cu-productions provided valuable insights into the Cd-selectivity of the constructs. The formation of heteronuclear species in Cu-enriched medium suggested that the Cd-thionein character of the proteins was moderate, as extreme Cd-thioneins typically produce homonuclear Cu-species under similar conditions (Garcia-Risco et al. 2021; Garcia-Risco et al. 2023). The absence of glycosylated Cd-MT species in most of the Cd-enriched productions (a derivatization typically observed for Cu-thioneins recombinantly produced in Cd-enriched media) (Calatayud et al. 2022) is consistent with the moderate Cd-thionein character of these metallothioneins. Overall, the chemical characterization of the metal–protein complexes from different AbiMT constructs, whether encoding native or artificial proteins, revealed a consistent preference for divalent over monovalent metal ions, with a moderate but distinctive selectivity for Cd(II) ions for all cases.

2.3 | Metal tolerance in living cells

We also evaluated the ability of large MTs and their truncated variants thereof to confer metal tolerance in a bacterial cell system. In a "spot assay," *E. coli* strains transformed with the same AbiMT4-pGEX constructs used to produce the metal–protein complexes (Figure 1; see section 2.1) were exposed to different concentrations of Cd, the cognate metal of 10d-AbiMT4, as well as Zn and Cu. We also evaluated the tolerance of *E. coli* cells transformed with 2d-AbiMT1-pGEX as a reference for bacteria expressing a standard gastropod MT, and with the empty pGEX vector as a negative control. Cell cultures were serially diluted (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}) and spotted onto LB agar plates containing a range of

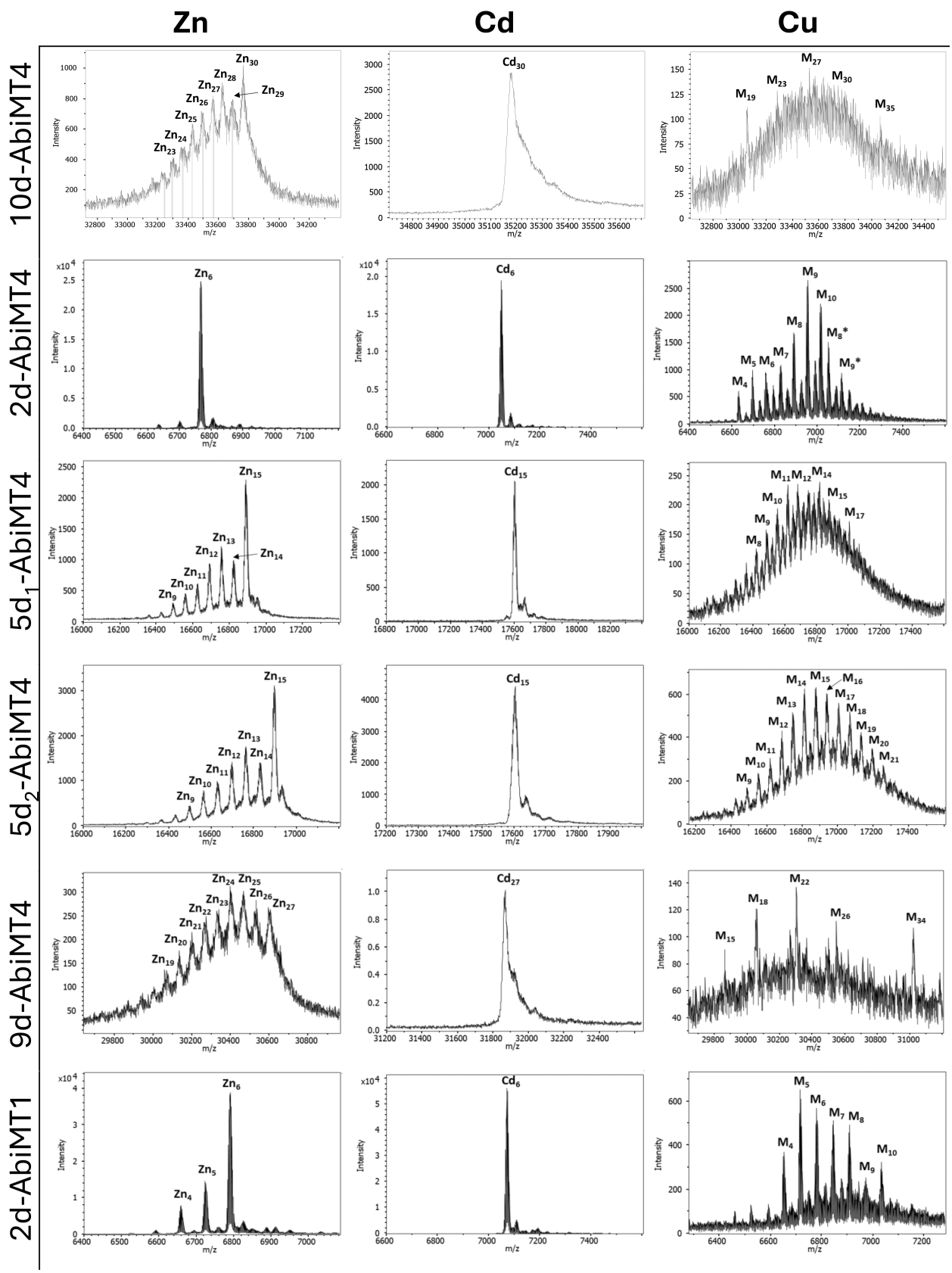

FIGURE 2 ESI-MS spectra. Deconvoluted ESI-MS spectra of the different proteins analyzed in this work, recorded at neutral pH, recombinantly produced in Zn-, Cd-, and Cu-enriched media. M stands for Cu + Zn. Glycosylated metal–MT species are depicted with an asterisk (*).

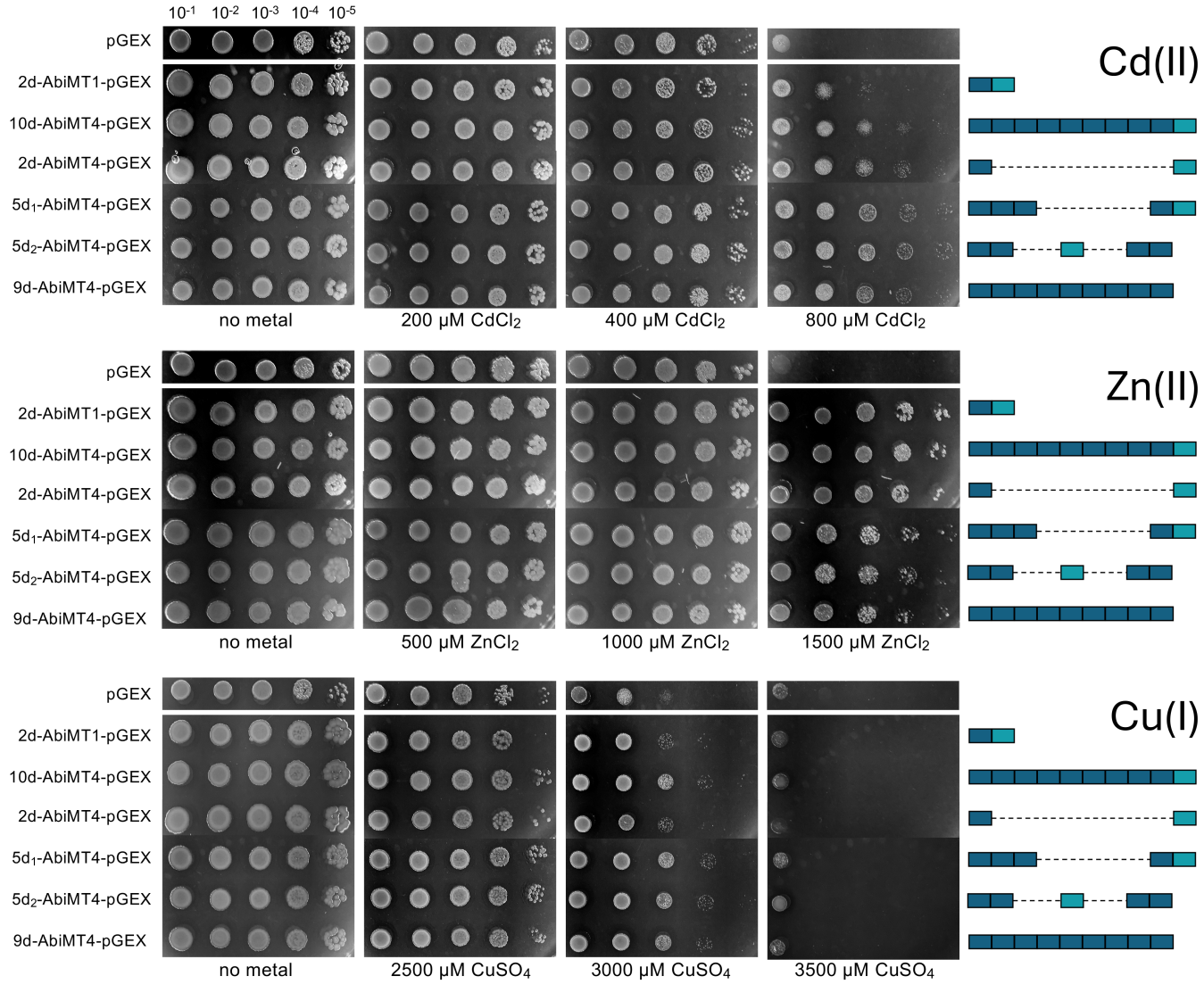

FIGURE 3 Metal tolerance assays in bacteria. Spot assays showing the effect of the heterologous expression of the native10d-AbiMT4 and 2d-AbiMT1, and four artificial AbiMT4 variants (9d-AbiMT4, $5d_1$-AbiMT4, $5d_2$-AbiMT4, and 2d-AbiMT4) in *E. coli* cell growth under Cd (top), Zn (middle), and Cu (bottom) supplementation. At the right, AbiMT constructs are schematically represented as in Figure 1. *Escherichia coli* BL21 cells were transformed with the expression vector pGEX-4T-1 (empty) or with the corresponding AbiMT-pGEX constructs. Cell cultures were serially diluted (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}) and 5 µl of each dilution were spotted onto LB-ampicillin agar plates containing a range of metal concentrations (see section 4 for details).

$CdCl_2$ (0, 200, 400, and 800 µM), $ZnCl_2$ (0, 500, 1000, and 1500 µM), and $CuSO_4$ (0, 2500, 3000, and 3500 µM) concentrations. While all cells tested grew well on LB medium without metal or at the lowest metal concentrations, only cells expressing a recombinant MT variant exhibited heightened tolerance to Cd and Zn compared to the negative control cells (Figure 3). For instance, the negative control cells showed little to no growth at the highest concentrations of Cd (800 µM) and Zn (1500 µM), whereas the cells transformed with AbiMT variants displayed clear growth (Figure 3, top and middle right panels). In contrast, there were negligible differences in Cu tolerance between cells expressing AbiMT and the negative control cells (Figure 3, bottom panels), consistent with the low preference of AbiMT for monovalent metal ions. Overall, the results demonstrated that native gastropod MTs—both the large 10d-AbiMT4 and the small 2d-AbiMT1—enhanced the tolerance of *E. coli* to divalent metal ions Cd^{2+} and Zn^{2+}, indicating their functionality when produced in a bacterial cell environment. Furthermore, the data indicated that artificially designed constructs with a variable number (2d-, 5d-, or 9d-AbiMT4), type (with or without β1), or order ($5d_1$- or $5d_2$-AbiMT4) of domains also increased the metal tolerance of the *E. coli* cells without any discernible growth differences with the wild-type 10d-AbiMT4 and 2d-AbiMT1 proteins (Figure 3). This indicated that all artificially designed MTs were functional proteins as well.

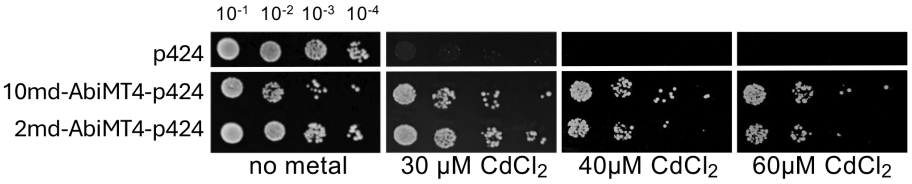

FIGURE 4 Metal tolerance assays in yeast. Effect of the heterologous expression of the native 10d-AbiMT4 and of the shortest 2d-AbiMT4 construct in *Saccharomyces cerevisiae* cell growth under Cd supplementation. Yeast 51.2cΔc5 cells (an MT-KO null strain) were transformed with the vector p424 (empty) or with the corresponding AbiMT4-p424 constructs. Transformed 51.2cΔc5 cells were grown, serially diluted (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}), and 5 µl of each dilution were spotted onto SC agar plates containing a range of $CdCl_2$ concentrations (see section 4 for details).

In order to validate our metal tolerance assays in *E. coli* expressing recombinant MTs, and to ensure the reliability of our results without any potential distortions caused by the cell system or the fusion of MTs with GST, we conducted "spot assays" using a different cell system (*Saccharomyces cerevisiae*) and producing the MTs alone, without fusing them with any additional sequences. In addition, to increase the sensibility of the assays, we used the 51.2cΔc5 knockout strain, a yeast strain lacking its endogenous MTs (yeast Cup1 and Crs5), which exhibits a "low metal tolerance" phenotype. Yeast MT-KO cells were transformed with two AbiMT4-p424 constructs: one expressing the full-length 10d-AbiMT4, and the other expressing the smallest artificial variant, 2d-AbiMT4 (Figure 1). Their growth was assessed to determine the functional complementation of the gastropod MTs in MT-KO yeast cells exposed to Cd. As control, we evaluated the Cd tolerance of 51.2cΔc5 yeast cells transformed with the empty p424 vector. Cell cultures were serially 10-fold diluted and spotted onto SC plates containing a range (0, 30, 40, and 60 µM) of $CdCl_2$ concentrations. Plates were incubated for 3 days at 30°C. MT-KO cells transformed with the empty p424 vector did not grow at any of the tested Cd concentrations, whereas cells expressing both AbiMT4s (the full-length or the smallest variant) grew well, even at the highest concentration of 60 µM (Figure 4). These experiments indicated that both the large and small AbiMT4 variants were functionally competent proteins capable of enhancing the metal tolerance of cells expressing them, whether they were *S. cerevisiae* or *E. coli* cells.

3 | DISCUSSION

3.1 | Functional properties of multidomain MTs

As previously mentioned, standard gastropod MTs are bidomain β3β1 proteins, where each domain folds into a well-defined three-divalent metal cluster (Beil et al. 2019). In accordance with this stoichiometry, the 2d-AbiMT1 (β3β1) used as a reference in this study bound six divalent metal ions, with a moderate selectivity for Cd ions (Figure 2). The same stoichiometry and metal selectivity were observed for the large 10d-AbiMT4 ($\beta3_9\beta1$), which bound 30 divalent metal (Cd) ions, as well as for the artificial variants with different number of β3 repeats: 27 ions for the 9d-AbiMT4, 15 ions for the $5d_1$- and $5d_2$-AbiMT4, and 6 ions for the 2d-AbiMT4 (Figure 2). These data demonstrate that the metal binding capacity of 10d-AbiMT4 and of the different variants is directly proportional to the number of repeats, as expected if each repeat functions as an autonomous domain folding into a Cd_3Cys_9 cluster. Mass-spectroscopic data further revealed that AbiMT4 is a Cd-selective MT, suggesting its involvement in the physiological handling of cadmium. Supporting this function, adult snails exposed to Cd in the laboratory showed transcriptional upregulation of the *AbiMT4* gene (Pedrini-Martha et al. 2020). This substantiates the idea that when an MT has never been isolated and no relevant functional data is available, "metal selectivity"—the preference of an MT to bind certain metal ions in in vitro experiments (Dallinger 2024)—can be a good proxy for its metal specificity, that is, its biological function in front of a certain metal in vivo (Bofill et al. 2009; Dallinger 2024; Palacios et al. 2011b). It is also noteworthy that the 10d-AbiMT4 and the different variants exhibited the same metal selectivity for Cd ions, suggesting that the metal preference of multidomain MTs likely arises from each individual domain and, at least for AbiMT4, is largely independent of the number, type, or position of domains within the protein.

In the metal tolerance assays, no noticeable growth differences were observed between the 10d-AbiMT4 and the four artificial MT4 constructs (Figure 3). These results suggest that the tolerance of the bacteria to metals depended primarily on the intracellular production of the MT domains, regardless of whether these domains were part of a large (10d- and 9d-AbiMT4), intermediate

(5d-AbiMT4), or small (2d-AbiMT4 and 2d-AbiMT1) protein, or the type (with or without β1) and order ($5d_1$- and $5d_2$-AbiMT4) of the domains in the recombinant MT. While we cannot completely rule out subtle differences between the MT constructs, our results align with findings that the metal-binding capacity of *E. coli* cells expressing MTs with different numbers of tandem repeats of human MT1 or Crab ShMT was similar and depended little on the number of repetitions (Ma et al. 2011; Ma et al. 2019). In summary, our results demonstrate that 10d-AbiMT4 possesses a modular architecture in which the tandemly repeated domains function independently, both structurally and functionally. Spectrometric and metal tolerance results consistently indicated that the metal-binding properties of AbiMT4 are not determined by the overall architecture of the whole protein, but primarily by the sum of the features of each individual domain.

3.2 | Evolutionary implications of the modular architecture of MTs

The modular architecture of MTs has significant evolutionary implications. In a previous work, we argued that the modularity of MTs contributes to the evolvability of small MTs by facilitating the emergence of new forms through domain swaps (Calatayud et al. 2022). Our current data extend this concept to large multidomain MTs, demonstrating that modularity provides the flexibility and versatility necessary for these proteins to remain functional despite changes in the number, type, or order of their domains. In this regard, it should be noted that a second large MT sequence has been identified in *A. biplicata* (Calatayud et al. 2021b; Pedrini-Martha et al. 2020). This isoform, named 9d-AbiMT3, consists of eight β3 repeats and one β1 domain ($\beta3_8\beta1$). Comparison of AbiMT3 and AbiMT4 sequences revealed that their β3 (and β1) domains are identical or nearly identical, with the primary difference being the absence of the sixth β3 domain (Pedrini-Martha et al. 2020). While it is possible that AbiMT3 and AbiMT4 are encoded by different genes, it is more plausible that AbiMT3 and AbiMT4 are allelic variants of the same gene, differing only in the number of repeats, which our results indicate does not significantly impact MT function. This possibility is supported by the identification of allelic variants with different repeat domain numbers in the multidomain OdiMT2 of *O. dioica* (Calatayud et al. 2021c).

In conclusion, our results suggest that the metal-binding properties of multidomain MTs arise from the combined contributions of each individual domain. Their modular organization endows these proteins with structural and functional plasticity, facilitating the gain, loss, or rearrangement of domains. This enhances their evolvability, enabling the creation of variants with new functional features in metal handling that contribute to the adaptation and evolution of organisms. However, several important biological questions remain unresolved. For example, while large multidomain MTs have independently arisen in mollusks and chordates, the reasons for their presence being restricted to only certain species are unclear. Additionally, the prevalence of multidomain MTs in other metazoan phyla and the selective advantages they confer on the species that possess them are not well understood. Further research on multidomain MTs is essential to address these questions comprehensively.

4 | MATERIALS AND METHODS

4.1 | Production and purification of recombinant metal–MT complexes

Synthetic cDNAs encoding proteins with different numbers and combinations of AbiMT domains (Figure 1) were provided by Synbio Technologies (Monmouth Junction, NJ). They were cloned into the *BamH*I/*Xho*I sites of the pGEX-4T-1 expression vector (GE Healthcare, Chicago, IL) and transformed in the protease-deficient *E. coli* BL21 strain. For heterologous protein production, 500 ml of LB medium containing 100 μg ml^{-1} ampicillin was inoculated with *E. coli* BL21 cells transformed with the recombinant plasmids. After growth overnight at 37°C/250 rpm, the cultures were used to inoculate 5 L of fresh LB medium with 100 μg ml^{-1} ampicillin. Gene expression was induced with 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG) for 3 h at 37°C/250 rpm. After the first 30 min of induction, cultures were supplemented with $ZnCl_2$ (300 μM), $CdCl_2$ (300 μM), or $CuSO_4$ (500 μM). 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 (140 mM NaCl, 2.7 mM KCI, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 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 s, and then centrifuged for 40 min at 17,200*g* (12,000 rpm) at 4°C. 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 PBS bubbled with argon to prevent oxidation. After three washes, GST–MT fusion proteins were digested with thrombin (Serva, Heidelberg), 20 U L^{-1} of culture overnight at 17°C, thus enabling separation of the

metal–MT complexes from the GST that remained bound to the sepharose matrix. Note that the digestion with thrombin added two additional residues, Gly and Ser to the N-terminal end of all purified proteins. These two amino acids do not interfere with the metal-binding features of the recombinant MTs (Cols et al. 1997). The eluted metal–MT complexes were concentrated on a 3 kDa Amicon Ultra-2 concentrator (Amicon, Merck), and fractionated on a Superdex-75 FPLC column (GE Healthcare) equilibrated with 20 mM Tris–HCl pH 7.0, or 50 mM ammonium acetate 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 the purified metal–MT preparations were characterized by means of Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES), using an Optima 4300DV instrument (Perkin-Elmer, Waltham, MA) by measuring S at 182.04 nm, Zn at 213.85 nm, Cd at 228.80 nm, and Cu at 324.75 nm. The protein concentration was determined assuming that the total content of S in the proteins is derived from both cysteine and methionine amino acids.

The molecular weight of the Zn-, Cd-, and Cu-MT species formed, as well as that of the corresponding apo proteins, was determined by Electrospray Ionization Time-of-Flight Mass Spectrometry (ESI-TOF MS) on a Micro TOF-Q instrument (Bruker Daltonics, Bremen) connected to a Series 1200 HPLC Agilent pump and controlled by the Compass Software. Samples were analyzed under neutral (pH 7.0) and acidic (pH 2.4) conditions, using as running buffer a 5:95 mixture of acetonitrile:ammonium acetate (15 mM) and a 5:95 mixture of acetonitrile:formic acid solution, respectively. Under acidic conditions, Zn(II) and Cd(II) are released while Cu(I) is kept in the protein complex. Therefore, the characterization of the apo proteins was performed at acidic conditions on the Zn- or Cd-loaded forms. For each analysis, 20 µl of protein solution (with concentrations ranging from 20 to 129 µM) were injected at 30–50 µl min^{-1} and analyzed under the following conditions: capillary counter-electrode voltage, 3500–5000 V; dry temperature, 90–110°C; dry gas, 6 L min^{-1}; m/z range 800–3000.

4.3 | Metal tolerance assays in bacteria

The same constructions used for the production of recombinant metal–MT complexes (Figure 1) were used for metal tolerance assays in bacteria. We also used an empty pGEX vector as a negative control. *Escherichia coli* BL21 cells transformed with the plasmids were grown in LB medium with 100 µg ml^{-1} ampicillin overnight at 37°C/250 rpm. Saturated cultures were diluted 7-fold in 4 ml of fresh medium with 100 µg ml^{-1} ampicillin and incubated at 37°C/250 rpm to an OD_{600} of 0.5–0.6 (≈45 min). MT expression was induced with 100 µM IPTG for 3 h at 37°C/250 rpm. Cell cultures were serially diluted (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}) and 5 µl of each dilution were spotted onto LB agar plates containing a range of $CdCl_2$ (0, 200, 400, and 800 µM), $ZnCl_2$ (0, 500, 1000, and 1500 µM), and $CuSO_4$ (0, 2500, 300, and 3500 µM) concentrations. Plates were incubated overnight at 37°C and then photographed.

4.4 | Cloning AbiMTs in yeast expression vector and metal tolerance assays in yeast MT-knockout cells

Synthetic cDNAs encoding the full-length 10d-AbiMT4 and the 2d-AbiMT4 proteins (Figure 1) were provided by Synbio Technologies (Monmouth Junction, NJ) and cloned into the *Spe*I/*Eco*RI sites of the yeast p424-GDP vector (ATCC, Manassas, VA), containing TRP1 as a selection marker, the constitutive glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter for gene expression, and the cytochrome-c-oxidase (CYC1) transcriptional terminator (Mumberg et al. 1995). A yeast Kozak sequence (TATAAAATG) (Li et al. 2017) was included in both cDNAs to improve translation efficiency. Recombinant plasmids were transformed using the LiAc/SS-DNA/PEG procedure (Stearns et al. 1990) into the *S. cerevisiae* 51.2cΔc5 strain (*MAT*a, *trp*1-1, *ura*3-52, *ade*-, *his*-, *CAN*^{R}, *gal*1, *leu*2-3, *112 met13*, *cup1Δ::URA3 crs5Δ::LEU2*) derived from VC-sp6 (Culotta et al. 1994) and characterized by a total MT deficiency. Transformed cells were selected by their capacity to grow in complete synthetic medium (SC) lacking Trp (SC–Trp). The empty p424 vector was used as control. For cadmium tolerance assays, cultures of 51.2cΔc5 cells transformed with the different constructs were grown in SC–Trp liquid medium at 30°C/250 rpm to an OD_{600} of 0.6. Four 10-fold dilutions were performed, and 5 µl of each dilution were spotted onto SC plates supplemented with cadmium at final concentrations of 0, 30, 40, and 60 µM. Plates were incubated for 3 days at 30°C and then photographed.

AUTHOR CONTRIBUTIONS

Reinhard Dallinger: Conceptualization; project administration; funding acquisition; writing – review and

editing; formal analysis. **Veronika Pedrini-Martha:** Conceptualization; methodology; investigation; funding acquisition; writing – review and editing. **Maria Lucia Burdisso:** Methodology; investigation; writing – review and editing. **Mercè Capdevila:** Conceptualization; formal analysis; funding acquisition; project administration; writing – review and editing. **Oscar Palacios:** Conceptualization; formal analysis; investigation; funding acquisition; project administration; writing – review and editing; visualization. **Ricard Albalat:** Conceptualization; methodology; formal analysis; investigation; funding acquisition; project administration; writing – original draft; writing – review and editing; visualization.

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. R.A. was supported by Grant PID2021-123258NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU. M.C. and Ò.P. were supported by BIO2015-67358-C2-2-P from the Spanish Ministerio de Ciencia e Innovación. R.A. was member of the "Grup de Recerca de la Generalitat de Catalunya" ref. 2021 SGR 00372, and M.C. and Ò.P. of the ref. 2021SGR00668. Significant parts of this study were co-financed by a cooperation grant to R.D. and V.P-M. from the Austrian Science Fund (FWF project P33973-B; DOI: https://doi.org/10.55776/P33973).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data underlying this article are available in NCBI database: https://www.ncbi.nlm.nih.gov/.

ORCID

Ricard Albalat https://orcid.org/0000-0003-0282-9595

REFERENCES

Apic G, Gough J, Teichmann SA. Domain combinations in archaeal, eubacterial and eukaryotic proteomes. J Mol Biol. 2001;310:311–25.

Atrian S. Metallothioneins in Diptera. In: Sigel H, Sigel A, Sigel RKO, editors. Metallothioneins and related chelators. Cambridge: De Gruyter; 2009. p. 155–82.

Baumann C, Beil A, Jurt S, Niederwanger M, Palacios O, Capdevila M, et al. 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. 2017;56:4617–22.

Beil A, Jurt S, Walser R, Schonhut T, Guntert P, Palacios O, et al. The solution structure and dynamics of Cd-metallothionein from *Helix pomatia* reveal optimization for binding Cd over Zn. Biochemistry. 2019;58:4570–81.

Björklund ÅK, Ekman D, Elofsson A. Expansion of protein domain repeats. PLoS Comput Biol. 2006;2:e114.

Björklund ÅK, Ekman D, Light S, Frey-Skött J, Elofsson A. Domain rearrangements in protein evolution. J Mol Biol. 2005;353:911–23.

Blindauer CA, Leszczyszyn OI. Metallothioneins: unparalleled diversity in structures and functions for metal ion homeostasis and more. Nat Prod Rep. 2010;27:720–41.

Bofill R, Capdevila M, Atrian S. Independent metal-binding features of recombinant metallothioneins convergently draw a step gradation between Zn- and Cu-thioneins. Metallomics. 2009;1:229–34.

Bofill R, Capdevila M, Cols N, Atrian S, Gonzalez-Duarte P. Zinc (II) is required for the in vivo and in vitro folding of mouse copper metallothionein in two domains. J Biol Inorg Chem. 2001;6:405–17.

Braun W, Wagner G, Worgotter E, Vasak M, Kagi JH, Wuthrich K. Polypeptide fold in the two metal clusters of metallothionein-2 by nuclear magnetic resonance in solution. J Mol Biol. 1986;187:125–9.

Calatayud S, Garcia-Risco M, Capdevila M, Cañestro C, Palacios O, Albalat R. Modular evolution and population variability of *Oikopleura dioica* metallothioneins. Front Cell Dev Biol. 2021c;9:702688.

Calatayud S, Garcia-Risco M, Palacios O, Capdevila M, Cañestro 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. 2021a;38:4435–48.

Calatayud S, Garcia-Risco M, Pedrini-Martha V, Eernisse DJ, Dallinger R, Palacios O, et al. Modularity in protein evolution: modular organization and de novo domain evolution in mollusk metallothioneins. Mol Biol Evol. 2021b;38:424–36.

Calatayud S, Garcia-Risco M, Pedrini-Martha V, Niederwanger M, Dallinger R, Palacios O, et al. The modular architecture of metallothioneins facilitates domain rearrangements and contributes to their evolvability in metal-accumulating mollusks. Int J Mol Sci. 2022;23:15824.

Calatayud S, Garcia-Risco M, Rojas NS, Espinosa-Sanchez L, Artime S, Palacios O, et al. Metallothioneins of the urochordate *Oikopleura dioica* have Cys-rich tandem repeats, large size and cadmium-binding preference. Metallomics. 2018;10:1585–94.

Coban A, Bornberg-Bauer E, Kemena C. Domain evolution of vertebrate blood coagulation Cascade proteins. J Mol Evol. 2022;90:418–28.

Cohen-Gihon I, Fong JH, Sharan R, Nussinov R, Przytycka TM, Panchenko AR. Evolution of domain promiscuity in eukaryotic genomes—a perspective from the inferred ancestral domain architectures. Mol Biosyst. 2011;7:784–92.

Cols N, Romero-Isart N, Capdevila M, Oliva B, Gonzalez-Duarte P, Gonzalez-Duarte R, et al. 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;68:157–66.

Cui Y, Wong WH, Bornberg-Bauer E, Chan HS. Recombinatoric exploration of novel folded structures: a heteropolymer-based model of protein evolutionary landscapes. Proc Natl Acad Sci U S A. 2002;99:809–14.

Culbertson EM, Levin TC. Eukaryotic CD-NTase, STING, and viperin proteins evolved via domain shuffling, horizontal transfer, and ancient inheritance from prokaryotes. PLoS Biol. 2023; 21:e3002436.

Culotta VC, Howard WR, Liu XF. CRS5 encodes a metallothionein-like protein in *Saccharomyces cerevisiae*. J Biol Chem. 1994;269: 25295–302.

Dallinger R. Metals and metallothionein evolution in snails: a contribution to the concept of metal-specific functionality from an animal model group. Biometals. 2024;37:671–96.

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.

Egli D, Domènech J, Selvaraj A, Balamurugan K, Hua H, Capdevila M, et al. The four members of the Drosophila metallothionein family exhibit distinct yet overlapping roles in heavy metal homeostasis and detoxification. Genes Cells. 2006;11: 647–58.

Ekman D, Björklund ÅK, Frey-Skött J, Elofsson A. Multi-domain proteins in the three kingdoms of life: orphan domains and other unassigned regions. J Mol Biol. 2005;348:231–43.

Espart A, Gil-Moreno S, Palacios O, Capdevila M, Atrian S. Understanding the 7-Cys module amplification of *C. neoformans* metallothioneins: how high capacity Cu-binding polypeptides are built to neutralize host nutritional immunity. Mol Microbiol. 2015;98:977–92.

Fong JH, Geer LY, Panchenko AR, Bryant SH. Modeling the evolution of protein domain architectures using maximum parsimony. J Mol Biol. 2007;366:307–15.

Forslund SK, Kaduk M, Sonnhammer ELL. Evolution of protein domain architectures. Methods Mol Biol. 2019;1910:469–504.

Garcia-Risco M, Calatayud S, Pedrini-Martha V, Albalat R, Dallinger R, Palacios O, et al. 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.

Garcia-Risco M, Calatayud S, Pedrini-Martha V, Albalat R, Palacios O, Capdevila M, et al. A de novo evolved domain improves the cadmium detoxification capacity of limpet metallothioneins. Sci Rep. 2023;13:8895.

Han NS, Seo J-H, Chung YC. Growth and copper resistance of recombinant *Saccharomyces cerevisiae* containing a metallothionein gene. Biotechnol Lett. 1992;14:7–10.

Iturbe-Espinoza P, Gil-Moreno S, Lin W, Calatayud S, Palacios O, Capdevila M, et al. The fungus *Tremella mesenterica* encodes the longest metallothionein currently known: gene, protein and metal binding characterization. PLoS One. 2016;11:e0148651.

Kersting AR, Bornberg-Bauer E, Moore AD, Grath S. Dynamics and adaptive benefits of protein domain emergence and arrangements during plant genome evolution. Genome Biol Evol. 2012; 4:316–29.

Kummerfeld SK, Teichmann SA. Relative rates of gene fusion and fission in multi-domain proteins. Trends Genet. 2005;21:25–30.

Levitt M. Nature of the protein universe. Proc Natl Acad Sci U S A. 2009;106:11079–84.

Li J, Liang Q, Song W, Marchisio MA. Nucleotides upstream of the Kozak sequence strongly influence gene expression in the yeast *S. cerevisiae*. J Biol Eng. 2017;11:25.

Li X, Ren Z, Crabbe MJC, Wang L, Ma W. Genetic modifications of metallothionein enhance the tolerance and bioaccumulation of heavy metals in *Escherichia coli*. Ecotoxicol Environ Saf. 2021;222:112512.

Liu Y, Wu H, Kou L, Liu X, Zhang J, Guo Y, et al. Two metallothionein genes in *Oxya chinensis*: molecular characteristics, expression patterns and roles in heavy metal stress. PLoS One. 2014;9:e112759.

Ma W, Li X, Wang Q, Ren Z, Crabbe MJC, Wang L. Tandem oligomeric expression of metallothionein enhance heavy metal tolerance and bioaccumulation in *Escherichia coli*. Ecotoxicol Environ Saf. 2019;181:301–7.

Ma Y, Lin J, Zhang C, Ren Y, Lin J. Cd(II) and As(III) bioaccumulation by recombinant *Escherichia coli* expressing oligomeric human metallothioneins. J Hazard Mater. 2011;185:1605–8.

Mumberg D, Muller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene. 1995;156:119–22.

Orihuela R, Monteiro F, Pagani A, Capdevila M, Atrian S. Evidence of native metal-S(2-)-metallothionein complexes confirmed by the analysis of Cup1 divalent-metal-ion binding properties. Chemistry. 2010;16:12363–72.

Otvos JD, Armitage IM. Structure of the metal clusters in rabbit liver metallothionein. Proc Natl Acad Sci U S A. 1980;77:7094–8.

Palacios O, Atrian S, Capdevila M. Zn- and Cu-thioneins: a functional classification for metallothioneins? J Biol Inorg Chem. 2011b;16:991–1009.

Palacios O, Espart A, Espin J, Ding C, Thiele DJ, Atrian S, et al. Full characterization of the Cu-, Zn-, and Cd-binding properties of CnMT1 and CnMT2, two metallothioneins of the pathogenic fungus *Cryptococcus neoformans* acting as virulence factors. Metallomics. 2014;6:279–91.

Palacios O, Pagani A, Pérez-Rafael S, Egg M, Höckner M, Brandstätter A, et al. Shaping mechanisms of metal specificity in a family of metazoan metallothioneins: evolutionary differentiation of mollusc metallothioneins. BMC Biol. 2011a;9:4.

Pan A, Tie F, Yang M, Luo J, Wang Z, Ding X, et al. Construction of multiple copy of α-domain gene fragment of human liver metallothionein IA in tandem arrays and its expression in transgenic tobacco plants. Protein Eng Design Sel. 1993;6:755–62.

Pedrini-Martha V, Koll S, Dvorak M, Dallinger R. Cadmium uptake, MT gene activation and structure of large-sized multidomain metallothioneins in the terrestrial door snail *Alinda biplicata* (Gastropoda, Clausiliidae). Int J Mol Sci. 2020;21:1631.

Rorick MM, Wagner GP. Protein structural modularity and robustness are associated with evolvability. Genome Biol Evol. 2011;3: 456–75.

Schmielau L, Dvorak M, Niederwanger M, Dobieszewski N, Pedrini-Martha V, Ladurner P, et al. 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. 2019;648:561–71.

Stearns T, Ma H, Botstein D. Manipulating yeast genome using plasmid vectors. Methods Enzymol. 1990;185:280–97.

Thiele DJ, Walling MJ, Hamer DH. Mammalian metallothionein is functional in yeast. Science. 1986;231:854–6.

Tordai H, Nagy A, Farkas K, Banyai L, Patthy L. Modules, multidomain proteins and organismic complexity. FEBS J. 2005;272:5064–78.

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.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dallinger R, Pedrini-Martha V, Burdisso ML, Capdevila M, Palacios O, Albalat R. Experimental recombining of repetitive motifs leads to large functional metallothioneins and demonstrates their modular evolvability potential. Protein Science. 2025;34(1): e5247. https://doi.org/10.1002/pro.5247