J Biol Inorg Chem (2011) 16:683–694
DOI 10.1007/s00775-011-0770-2

ORIGINAL PAPER

# Protein and metal cluster structure of the wheat metallothionein domain $\gamma$-$E_c$-1: the second part of the puzzle

**Jens Loebus · Estevão A. Peroza · Nancy Blüthgen · Thomas Fox · Wolfram Meyer-Klaucke · Oliver Zerbe · Eva Freisinger**

Received: 15 November 2010 / Accepted: 27 February 2011 / Published online: 25 March 2011
© SBIC 2011

**Abstract** Metallothioneins (MTs) are small cysteine-rich proteins coordinating various transition metal ions, including $Zn^{II}$, $Cd^{II}$, and $Cu^{I}$. MTs are ubiquitously present in all phyla, indicating a successful molecular concept for metal ion binding in all organisms. The plant MT $E_c$-1 from *Triticum aestivum*, common bread wheat, is a $Zn^{II}$-binding protein that comprises two domains and binds up to six metal ions. The structure of the C-terminal four metal ion binding $\beta_E$ domain was recently described. Here we present the structure of the N-terminal second domain, $\gamma$-$E_c$-1, determined by NMR spectroscopy. The $\gamma$-$E_c$-1 domain enfolds an $M^{II}_2Cys_6$ cluster and was characterized as part of the full-length $Zn_6E_c$-1 protein as well as in the form of the separately expressed domain, both in the $Zn^{II}$-containing isoform and the $Cd^{II}$-containing isoform. Extended X-ray absorption fine structure analysis of $Zn_2\gamma$-$E_c$-1 clearly shows the presence of a $ZnS_4$ coordination sphere with average Zn–S distances of 2.33 Å. $^{113}$Cd NMR experiments were used to identify the $M^{II}$-Cys connectivity pattern, and revealed two putative metal cluster conformations. In addition, the general metal ion coordination abilities of $\gamma$-$E_c$-1 were probed with $Cd^{II}$ binding experiments as well as by pH titrations of the $Zn^{II}$ and $Cd^{II}$ forms, the latter suggesting an interaction of the $\gamma$ domain and the $\beta_E$ domain within the full-length protein.

**Keywords** Plant metallothionein · Metal–thiolate cluster · Electronic absorption spectroscopy · Extended X-ray absorption fine structure · NMR spectroscopy

J. Loebus and E. A. Peroza contributed equally.

An interactive 3D complement page in Proteopedia is available at http://proteopedia.org/wiki/index.php/Journal:JBIC:9.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00775-011-0770-2) contains supplementary material, which is available to authorized users.

J. Loebus · E. A. Peroza · N. Blüthgen · T. Fox · E. Freisinger (✉)
Institute of Inorganic Chemistry,
University of Zurich, 8057 Zurich, Switzerland
e-mail: freisinger@aci.uzh.ch

W. Meyer-Klaucke
European Molecular Biology Laboratory (EMBL),
Outstation Hamburg at Deutsches Elektronen-Synchrotron (DESY), 22603 Hamburg, Germany

O. Zerbe (✉)
Institute of Organic Chemistry,
University of Zurich, 8057 Zurich, Switzerland
e-mail: zerbe@oci.uzh.ch

## Introduction

Metallothioneins (MTs) are low molecular mass (2–10 kDa) and Cys-rich proteins with a preference for the coordination of metal ions with $d^{10}$ electron configuration, e.g., $Zn^{II}$, $Cu^{I}$, and $Cd^{II}$ [1]. Their occurrence is reported throughout the animal kingdom, in plants, in several eukaryotic microorganisms, as well as in some prokaryotes [2]. The plant MT $E_c$-1 from wheat consists of 81 amino acids. All of the 17 Cys and two His residues are involved in the coordination of six divalent metal ions that are arranged in two separate metal-binding domains (Fig. 1) [3].

$E_c$-1 is the first and so far only plant MT which has been successfully isolated from plant material [4]. $E_c$-1 is most abundant in wheat embryos and is present in the $Zn^{II}$ form [5]. Unlike most MTs, $E_c$-1 also recruits two His residues for $Zn^{II}$ binding, a feature so far only observed in a cyanobacterial MT form [6]. This ligand specificity usually

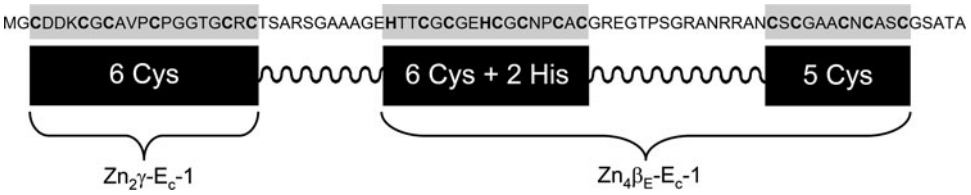


**Fig. 1** Amino acid sequence of full-length wheat $E_c$-1 with the Cys-rich metal ion coordinating regions highlighted in *gray* (*top*) as well as a schematic representation giving the sort and number of coordinating amino acids. The N-terminal Cys-rich region harbors the $Zn_2Cys_6$ cluster described herein, and the central and C-terminal regions together form the $Zn_4\beta_E$ domain

distinguishes MTs from $Zn^{II}$-binding enzymes, where the $Zn^{II}$ coordination sphere often consists of a mixture of sulfur-, nitrogen-, and oxygen-donor ligands. In consequence, the resulting metal clusters show a limited variety of possible structures. This is demonstrated by the fact that only two basic cluster arrangements have been structurally described for divalent metal ions so far: The $M^{II}_3Cys_9$ cluster of the $\beta$ domain of vertebrate, crustacean, and echinoderm MTs and the $M^{II}_4Cys_{11}$ cluster of the α domain from vertebrate and echinoderm MTs (e.g., Protein Data Bank codes 4MT2, 1DMC, 1DME, 1QJK, and 1QJL [7–9]). The $Zn_4Cys_9His_2$ cluster of the cyanobacterium *Synechococcus* PCC 7942 is structurally related to the $Zn_4Cys_{11}$ cluster with exchange of two terminal thiolate ligands for imidazole moieties [6]. A $Zn_3Cys_9$ cluster with similarity to the $M^{II}_3Cys_9$ cluster mentioned above can be also found in the $\beta_E$ domain of wheat $E_c$-1, but the $Zn^{II}$-coordinating residues are interleaved with the two Cys and His ligands of the additional mononuclear $Zn^{II}$ site [10]. This $ZnCys_2His_2$ site, although known from certain zinc-finger proteins, was unprecedented in the MT superfamily and is so far uniquely found in the plant $E_c$ proteins. The limited structural variability of the metal ion binding sites in MTs is consistent with the constricted biological functionality known so far, mainly the participation in metal ion homeostasis and detoxification as well as protection against oxidative stress [11–13]. Wheat $E_c$-1 seems to have a potential role in plant development as inferred from its high abundance during embryogenesis [8]. Interestingly, the only regulatory element found so far in the upstream 5′ flanking region of the wheat $E_c$-1 messenger RNA (mRNA) is an abscisic acid responsive element analog [14]. Even more striking than the putative regulation by one of the major plant hormones involved in abscission, grain filling, desiccation, and embryogenesis [15] is the absence of a metal responsive regulatory element [14]. In addition, it is known that 25% of an entire $^{35}S$-labeled Cys pool is found in $E_c$-1 when wheat grain embryo mRNA is expressed in a cell-free expression system [4]. This indicates either a high rate of mRNA translation or a massive accumulation of $E_c$-1 mRNA and/or a failure of its degradation in the dried embryo, possibly regulated at the transcriptional level. From further studies, using 8K complementary DNA microarray technology, the $E_c$-1 transcript abundance was shown to correlate with grain dry weight only [16]. This behavior is found for only three other proteins, the defense proteins $\gamma$-purothionin and remorin-like protein and asparagine synthetase 2, a "housekeeping" enzyme. In that study, it was further shown that gene expression during wheat grain development can be divided into ten clusters. The $E_c$-1 transcript belongs to the tenth cluster, which comprises only ten of the 2,295 differentially expressed genes. The $E_c$-1 transcript level peaks at 35 days after anthesis, hence in the maturation and desiccation state, and disappears abruptly during the first hour of imbibition [14, 16]. Generally, such rapid decline of transcripts is indicative of redundant mRNA remaining after anthesis. This view is supported by the concomitant decrease of the $E_c$-1 protein level. Yet, despite the wealth of information, the role of $E_c$-1 still remains elusive.

In the following, we present the first 3D structure of the N-terminal, 24 amino acid residues comprising $\gamma$ domain of wheat $E_c$-1 as determined by NMR spectroscopy. In total, three structures of different $\gamma$-$E_c$-1 forms are presented: the $Cd^{II}$ and $Zn^{II}$ forms of the separately expressed domain and the $Zn_2\gamma$-$E_c$-1 form as part of the full-length protein. All forms contain a $M^{II}_2Cys_6$ cluster, which is unprecedented for any MT. Extended X-ray absorption fine structure (EXAFS) studies confirm such an arrangement. Moreover, the metal ion binding properties of $\gamma$-$E_c$-1 are probed via pH titration and metal ion reconstitution experiments.

## Materials and methods

### Chemicals and solutions

$^{113}CdCl_2$ and $^{15}NH_4Cl$ were purchased from Cambridge Isotope Laboratories (Innerberg, Switzerland), $d_{11}$-tris(hydroxymethyl)aminomethane ($d_{11}$-Tris) was purchased from Euriso-top (Saint-Aubin, France), the enzymes used for plasmid construction and protein cleavage were purchased from Promega (Catalys, Wallisellen, Switzerland), Roche (Rotkreuz, Switzerland), GE Healthcare Europe (Glattbruck, Switzerland), or New England Biolabs

(Ipswich, MA, USA), Luria–Bertani broth (Miller) was purchased from Chemie Brunschwig (Basel, Switzerland), and Chelex® 100 resin was purchased from Bio-Rad (Reinach, Switzerland). All other chemicals were ACS grade or comparable and were obtained from Sigma-Aldrich Chemie (Buchs, Switzerland), Calbiochem (VWR International, Lucerne, Switzerland), or Acros Organics (Chemie Brunschwig, Basel, Switzerland). All solutions were prepared using degassed Millipore water. If appropriate, solutions were saturated with nitrogen or argon. Whenever complete absence of oxygen was required, Millipore water was degassed by three consecutive freezing–thawing cycles under vacuum.

### Synthetic peptide

A synthetic $\gamma$-$E_c$-1 peptide (purity greater than 90%) consisting of the first 25 amino acids of the full-length $E_c$-1 protein MGCDD KCGCA VPCPG GTGCR CTSAR was purchased from Sigma-Genosys (Haverhill, UK) and used for EXAFS, electrospray ionization mass spectrometry (ESI–MS), pH and metal ion titration experiments, and for the 2D $^1$H–$^1$H total correlation spectroscopy (TOCSY) and nuclear Overhauser enhancement spectroscopy (NOESY) NMR experiments of the $Cd_2\gamma$-$E_c$-1 form. All other experiments were conducted with the peptide overexpressed using the pGEX-4T-gEc1 construct.

### Plasmid construction

The complementary DNA sequence encoding for the first 24 amino acids of wheat $E_c$-1 without the N-terminal translation initiator Met was optimized for *Escherichia coli* codon usage. Two additional Gly and Ser residues were added to the N-terminus of the protein to ensure optimal thrombin cleavage, yielding the sequence GS GCDD KCGCA VPCPG GTGCR CTSAR.

The resulting construct was cloned into the pGEX-4T expression vector (GE Healthcare) using the *Bam*H1 and *Eco*R1 restriction sites, and the construct identity (pGEX-T4-gEc1) was subsequently verified by DNA sequencing. The constructed plasmid was transferred into the protease-deficient *E. coli* expression strain BL21(DE3).

### Protein expression and purification

$\gamma$-$E_c$-1 was overexpressed in the form of the glutathione *S*-transferase (GST)–MT fusion protein according to the GST purification manual (GE Healthcare). After induction with 1 mM isopropyl-$\beta$-D-thiogalactopyranoside at an optical density at 600 nm of 1, cells were harvested after 4–7 h at 37 °C and lysed by sonification. The supernatant was loaded on a pre-equilibrated GST-affinity column (GE Healthcare). After washing, the fusion protein was stripped from the column using 10 mM glutathione and cleaved with 1 unit thrombin per milligram of GST–MT fusion protein for 60 h at 25 °C. Final purification of $\gamma$-$E_c$-1 was performed by size-exclusion chromatography using a Superdex 30 pg column (GE Healthcare), and the molecular identity was verified by ESI–MS (Fig. 2). Crucial for GST–MT protein cleavage with thrombin was the demetalation of the fusion protein with 10 mM EDTA during the GST-affinity column washing step. All chromatographic steps were conducted in 100 mM phosphate-buffered saline at pH 7.3. Purified and completely oxidized $\gamma$-$E_c$-1 was dialyzed twice against 20 mM Tris/HCl pH 8.0, lyophilized, and stored at −80 °C. Average yields were 4 mg of purified protein per liter of cell culture medium. The full-length $E_c$-1 protein was prepared as described elsewhere [3].

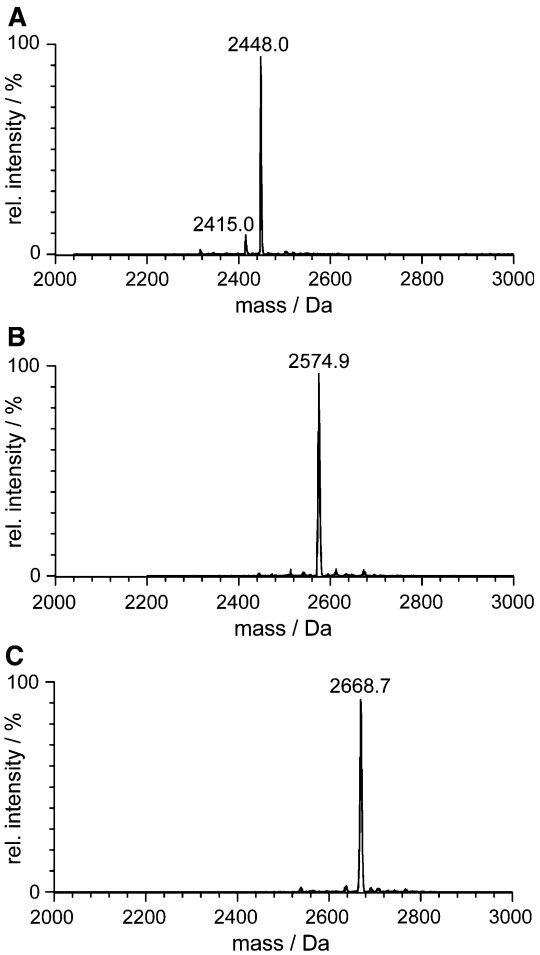


**Fig. 2** Deconvoluted ESI–MS spectra of **a** $\gamma$-$E_c$-1 in the metal-depleted apo form at pH 2 (calculated mass 2,448.0 Da), **b** $Zn_2\gamma$-$E_c$-1 at pH 7.5 (calculated mass 2,574.7 Da), and **c** $Cd_2\gamma$-$E_c$-1 at pH 7.5 (calculated mass 2,668.8 Da)

### Apo$\gamma$-$E_c$-1 preparation

Apo$\gamma$-$E_c$-1 was prepared freshly prior to each experiment. Typically 1–3 mg of oxidized $\gamma$-$E_c$-1 was incubated with 200 mM dithiothreitol in a 100 mM Tris/HCl solution (pH 8.0) for 1 h prior to acidification to pH 2 with 1 M HCl. The sample was applied to a G10 size-exclusion column (GE Healthcare) pre-equilibrated with 10 mM HCl and eluted under constant argon flow. The residual $Zn^{II}$, $Cd^{II}$, and $Cu^{I/II}$ content of apo$\gamma$-$E_c$-1 was below the detection limit of flame atomic absorption spectroscopy (F-AAS) (0.001 ppm). Prior to the subsequent metal ion reconstitution step, the solution of apo$\gamma$-$E_c$-1 was argon-saturated for 1 h in a $N_2$-flushed glove box and the protein concentration was determined via thiol quantification using the 2,2′-dithiodipyridine assay [17].

### Preparation of $Zn_2\gamma$-$E_c$-1, $Zn_6E_c$-1, and $Cd_2\gamma$-$E_c$-1

For all experiments 2 or 6 Eq of metal ions were titrated to the respective apo form in a $N_2$-flushed glove box. Subsequently, the pH was raised to 8.6 using Tris/HCl or $d_{11}$-Tris/HCl for the NMR samples. Reconstituted samples were dialyzed against a 20 mM solution of the respective Tris/HCl buffer or 5 mM $NH_4Ac$ for ESI–MS measurements and were concentrated by lyophilization.

### pH titrations followed by UV spectroscopy

800 µL of the respective reconstituted $E_c$-1 form (approximately 10 µM each) in 1 mM Tris/HCl pH 8.6 and 10 mM NaCl were titrated with diluted HCl as described in [3]. For the concurrent titration of both domains, $\gamma$-$E_c$-1 and $\beta_E$-$E_c$-1 were mixed in equimolar amounts. Plots of molar absorptivity at 230 nm for the $Zn^{II}$ forms and at 250 nm for the $Cd^{II}$-loaded species against pH were fitted with the program Origin 7.0® (OriginLab, Northampton, MA, USA) using three different functions, considering either one or two common apparent p$K_a$ values for the Cys residues in the presence of the respective metal ions as described in [3, 18].

### Titration of apo$\gamma$-$E_c$-1 with $Cd^{II}$

For each titration point, 90 µL of 32 µM apo$\gamma$-$E_c$-1 were mixed with the appropriate amount of a 1.25 mM $CdCl_2$ solution in a $N_2$-purged glove box. The pH was raised to approximately 8.6 with 100 mM Tris that had been pre-treated with Chelex 100, resulting in samples with 20 µM $E_c$-1, 20 mM Tris/HCl, and 10 mM NaCl. Samples were transferred into cuvettes, sealed, and UV spectra were recorded.

### Mass spectrometry

Samples of $Zn_2\gamma$-$E_c$-1 in 100 mM $NH_4Ac$ (pH 8) were treated with 2 Eq $Zn^{II}$ or 4 Eq $Cd^{II}$ and injected directly or with a prior acidification step into an Ultima API quadrupole time-of-flight spectrometer (Waters, UK). As a solvent, 10 mM $NH_4Ac$ in 50% MeOH (pH 7.5) or 50% acetonitrile with 0.2% formic acid (pH 2–3) was used. Scans were accumulated and further processed with the software MassLynx 3.5 (Micromass). Deconvolution of mass spectra was done by applying the maximum entropy algorithm of the MassLynx tool MaxEnt1. Electrospray parameters were capillary voltage 2.8 V, cone voltage 60 V, and source temperature 80 °C.

### X-ray absorption spectroscopy

To determine the average $Zn^{II}$-binding motif, K-edge X-ray absorption spectra of $Zn_2\gamma$-$E_c$-1 were recorded at beamline D2 of the EMBL Outstation Hamburg at DESY, Germany, as described in [19]. Data reduction, such as background removal, normalization, and extraction of the fine structure, was performed with KEMP [20] assuming a threshold energy of $E_{0,Zn} = 9{,}662$ eV. The extracted K-edge EXAFS data were converted to photoelectron wave vector $k$-space and weighted by $k^3$. Initial evaluation of the spectra by ABRA [21], which is based on EXCURV [22], included a systematic screening of approximately 400 potential binding motifs. The subsequent meta-analysis identified structural zinc sites, and thus in the final refinement the total number of ligands was fixed at four. The absence of multiple scattering contributions indicative of the binding of imidazole rings to the $Zn^{II}$ ion limited the refinement to the following parameters for each structural model: the atomic distances ($R$), the Debye–Waller factors ($2\sigma^2$), and a residual shift of the energy origin (EF) were refined, minimizing the fit index ($\Phi$). An amplitude reduction factor (AFAC) of 1.0 was used throughout the data analysis.

### NMR spectroscopy

$Zn_2^{15}N$-$\gamma$-$E_c$-1, $Cd_2^{15}N$-$\gamma$-$E_c$-1, $^{113}Cd_2\gamma$-$E_c$-1, and $Zn_6^{15}N$-$E_c$-1 samples were prepared as described above. The lyophilized proteins were dissolved in 10% $D_2O$/90% $H_2O$, 15 mM $d_{11}$-Tris/HCl pH 6.9, and 50 mM NaCl to a final concentration of 1 mM protein for $^1$H–NMR and 3 mM protein for $^{113}$Cd-NMR studies. $^1$H–NMR experiments to elucidate the protein backbone structure were recorded at 25 °C with Bruker Avance 700- and 600-MHz spectrometers. $^{113}$Cd-NMR experiments to investigate the binding sites of the $Cd^{II}$ ions were performed with a Bruker DRX 500-MHz spectrometer. Assignment of resonances in

$Cd_2\gamma$-$E_c$-1 and $Zn_6E_c$-1 was performed using 3D $^{15}$N-resolved TOCSY [23, 24] and NOESY [25, 26] spectra recorded with 80- and 120-ms mixing times, respectively. Distance restraints were derived from the 120-ms mixing time 3D $^{15}$N-resolved NOESY and 2D NOESY experiments. Resonance assignments for the $Zn_2\gamma$-$E_c$-1 domain were conducted using 2D TOCSY and 2D NOESY spectra with 80- and 120-ms mixing times, respectively. Additionally, a 120-ms mixing time 3D $^{15}$N-resolved NOESY spectrum as well as the information from the $Cd_2\gamma$-$E_c$-1 and $Zn_6E_c$-1 forms were used to validate the assignment. Distance restraints were again derived from the 3D $^{15}$N-resolved NOESY and 2D NOESY experiments. In all cases zero-quantum interference in the spectra was suppressed using an appropriate filter [27, 28]. $^{15}$N,$^{1}$H correlation maps were derived from a gradient-enhanced [$^{15}$N,$^{1}$H] heteronuclear single quantum coherence (HSQC) experiment using the Rance–Palmer trick for sensitivity enhancement [29, 30]. 1D $^{113}$Cd-NMR and 2D [$^{113}$Cd,$^{1}$H]-HSQC spectra, as well as 2D [$^{113}$Cd,$^{113}$Cd] correlation spectroscopy (COSY) experiments were recorded to investigate the metal cluster [31]. $^{3}J$[H$\beta$,Cd] couplings derived from a 2D [$^{113}$Cd,$^{1}$H]-HSQC spectrum allowed the individual Cd-Cys connectivities to be established.

Sequence-specific resonance assignment was performed using the method developed by Wüthrich and coworkers [32]. Assignments were achieved on the basis of information from 2D TOCSY, NOESY, 2D [$^{15}$N,$^{1}$H]-HSQC, 3D $^{15}$N-resolved NOESY, and 3D $^{15}$N-resolved TOCSY experiments. The 2D and 3D spectra were evaluated with the programs XEASY [33] and CARA [34], respectively. As a first step, the spin systems were identified in the 2D TOCSY or 3D $^{15}$N-resolved TOCSY experiments. Subsequently, spin systems were linked on the basis of nuclear Overhauser enhancement (NOE) information derived from 2D NOESY and 3D $^{15}$N-resolved NOESY. Once longer stretches had been identified, they were mapped onto the sequence of $\gamma$-$E_c$-1.

For the structure calculations, NOE peaks were picked and integrated using the program XEASY for 2D experiments and CARA for 3D experiments employing identical lower integration thresholds. Torsion angle dynamics [35] were performed with the *noeassign* [36] algorithm of the program CYANA 2.1 [37]. Structure calculations were started from 100 conformers with randomized torsion angle values. The 20 conformers with the lowest final target function value were further subjected to restrained energy minimization in explicit solvent against the AMBER force field [38] using the program OPALp [39, 40]. The resulting structures were deposited in the Protein Data Bank under accession codes 2I61 and 2I62. Structure figures were generated with the program MOLMOL [41].

## Results and discussion

### Initial quantification of metal ion binding

Reconstituted forms of $\gamma$-$E_c$-1 with $Zn^{II}$ and $Cd^{II}$ were analyzed for their metal ion content using F-AAS and the 2,2′-dithiodipyridine assay and yielded $M^{II}$ to thiol group ratios of 1:3, indicating the binding of two divalent metal ions per protein. The proposed metal stoichiometry was confirmed by mass spectrometry. For this 2 Eq of $Zn^{II}$ or 4 Eq of $Cd^{II}$ ions were added to a solution of $Zn_2\gamma$-$E_c$-1, and the resulting mixture was analyzed by ESI–MS at either acidic or neutral pH. The spectrum at acidic pH shows the apo$\gamma$-$E_c$-1 species as expected, whereas at neutral pH exclusively $Zn_2\gamma$-$E_c$-1 or $Cd_2\gamma$-$E_c$-1 is observed despite the addition of an excess of the respective metal ion (Fig. 2).

To corroborate this result, apo$\gamma$-$E_c$-1 was titrated with increments of $Cd^{II}$ and UV spectra were recorded (Fig. 3). They show the formation of the typical ligand-to-metal charge transfer bands at 245–250 nm indicative of $Cd^{II}$ coordination to thiolate groups.

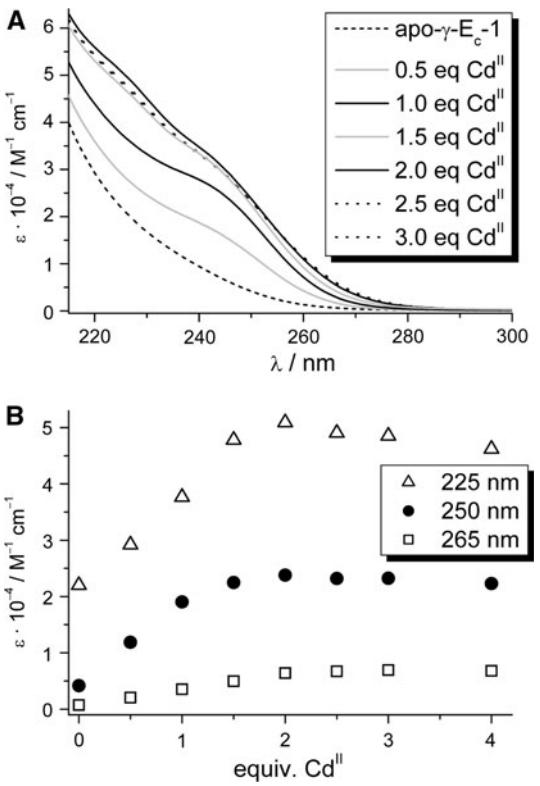


**Fig. 3** **a** UV spectra of the stepwise reconstitution of apo$\gamma$-$E_c$-1 with $Cd^{II}$ ions showing the evolution of the S → $Cd^{II}$ ligand-to-metal charge transfer bands around 250 nm. **b** Molar absorptivity at 225, 250, and 265 nm against the number of equivalents of $Cd^{II}$ ions added, with the maximum value being reached in all cases after addition of 2 Eq

These bands increase in absorptivity up to the addition of 2 Eq of $Cd^{II}$ and remain constant thereafter. It had already been shown that the full-length $E_c$-1 protein is able to coordinate six divalent metal ions and that four of them can be accommodated in the C-terminal $\beta_E$ domain [3, 10]. That the remaining two metal ions are bound within the N-terminal $\gamma$ domain was evidenced by ESI–MS measurements on a proteolytically digested $Zn_6E_c$-1 sample revealing the presence of a $Zn_2\gamma$-$E_c$-1 species [19]. In the same publication, a [$^{113}$Cd,$^{113}$Cd]-COSY spectrum of $^{113}Cd_6E_c$-1 showed cross-peaks between two $^{113}$Cd signals that originate from metal ions coordinated within the $\gamma$ domain and led to the proposal of a $Cd_2Cys_6$ cluster. The results with the separate $\gamma$-$E_c$-1 peptide presented here clearly demonstrate that the binding ability for two divalent metal ions is not restricted to the domain within the full-length protein. This validates our experimental approach to use the separate $\gamma$-$E_c$-1 sequence for the in-depth spectroscopic characterization of the $\gamma$ domain of wheat $E_c$-1.

## pH titrations of $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1

pH titrations followed by UV spectroscopy were performed to investigate the pH-dependent metal ion release. Although not tantamount, the apparent p$K_a$ values of the Cys residues obtained in presence of the respective metal ion are a good indication for the relative binding affinity of the metal ion to the MT. A number of apparent p$K_a$ values for different $Zn^{II}$ and $Cd^{II}$ MTs were recently reviewed [13]. In the context of the study presented here, the pH stability of $\gamma$-$E_c$-1 in comparison to the full-length protein and the $\beta_E$ domain is of special interest. The UV spectra of the pH titration of $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1 are depicted in Fig. 4 as are the plots of molar absorptivity at 230 nm ($Zn^{II}$ form) and 250 nm ($Cd^{II}$ form) against the respective pH values.

Fitting of the data was performed as described in [6, 21], and the results are presented in more detail in the electronic supplementary material. As the p$K_a$ values obtained for both $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1 were significantly higher than determined for the respective $\beta_E$ and full-length forms (Fig. 5), also a titration of an equimolar mixture of both domains was performed. The resulting p$K_a$ values for the mixed domains lie in between the values obtained for the respective $\gamma$ and $\beta_E$ domains, but are still significantly higher than the values for the full-length $E_c$-1 forms (Fig. 5).

Hence, it appears that not the mere presence but rather the close proximity of the respective other domain leads to the observed increased pH stability of the full-length protein. Although possible, it seems unlikely that the nature of the five amino acids SGAAA between the $\gamma$ domain and the $\beta_E$ domain, which were removed in the separately expressed domains, has a major influence as the residues are neither charged nor especially bulky or hydrophobic. So far, also no indications for any sort of interactions between the two domains have been observed, judging from the lack of corresponding NOE signals in the NMR experiments. Furthermore, $^{15}$N dynamics data, in particular

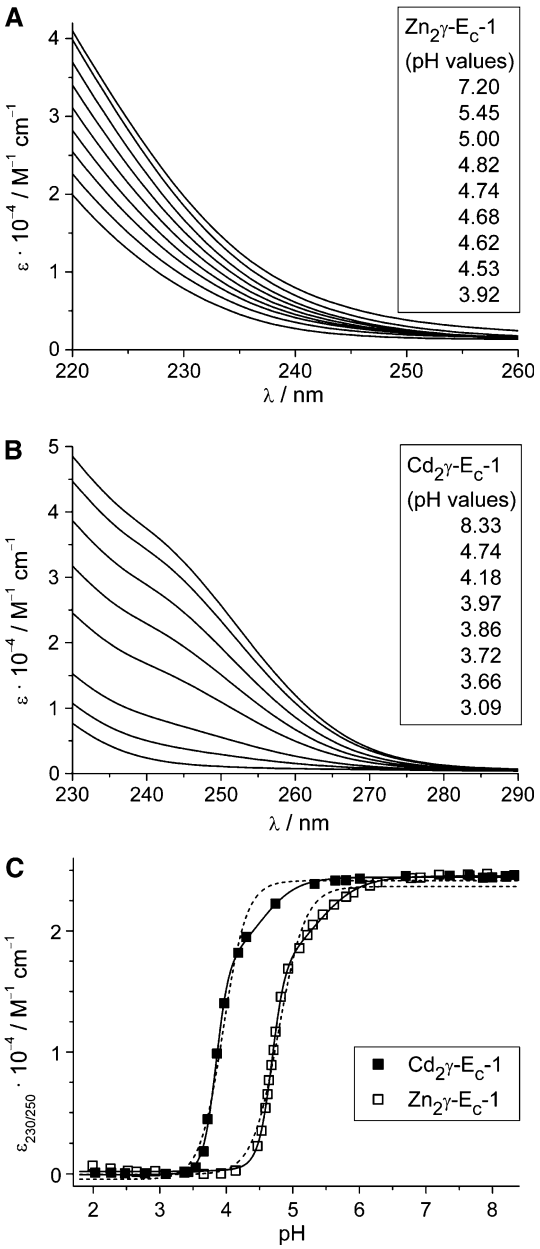


**Fig. 4** Representative UV spectra of the titration of **a** $Zn_2\gamma$-$E_c$-1 and **b** $Cd_2\gamma$-$E_c$-1 with increasing amounts of HCl. **c** Molar absorptivity at 230 nm for the $Zn^{II}$ form and at 250 nm for the $Cd^{II}$ form versus pH. To allow better comparability, the values obtained for the apo forms in both titrations were shifted to zero and in addition the plot of the $Zn^{II}$ form was normalized to the values obtained for the $Cd^{II}$ form. Curve fits were performed with equations considering one (*dashed lines*) or two (*solid lines*) apparent p$K_a$ values as described in the electronic supplementary material

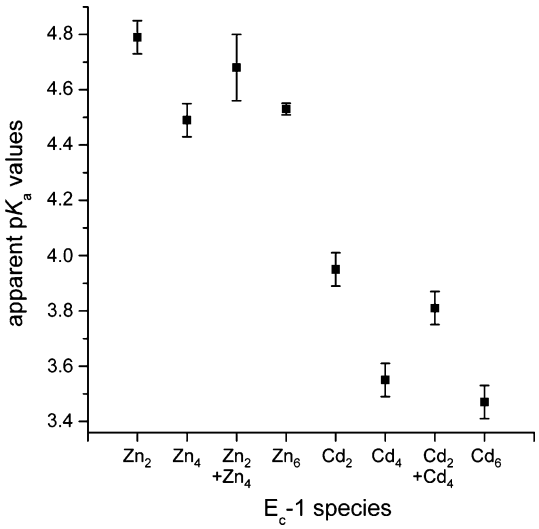


**Fig. 5** Apparent p$K_a$ values of the Cys residues in $\gamma$-E$_c$-1, $\beta_E$-E$_c$-1, and full-length wheat E$_c$-1 in presence of $Zn^{II}$ or $Cd^{II}$ ions. The error bars show the $3\sigma$ error range. From *left* to *right*: 4.79(6), 4.49(6), 4.68(12), 4.530(21), 3.95(6), 3.55(6), 3.81(6), and 3.47(6)

values of the $^{15}N\{^1H\}$-NOE, indicate that the linker comprising residues 26–30 is fully flexible [10]. One possible explanation for the low p$K_a$ values of the full-length protein is that the spatial proximity of the respective other domain leads to a reduced solvent accessibility and hence to a deferred metal ion displacement by protons. Alternatively, formation of intermediate species in the full-length protein at decreasing pH values could occur, slowing down the metal ion release process, i.e., migration of metal ions between the two domains or even transient generation of a new cluster arrangement. However, such species would not be detected in the structural investigation presented here, as the NMR experiments were performed at pH 6.9, whereas metal ion release only starts below pH 6.5 for the $Zn^{II}$ form and below pH 5.5 for the $Cd^{II}$ form. It is interesting to note that the stabilizing effect of the respective other domain is obviously much more pronounced for the $\gamma$ domain, whereas the $\beta_E$ domain shows within the error limits the same apparent p$K_a$ values as the full-length protein. Shifts in p$K_a$ values are subtle probes of thermodynamic protein stability, and our data indicate that the $\gamma$ domain is intrinsically less stable than the $\beta_E$ domain when in isolation, although both are structured and capable of metal binding.

An additional surprising result was obtained when the pH titration data for $Zn_2\gamma$-E$_c$-1 and $Cd_2\gamma$-E$_c$-1 were evaluated more closely. The absorptivity values obtained from the curve fit with the equation considering two apparent p$K_a$ values reveal a decrease by approximately one third for the first protonation step, characterized by p$K_{a2}$, i.e., $\Delta\varepsilon$ 2,900 ± 400 $M^{-1}$ $cm^{-1}$ for $Zn_2\gamma$-E$_c$-1 and $\Delta\varepsilon$ 7,200 ± 1,700 $M^{-1}$ $cm^{-1}$ for $Cd_2\gamma$-E$_c$-1, and a decrease by two thirds for the second step, characterized by p$K_{a1}$, i.e., $\Delta\varepsilon$ 5,700 ± 400 $M^{-1}$ $cm^{-1}$ for $Zn_2\gamma$-E$_c$-1 and $\Delta\varepsilon$ 17,300 ± 1,700 $M^{-1}$ $cm^{-1}$ for $Cd_2\gamma$-E$_c$-1 (see the electronic supplementary material). Disregarding the contribution of bridging thiolate ligands to the ligand-to-metal charge transfer bands, which is considerably less than the contribution of terminal thiolate ligands, the absorptivity decreases suggest the loss of two terminal metal–thiolate bonds in the first step and of four metal–thiolate bonds in the second step. This suggests that in the first step one metal ion is released and hence the contribution of two terminal and two bridging metal–thiolate bonds is lost, whereas in the second step the second metal ion is released and hence the contribution of the residual four metal–thiolate bonds disappears. As a result, the two metal ions in $\gamma$-E$_c$-1 seem to be released at pH values approximately 0.6–0.7 units apart and this might also be an indication that $\gamma$-E$_c$-1 contains two metal ion binding sites with different affinities.

## Extended X-ray absorption fine structure spectroscopy of $Zn_2\gamma$-E$_c$-1

EXAFS spectra were recorded to analyze the coordination environment of the $Zn^{II}$ ions in $Zn_2\gamma$-E$_c$-1. The results reveal the presence of four sulfur ligands and no contribution of lighter ligands with nitrogen- or oxygen-donor atoms within the error limits. The Zn–S distances of 2.332(3) Å are in the normal range for $Zn^{II}$ coordination by thiol ligands (Fig. 6, Table 1).

Assuming the presence of a Zn–Zn interaction with a distance of 3.163(6) Å improves the fit index significantly from 0.2988 to 0.2141. However, the error range for the corresponding average number of ligands is relatively high.

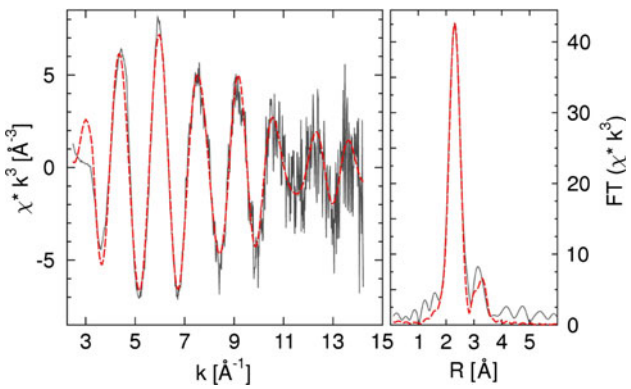


**Fig. 6** EXAFS (*left*) and corresponding Fourier transform (*right*) of $Zn_2\gamma$-E$_c$-1. The EXAFS is dominated by a single frequency, originating from sulfur backscattering. In the refinement, no other first-shell contribution could be identified. In line with this result, the Fourier transform is dominated by a single peak at 2.3 Å. The additional peak above 3 Å is refined as a metal–metal contribution, indicative of bridging sulfur ligands. The corresponding parameters are given in Table 1

**Table 1** EXAFS refinement parameters for $Zn_2\gamma$-$E_c$-1 ($\Delta E$ = 13–770 eV, $E_{0,Zn}$ = 9,662 eV)

| *N* | M···L | *R* (Å) | $2\sigma^2$ (Å$^2$) | EF (eV) | $\Phi$ |
|---|---|---|---|---|---|
| 4 | Zn···S | 2.332(3) | 0.0134(5) | −7.6(5) | 0.2141 |
| 0.5(3) | Zn···Zn | 3.163(6) | 0.006(3) | | |

Best models with average numbers (*N*) of ligand atoms (*L*), their distance to the metal ion (*R*), the respective Debye–Waller factor ($2\sigma^2$), the Fermi energy for all shells (*EF*), and the fit index ($\Phi$) indicating the quality of the fit are shown. The total error is estimated to be 0.01 Å or smaller for the first-shell distance and 0.05–0.1 Å for the metal–metal contribution. The numerical error margins are given on the $2\sigma$ level in *parentheses*

These findings together with the results from the concentration and ESI–MS measurements presented above strongly suggest the formation of a $Zn_2Cys_6$ metal–thiolate cluster with four terminal and two bridging thiolate groups. Such a cluster is in accordance with the $Cd_2Cys_6$ cluster proposed to be formed in $^{113}Cd_6$Ec-1 on the basis of the [$^{113}$Cd,$^{113}$Cd]-COSY spectrum mentioned already [19].

## NMR solution structures of the separate $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1 peptides

Except for the first two amino acids Gly and Ser, which were engineered to improve proteolytic cleavage of the GST-fusion protein by thrombin, all residues could be identified by 3D $^{15}$N-resolved NOESY and TOCSY NMR experiments in case of the $Cd^{II}$ form or by 3D $^{15}$N-resolved NOESY, 2D TOCSY, and [$^{15}$N,$^{1}$H]-HSQC experiments for the $Zn^{II}$ form. Experiments with the NMR-active $^{113}$Cd nucleus were performed to probe the metal ion coordination sphere. [$^{113}$Cd,$^{1}$H]-HSQC spectra allow the observation of cross-peaks based on $^3J$ couplings between the H$\beta$ protons of the Cys residues and the respective coordinated $Cd^{II}$ ions. In theory, each of the 12 H$\beta$ protons of the six Cys residues present in the peptide should display a $^3J$ coupling to one $^{113}Cd^{II}$ ion, or in the case of bridging thiolate ligands, to two $Cd^{II}$ ions. As shown in Fig. 7, indeed most of the Cys residues correlate with metal ions; however, not two as expected but rather three possible bridging Cys residues were identified.

Despite differences in size and electronegativity, $Cd^{II}$ and $Zn^{II}$ forms of MTs have been interchangeably used for structural studies [31, 42, 43]. In the case of the $\gamma$-$E_c$-1 domain, such an assumption is justified by the reasonable agreement between proton chemical shifts in the two forms (Fig. 8).

Chemical shift assignment was based on the sequence-specific sequential resonance assignment procedure developed by Wüthrich and coworkers [32]. Overall, the completeness of proton assignment was 99%. No long-range NOEs reflecting contacts of residues from the C-terminal and N-terminal regions were observed in the NOESY spectra. Pro-12 (in contrast to Pro-14) was shown to be connected via a *cis* peptide bond to the previous residue in both metal isoforms (Fig. S1), as deduced from comparably strong NOEs of the sequential $\alpha$ protons. The comparison of backbone amide $^1$H and $^{15}$N chemical shifts for residues of the $\gamma$ domain indicates that differences as large as 0.4 or 1.1 ppm are observed between the Zn and Cd species, respectively (Fig. S2). The largest differences are not limited to residues that are involved in metal coordination, but also include some of the residues in the small loop regions. A comparison of the amide proton and nitrogen chemical shifts of $Zn_2\gamma$-$E_c$-1 with the values previously determined for the full-length $Zn_6E_c$-1 protein reveals that chemical shift differences are negligible and limited to the terminal residues (Fig. S2). The latter are expected to be different owing to the slightly N-terminally modified sequence (GlySer instead of Met) or owing to the additional presence of the C-terminal $\beta_E$ domain in the full-length species. As mentioned already, a somewhat surprising result of the analysis is that although the analysis of the p$K_a$ values indicated differences between corresponding segments in the isolated $\gamma$ domain compared with the full-length protein, no such differences could be detected in the backbone amide chemical shifts. We

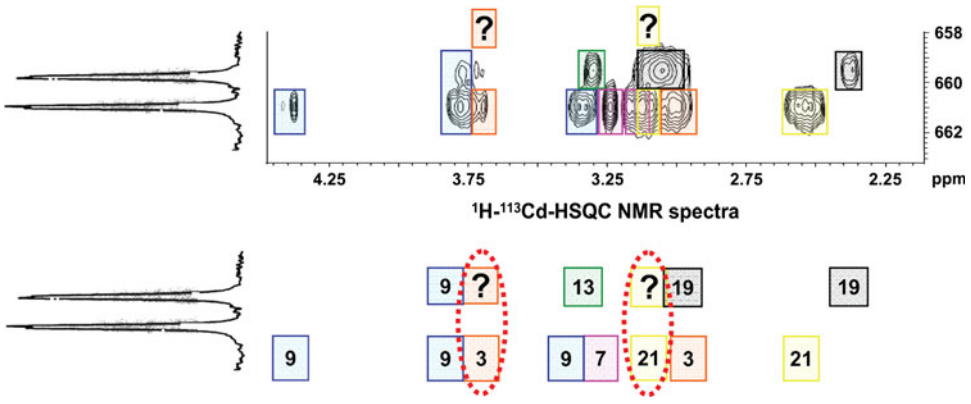


**Fig. 7** The 1D $^{113}$Cd NMR spectrum of $Cd_2\gamma$-$E_c$-1 shows two doublets representing the two $Cd^{II}$ ions. The 2D [$^{113}$Cd,$^{1}$H]-HSQC NMR spectrum allows two possible solutions for the bridging Cys residues: Cys-9/Cys-3 and Cys-9/Cys-21. The assignment of cross-peaks to Cys residues is schematically depicted at the *bottom*

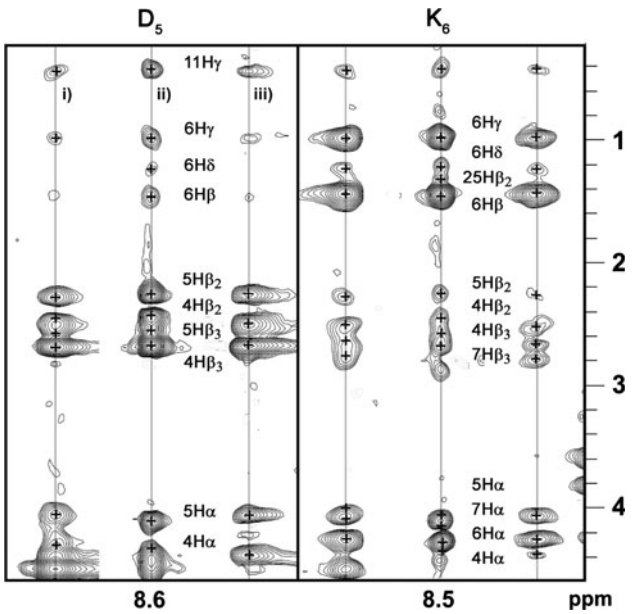


**Fig. 8** Comparison of the 3D $^{15}$N–$^{1}$H–$^{1}$H NOESY NMR slices of Asp-5 and Lys-6 depicted for the separate $Zn_2\gamma$-$E_c$-1 (*i*) and $Cd_2\gamma$-$E_c$-1 (*ii*) peptides as well as for the $Zn_2\gamma$-$E_c$-1 domain (*iii*) within the full-length protein. Assigned NOE cross-peaks are labeled. A high degree of NOE pattern homology is apparent

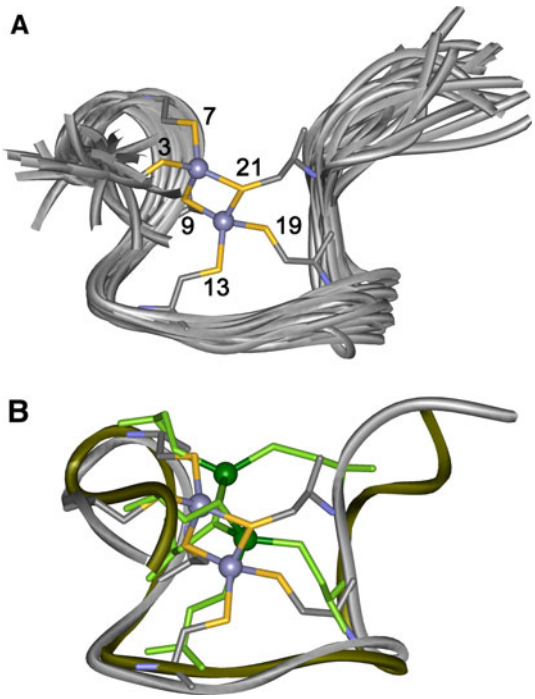


**Fig. 9** **a** Structure bundle for $Zn_2\gamma$-$E_c$-1 showing the Cys-9/Cys-21 metal cluster arrangement. The backbones are shown in *gray*, the Cys residues of one representative structure in stick mode, and the two corresponding $Zn^{II}$ ions as *light-blue spheres*. Cys residues are numbered according to their position in the amino acid sequence given in Fig. 1. **b** Backbone overlay of two representative structures of $Zn_2\gamma$-$E_c$-1 with Cys-9/Cys-21 connectivity as in **a** and with Cys-9/Cys-3 arrangement (*olive backbone*, Cys residues as *green sticks*, $Zn^{II}$ ions as *dark-green spheres*)

speculate that there may be contacts between the two domains that, however, are very transient (and possibly also unspecific) in nature.

Initial structures calculated without addition of explicit metal–Cys(S$\gamma$) restraints (Fig. S3) converge for the amino acid residues Gly-2 to Gly-18. However, the positions of the C-terminal residues and of the thiol groups of the Cys residues deviate substantially. When upper distance restraints were added that enforced tetrahedral geometry and metal–sulphur distances derived from EXAFS experiments, the calculations for the C-terminal part of the structure also converged. Although Cys-9 (numbering according to the amino acid sequence given in Fig. 1) could be unambiguously identified as a bridging Cys residue on the basis of the [$^{113}$Cd,$^{1}$H]-HSQC experiment, the nature of the second bridging Cys residue could not be experimentally established. Accordingly, independent structure calculations were performed assuming the bridging residues to be (1) Cys-9 and Cys-21, (2) Cys-9 and Cys-3, and (3) Cys-9 and Cys-13, with all other Cys residues coordinating in a terminal fashion. The calculation using Cys-9 and Cys-13 as bridging residues resulted in no low-energy conformer and was therefore excluded from further analysis. The resulting structures containing the metal cluster arrangements Cys-9/Cys-21 and Cys-9/Cys-3 are representatively shown for $Zn_2\gamma$-$E_c$-1 in Fig. 9, and the statistics from the corresponding structure calculations are summarized in Table 2.

Similar to previously determined structures of MTs, the $\gamma$ domain of $E_c$-1 is devoid of regular secondary structure. Superposition of backbone atoms result in root mean square deviations of 0.62–0.8 Å for backbone atoms of residues 2–22, and 1.3–1.5 for all heavy atoms, and the values are very similar for the $Cd^{II}$-loaded and $Zn^{II}$-loaded isoforms (see Table 2). The overall fold resembles a hook, in which the stem part is formed by the segments containing Cys7-9 and Cys19-21, and the loop by the coordination of Cys-13 to the metal. Two loops bridge the latter residue and the next metal-anchoring residues Cys-9 and Cys-19. Coordination of Cys-3 brings the N-terminal segment into proximity. No significant differences in conformation that were unambiguously supported by the NOEs were observed between the $Cd^{II}$-loaded and $Zn^{II}$-loaded peptides despite the fact that in part substantial chemical shift differences are observed (vide supra). A superposition of the backbone atoms of conformers, in which either Cys-3 or Cys-21 was constrained to be the bridging ligand, revealed that only small structural adaptations were necessary to transform one form into the other. Considering that only upper-distance limits were derived from the

**Table 2** Statistics of structure calculations

| | Cd3[a] | Cd21[a] | Zn3[a] | Zn21[a] | Zn3 (full)[a] | Zn21 (full)[a] |
|---|---|---|---|---|---|---|
| NMR distance restraints | | | | | | |
| Total NOE | 374 | 372 | 462 | 463 | 307 | 310 |
| Short range: $\|i - j\| \leq 1$ | 322 | 321 | 396 | 395 | 268 | 268 |
| Medium range: $1 > \|i - j\| > 5$ | 61 | 58 | 69 | 70 | 47 | 44 |
| Long range: $\|i - j\| \geq 5$ | 15 | 13 | 20 | 20 | 13 | 15 |
| Maximal distance restraint violation (Å) | 0.11 | 0.11 | 0.11 | 0.19 | 0.16 | 0.44 |
| AMBER energies (kcal/mol) | | | | | | |
| Total (mean ± SD) | −679 ± 76 | −711 ± 43 | −697 ± 70 | −693 ± 75 | −679 ± 52 | −641 ± 67 |
| van der Waals | −14 ± 4 | −7 ± 5 | −12 ± 5 | −12 ± 6 | −9 ± 5 | −8 ± 7 |
| RMSDs from idealized geometry | | | | | | |
| Bond lengths (Å) | 0.0142 ± 0.0002 | 0.0138 ± 0.0002 | 0.0136 ± 0.0002 | 0.0142 ± 0.0002 | 0.0138 ± 0.0002 | 0.0138 ± 0.0002 |
| Bond angles (°) | 2.15 ± 0.06 | 2.03 ± 0.06 | 1.94 ± 0.06 | 2.42 ± 0.06 | 2.11 ± 0.06 | 2.24 ± 0.06 |
| Ramachandran plot statistics (%) | | | | | | |
| Residues in most favored regions | 80.9 | 70.9 | 74.4 | 69.4 | 65.6 | 63.5 |
| Residues in additionally allowed regions | 18.2 | 28.5 | 24.7 | 28.5 | 32.4 | 32.4 |
| Residues in generously allowed regions | 0.9 | 0.6 | 0.9 | 1.8 | 2.1 | 3.2 |
| Residues in disallowed regions | 0 | 0 | 0 | 0.3 | 0 | 0.9 |
| RMSDs from the mean coordinates (Å) | | | | | | |
| N, Cα, and C′ of residues 2–22 | 0.70 ± 0.14 | 0.69 ± 0.33 | 1.06 ± 0.34 | 0.97 ± 0.25 | 1.29 ± 0.44 | 1.50 ± 0.45 |
| Heavy atoms of residues 2–22 | 1.37 ± 0.26 | 1.45 ± 0.37 | 1.76 ± 0.35 | 1.54 ± 0.27 | 1.97 ± 0.43 | 2.26 ± 0.46 |

*NOE* nuclear Overhauser enhancement, *RMSD* root mean square deviation, *SD* standard deviation

[a] The number describes the residue that presents the second bridging Cys moiety (in addition to Cys-9; see the text). Numbering is performed with respect to the sequence of the full-length protein as given in Fig. 1

NOEs and taking the inherent dynamics of the system as well as the low proton density in MTs, which are largely devoid of regular secondary structure or tertiary contacts, into account, we feel that no sound statements on structural differences of the $Zn^{II}$-loaded or $Cd^{II}$-loaded species can be made on the basis of the present data. Moreover, the lack of explicit $^{113}$Cd-Cys(H$\beta$) cross-peaks in the [$^{113}$Cd,$^{1}$H]-HSQC experiment for these residues, and the fact that the target functions of the calculated conformers are very similar precludes unambiguous determination of the nature of the second bridging Cys residue. Whether this is because the coordination mode in the peptide changes dynamically or whether the NMR data are simply insufficient to describe a unique coordination mode remains unclear presently.

An $M^{II}_2Cys_6$ cluster as identified in the $\gamma$ domain is unprecedented for MTs so far, but a very similar $Zn_2Cys_6$ cluster was previously observed in the transcription factor GAL4 from *Saccheromyces cerevisiae* [44]. Also here, the metal–Cys connectivities were probed by replacement of $Zn^{II}$ ions by $^{113}Cd^{II}$. Since the two $^{113}$Cd resonances at 669 and 707 ppm are well separated, it was possible to identify the bridging Cys residues using selectively decoupled [$^{113}$Cd,$^{1}$H]-HSQC spectra. However, such an experiment is not feasible in the case of $Cd_2\gamma$-$E_c$-1 owing to the very small chemical shift difference of only 2 ppm. Nevertheless, two further peculiarities in the NMR spectra corroborate the concomitant presence of two different, probably interchanging cluster arrangements. Firstly, in $^{113}Cd_2\gamma E_c$-1 TOCSY spectra, cross-peaks due to the geminal H$\beta_2$–H$\beta_3$ correlation for the putative bridging Cys-3 and Cys-21 are significantly broadened, and the corresponding correlation for Cys-9 is broadened beyond detection, indicating the presence of exchange processes. Secondly, in contrast to spectra recorded for the full-length protein [19], no mutual coupling was observed in the [$^{113}$Cd,$^{113}$Cd]-COSY spectrum of $^{113}Cd_2\gamma$-$E_c$-1, which might again be explained by intermediate exchange processes occurring in the isolated domain.

### NMR solution structure of $Zn_2\gamma$-$E_c$-1 as part of the full-length $Zn_6E_c$-1 protein and comparison with the separate $Zn_2\gamma$-$E_c$-1 peptide

Spectroscopic and spectrometric studies [3, 10, 19, 45] have revealed that wheat $Zn_6E_c$-1 is a two-domain protein. The larger C-terminal domain, termed extended $\beta$ or $\beta_E$, consists of 51 amino acids and embeds a mononuclear $ZnCys_2His_2$ site as well as a trinuclear $Zn_3Cys_9$ metal–thiolate cluster with similarity to the $\beta$ domain of vertebrate MTs. As described already, the smaller 24 amino acids long N-terminal domain, $\gamma$-$E_c$-1, folds around a $Zn_2Cys_6$ cluster. Chemical shift mapping accompanied by $^{15}N$ relaxation experiments was used to confirm identical folds of the $\beta_E$ domain in the form of the separately expressed peptide as well as being part of the full-length $Zn_6E_c$-1 protein. To confirm that this is also true for the $\gamma$ domain, chemical shift assignment and identification of close contacts from NOESY spectra were performed for the isolated peptides ($Cd^{II}$ and $Zn^{II}$ isoforms) of the $\gamma$ domain as well as for the corresponding part in the full-length protein. Similar chemical shifts and NOESY cross-peaks (Fig. 8) indicate analogous peptide folding. A comparison of the solution structure bundle calculated for the embedded and for the independent $\gamma$ domain ($Zn^{II}$ isoform, Cys-9/Cys-21 bridging) is given in the electronic supplementary material. Indeed, only minor differences between the two conformations are observed.

### Comparison of $Zn_2\gamma$-$E_c$-1 with the $Zn_2Cys_6$ cluster in GAL4

As mentioned already, $Zn_2Cys_6$ clusters have so far only been structurally described in yeast transcription factors [44, 46, 47]. Amino acid sequence alignments of the latter reveal a completely conserved Cys distribution pattern and a high conservation of Lys residues, which play a major role in the interaction of these proteins with DNA (Fig. 10). In contrast, the Cys distribution pattern of the $\gamma$-$E_c$-1 domain differs significantly from that of the transcription factors, and only three positively charged residues, one Lys and two Arg, are present. In addition, whereas the fold of the protein backbone in the yeast transcription factors can be described as a loop, the backbone of $\gamma$-$E_c$-1 is S-shaped or resembles a hook (Figs. 9, 10). Taking these findings together, despite the similarity of the metal–thiolate clusters, recognition of DNA by $\gamma$-$E_c$-1 in the same fashion as observed for the yeast transcription factors seems unlikely.

## Conclusions

Complementing our investigation of the $Zn_4\beta_E$-$E_c$-1 domain including the determination of its solution structure by NMR [10, 19], we have now completed the second part of the puzzle by presenting a study of the properties and the solution structure of the $\gamma$ domain of wheat $E_c$-1. ESI–MS experiments in conjunction with F-AAS measurements and metal ion titrations followed by UV spectroscopy clearly confirm the ability of the N-terminal $E_c$-1 fragment to coordinate two $Zn^{II}$ or $Cd^{II}$ ions even in the absence of the $\beta_E$ domain. Tetrahedral tetrathiolate coordination of the bound $Zn^{II}$ ions was established by EXAFS measurements in addition to the presence of a short Zn–Zn distance of 3.16 Å. pH titrations of $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1 revealed higher apparent p$K_a$ values of the Cys residues than previously determined for the $\beta_E$ domain and full-length $E_c$-1. This earlier protonation of thiolate ligands is paralleled by an increased peptide backbone flexibility compared with that of the $\beta_E$ domain [10]. A pH titration of an equimolar mixture of the $\gamma$ domain and $\beta_E$ domain yielded intermediate p$K_a$

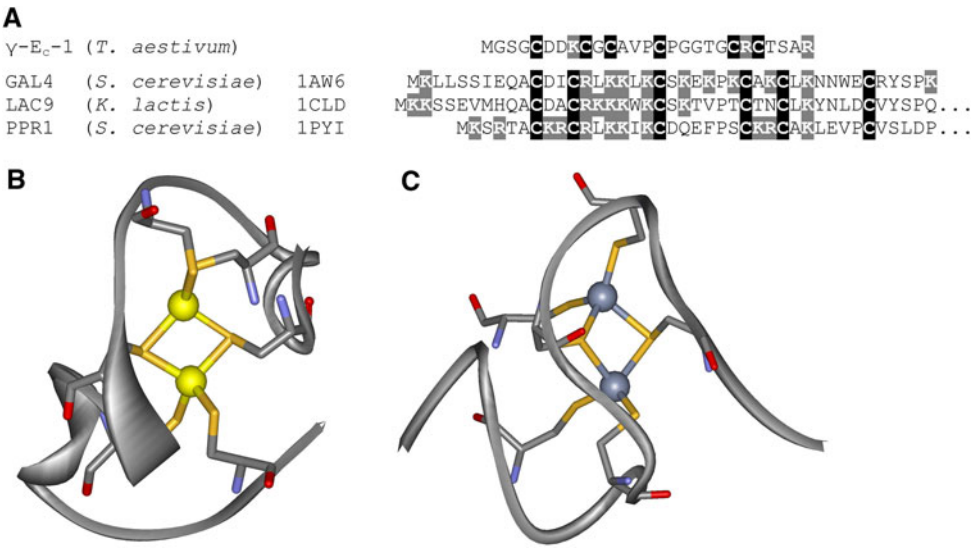


**Fig. 10** **a** Amino acid sequence of $\gamma$-$E_c$-1 and alignment of the sequences from the three yeast transcription factors GAL4, LAC9, and PPR1 with the species name and Protein Data Bank accession code given. Cys residues are highlighted with a *black background* and Lys and Arg residues are highlighted with a *gray background*. NMR solution structure of **b** $Cd_2$GAL4 [44] and **c** $Zn_2\gamma$-$E_c$-1 with the metal ions drawn as *yellow spheres* or *blue-grey spheres* and the Cys residues presented in stick mode. The N-terminus of each structure is positioned at the (*upper*) *right side*, respectively

values, which, however, differ from the values obtained with the full-length protein. This might indicate a yet unidentified interaction between the two domains in the full-length protein that increases the pH stability especially of the $M_2^{II}Cys_6$ cluster of the $\gamma$ domain. Owing to the low percentage or even lack of regular secondary structure in MTs, the metal clusters critically contribute to the overall protein fold. Hence, only when the metal-coordinating residues have been identified can the structure be determined correctly. In the case of the $\gamma$ domain, [$^{113}$Cd,$^{1}$H]-HSQC spectra indicated three possible metal ion-to-Cys connectivities, one of which could be eliminated during the structure calculation. On the basis of the $^{113}$Cd NMR and $^{1}$H NMR studies of the separate $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1 peptides as well as of the embedded domain in the full-length protein, we propose the presence of a highly dynamic metal cluster, possibly switching between two slightly different cluster arrangements recruiting either Cys-3 or Cys-21 as the second bridging thiolate ligand. This flexibility is in line with the observed decreased rigidity of the $\gamma$ domain compared with the $\beta_E$ domain as observed in $^{15}$N relaxation experiments [10]. Overall, the structures of the separate $Zn_2\gamma$-$E_c$-1 and $Cd_2\gamma$-$E_c$-1 peptides show only minor differences within the error limits In addition, when the chemical shifts of the backbone amide protons and nitrogen atoms in the 24-residue N-terminal segment of $Zn_6E_c$-1 and in the separate $Zn_2\gamma$-$E_c$-1 peptide were monitored, only small differences were observed. Hence, the solution structure of the separately expressed $\gamma$-$E_c$-1 peptide can be reliably taken as a model for the $\gamma$ domain in the full-length $E_c$-1 protein.

**Acknowledgments** We thank Peter Güntert for refining the CYANA structures with a full force field. This work was supported by the Swiss National Science Foundation (SNSF Professorship PP002-119106/1 to E.F.).

## References

1. Vallee BL (1979) In: Kägi JHR, Nordberg M (eds) Metallothionein. Birkhäuser, Basel, pp 19–40
2. Binz P-A, Kägi JHR (1999) In: Klaassen C (ed) Metallothionein IV. Birkhäuser, Basel, pp 7–13
3. Peroza EA, Freisinger E (2007) J Biol Inorg Chem 12:377–391
4. Hanley-Bowdoin L, Lane BG (1983) Eur J Biochem 135:9–15
5. Lane BG, Kajioka R, Kennedy TD (1987) Biochem Cell Biol 65:1001–1005
6. Blindauer CA, Harrison MD, Parkinson JA, Robinson AK, Cavet JS, Robinson NJ, Sadler PJ (2001) Proc Natl Acad Sci USA 98:9593–9598
7. Braun W, Vašák M, Robbins AH, Stout CD, Wagner G, Kägi JHR, Wüthrich K (1992) Proc Natl Acad Sci USA 89:10124–10128
8. Narula SS, Brouwer M, Hua Y, Armitage IM (1995) Biochemistry 34:620–631
9. Riek R, Prêcheur B, Wang Y, Mackay EA, Wider G, Güntert P, Liu A, Kägi JHR, Wüthrich K (1999) J Mol Biol 291:417–428
10. Peroza EA, Schmucki R, Güntert P, Freisinger E, Zerbe O (2009) J Mol Biol 387:207–218
11. Palmiter RD (1998) Proc Natl Acad Sci USA 95:8428–8430
12. Coyle P, Philcox JC, Carey LC, Rofe AM (2002) Cell Mol Life Sci 59:627–647
13. Freisinger E (2008) Dalton Trans 6663–6675
14. Kawashima I, Kennedy TD, Chino M, Lane BG (1992) Eur J Biochem 209:971–976
15. Finkelstein RR, Gampala SSL, Rock CD (2002) Plant Cell 14:S15–S45
16. Laudencia-Chingcuanco DL, Stamova BS, You FM, Lazo GR, Beckles DM, Anderson OD (2007) Plant Mol Biol 63:651–668
17. Pedersen AO, Jacobsen J (1980) Eur J Biochem 106:291–295
18. Freisinger E (2007) Inorg Chim Acta 360:369–380
19. Peroza EA, Al Kaabi A, Meyer-Klaucke W, Wellenreuther G, Freisinger E (2009) J Inorg Biochem 103:342–353
20. Korbas M, Marsa DF, Meyer-Klaucke W (2006) Rev Sci Instrum 77:1–5
21. Wellenreuther G, Parthasarathy V, Meyer-Klaucke W (2010) J Synchrotron Radiat 17:25–35
22. Binsted N, Strange RW, Hasnain SS (1992) Biochemistry 31:12117–12125
23. Braunschweiler L, Ernst RR (1983) J Magn Reson 53:521–528
24. Bax A, Davis DG (1985) J Magn Reson 65:355–360
25. Kumar A, Ernst RR, Wüthrich K (1980) Biochem Biophys Res Commun 95:1–6
26. Macura S, Ernst RR (1980) Mol Phys. 41:95–117
27. Otting G (1990) J Magn Reson 86:496–508
28. Rance M, Bodenhausen G, Wagner G, Wüthrich K, Ernst RR (1985) J Magn Reson 62:497–510
29. Palmer AG, Cavanagh J, Wright PE, Rance M (1991) J Magn Reson 93:151–170
30. Kay LE, Keifer P, Saarinen T (1992) J Am Chem Soc 114:10663–10665
31. Vašák M (1998) Biodegradation 9:501–512
32. Wüthrich K (1986) NMR of proteins and nucleic acids. Wiley, New York
33. Bartels C, Xia TH, Billeter M, Güntert P, Wüthrich K (1995) J Biomol NMR 6:1–10
34. Keller RLJ (2004) Computer aided resonance assignment tutorial. Cantina, Goldau
35. Güntert P, Mumenthaler C, Wüthrich K (1997) J Mol Biol 273:283–298
36. Herrmann T, Güntert P, Wüthrich K (2002) J Mol Biol 319:209–227
37. Güntert P (2003) Progr Nucl Magn Reson Spectrosc 43:105–125
38. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
39. Koradi R, Billeter M, Güntert P (2000) Comput Phys Commun 124:139–147
40. Luginbühl P, Güntert P, Billeter M, Wüthrich K (1996) J Biomol NMR 8:136–146
41. Koradi R, Billeter M, Wüthrich K (1996) J Mol Graphics 14:51–55
42. Braun W, Wagner G, Worgotter E, Vašák M, Kägi JHR, Wüthrich K (1986) J Mol Biol 187:125–129
43. Messerle BA, Schaffer A, Vašák M, Kägi JHR, Wüthrich K (1992) J Mol Biol 225:433–443
44. Baleja JD, Thanabal V, Wagner G (1997) J Biomol NMR 10:397–401
45. Leszczyszyn OI, Schmid R, Blindauer CA (2007) Proteins 68:922–935
46. Gardner KH, Anderson SF, Coleman JE (1995) Nat Struct Mol Biol 2:898–905
47. Marmorstein R, Harrison S (1994) Genes Dev 8:2504–2512