



UNIVERSITA' DEGLI STUDI DI VERONA

*DEPARTMENT OF*  
*Biotechnology*

*GRADUATE SCHOOL OF*  
*Natural Sciences and Engineering*

*DOCTORAL PROGRAM IN*  
*Applied Biotechnology*

XXXII Cycle

Plant calmodulin-like proteins: calcium binding and target interactions




S.S.D. BIO/10 BIOCHIMICA

Tutor: Prof. Paola Dominici  
Co-Tutor: Dr. Alessandra Astegno  
Coordinator: Prof. Matteo Ballottari

Doctoral Student: Matteo Trande

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License, Italy. To read a copy of the licence, visit the web page:

<http://creativecommons.org/licenses/by-nc-nd/3.0/>

-  **Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
-  **NonCommercial** — You may not use the material for commercial purposes.
-  **NoDerivatives** — If you remix, transform, or build upon the material, you may not distribute the modified material.

*Plant calmodulin-like proteins: calcium binding and target interactions*

Matteo Trande

PhD thesis

Verona, DD MONTH YEAR

ISBN XXXXXXXX



*To my Family*

# *Abstract*

University of Verona  
Ph.D. Course in Biotechnology, 32nd cycle

Plant calmodulin-like proteins: calcium binding and target interactions

by Matteo TRANDE

Calcium ( $\text{Ca}^{2+}$ ) is an ubiquitous key second messenger in all living organisms, from bacteria to differentiated neuronal cells and is involved in a large variety of biological processes such as apoptosis, cell differentiation and proliferation, fertilization, transcription factors activation and many others. Plants in particular are known to be capable of exploiting  $\text{Ca}^{2+}$  signalling in a more sophisticated way than other organisms as demonstrated by the larger set of  $\text{Ca}^{2+}$ -binding proteins (CBPs).

Although the study of CBP has been an open and highly studied field for over a century, a biochemical and structural characterization of many CBPs from different organisms is still missing.

Here, we provided a deep characterisation of two plant (*Arabidopsis thaliana*)  $\text{Ca}^{2+}$  sensors, a specific class of CBPs, to understand the differences in their ability to respond to  $\text{Ca}^{2+}$  and the regulation of downstream targets.

Thus we analysed the metal and target binding properties of Calmodulin like protein 7 and 19 (CML7 and CML19). By exploiting spectroscopic techniques (NMR and CD) we evaluated the conformational changes induced by metal binding and assigned the specific affinities for each binding site by a combination of calorimetric and spectroscopic techniques (ITC and NMR). To further understand the  $\text{Ca}^{2+}$  sensing properties and the ability to interact with target proteins, we analysed the interaction of CML7 and CML19 with synthetic peptides, representing the binding region on physiological or model targets.

Because the majority of CBPs are characterised by multiple metal binding sites, clustered affinities and cooperativity events, the determination of  $\text{Ca}^{2+}$ /CBPs stoichiometry requires time and sample consuming procedures and sophisticated instrumentations (i.e. ITC and NMR).



Thus, we report the optimisation of a new method, based on X-ray fluorescence, to overcome the most common limitations of the currently used techniques, allowing the quick estimation of the  $\text{Ca}^{2+}$ /protein stoichiometry using low amounts of sample.

Finally, initial studies on the recombinant production and characterisation of a neuronal family of CBPs are reported. These protein are called NECABs and are characterised by a peculiar domain composition and might be involved in the development of neurodegenerative diseases.

Università di Verona

# *Sommario*

Università di Verona  
Corso di Dottorato in Biotecnologie, 32esimo ciclo

Proteine calmodulin-like di pianta: legame al calcio e interazione  
con il target

di Matteo TRANDE

Il calcio ( $\text{Ca}^{2+}$ ) è un messaggero secondario ubiquitario e fondamentale in tutti gli organismi viventi, partendo dai batteri fino a cellule neuronali differenziate ed è coinvolto in una grande varietà di processi biologici, come ad esempio, differenziazione e proliferazione cellulare, fertilizzazione, attivazione di fattori di trascrizione e molti altri. Le piante in modo particolare sono capaci di sfruttare i segnali del  $\text{Ca}^{2+}$  in modo più sofisticato rispetto ad altri organismi, come dimostrato da un maggior numero di proteine leganti il calcio (PLC).

Sebbene lo studio delle PLC sia un acceso ambito di studio da più di un secolo, le informazioni biochimiche e strutturali di molte PLC di diversa origine sono ad oggi mancanti o incomplete.

In questa tesi riportiamo un'approfondita caratterizzazione di due  $\text{Ca}^{2+}$  sensori, una particolare classe di PLC, di pianta (*Arabidopsis thaliana*) al fine di meglio comprendere le differenze nella loro capacità di rispondere al  $\text{Ca}^{2+}$  e nella regolazione di proteine bersaglio.

Le proprietà di legame al metallo e ai bersagli sono state così studiate per le proteine Calmodulin like 7 e 19 (CML7 e CML19). Basandoci su tecniche spettroscopiche (NMR e CD) abbiamo valutato i cambiamenti conformazionali indotti dal legame al catione e assegnato le specifiche affinità per ogni sito di legame attraverso una combinazione di tecniche calorimetriche e spettroscopiche (ITC e NMR). Al fine di comprendere più nel dettaglio l'interazione con proteine bersaglio, abbiamo valutato la capacità di legame di CML7 e CML19 con peptidi sintetici, rappresentativi della regione di legame su bersagli fisiologici o modello.

Inoltre, poiché la maggior parte delle PLC sono caratterizzate dalla presenza di siti multipli di legame, affinità clusterizzate ed eventi di cooperatività, la determinazione della stechiometria  $\text{Ca}^{2+}$ /PLC richiede procedure preparative lunghe e dispendiose in termini di quantità di proteina, nonché l'utilizzo di strumentazioni complesse (ITC e NMR).

Perciò, si è ottimizzato un metodo, basato su fluorescenza a raggi X, allo scopo di superare le comuni limitazioni delle tecnologie attualmente in uso e permettere una rapida valutazione della stechiometria, impiegando limitate quantità di campione.

Infine, sono riportati anche studi preliminari di produzione ricombinante e caratterizzazione di una famiglia di PLC di origine neuronale umana. Tali proteine sono chiamate NECAB e sono caratterizzate dalla presenza di peculiari domini strutturali e potrebbero essere coinvolte nello sviluppo di malattie neurodegenerative.

# INDEX OF CONTENTS

## **Chapter 1 Introduction**

- 1.1 Ca<sup>2+</sup> homeostasis in cells and the model Ca<sup>2+</sup> sensor: Calmodulin 12
- 1.2 The EF-hand domain, a highly conserved Ca<sup>2+</sup>-binding motif 15

## **Chapter 2 Aim of the thesis** 19

## **Chapter 3 Methods**

- 3.1 DNA constructs 22
- 3.2 Proteins and peptides production 23
- 3.3 Nuclear Magnetic Resonance (NMR) spectroscopy 26
- 3.4 Calorimetric studies 27
- 3.5 Circular Dichroism (CD) spectroscopy 27
- 3.6 Fluorescence spectroscopy 28
- 3.7 Limited proteolysis 28
- 3.8 Electrophoretic mobility shift assays 29
- 3.9 Size Exclusion Chromatography (SEC) 29
- 3.10 Homology modelling 29
- 3.11 *A. thaliana* protoplast isolation and transformation 30
- 3.12 Confocal laser scanning microscopy 31
- 3.13 His-tag immunoblotting 31

## **Chapter 4 Calmodulin Like Protein 7 (CML7) from *Arabidopsis thaliana***

- 4.1 Introduction
  - 4.1.1 CaM-like proteins (CMLs) are a unique class of Ca<sup>2+</sup> sensors in plants 33
  - 4.1.2 Calmodulin Like Protein 7 (CML7) sequence analysis 36
- 4.2 Results and discussion
  - 4.2.1 CML7 production 38
  - 4.2.2 Apo-, Mg<sup>2+</sup>- and Ca<sup>2+</sup>-bound states of CML7 exist in distinct structural conformations 40
  - 4.2.3 CML7 has two high affinity Ca<sup>2+</sup> sites in the C-lobe and two weaker in the N-lobe 45

4.2.4 CML7 binds to the model target peptide melittin in a Ca <sup>2+</sup> -dependent manner	49
4.2.5 Homology modelling of CML7	52
4.2.6 CML7 subcellular localisation	53
4.3 Conclusions	54

## **Chapter 5 Calmodulin Like Protein 19 (CML19) from *Arabidopsis thaliana***

5.1 Introduction	
5.1.1 The role of Centrins in the Nucleotide Excision repair system (NER)	58
5.1.2 Centrins in <i>Arabidopsis thaliana</i> , the role of CML19	59
5.2 Results and discussion	
5.2.1 Ca <sup>2+</sup> -induced conformational changes	60
5.2.2 Ca <sup>2+</sup> binding to CML19	64
5.2.3 Interaction with RAD4 peptide	66
5.3 Conclusions	70

## **Chapter 6 A method for metal/protein stoichiometry determination**

### **Using Thin-Film Energy Dispersive X-ray Fluorescence**

6.1 Introduction	
6.1.1 The commonly used techniques for metallic cofactor identification and quantification and their limitations	73
6.1.2 XRF-based techniques	73
6.2 Specific methods	75
6.3 Results and discussion	
6.3.1 Experimental design and instrument calibration	77
6.3.2 Protein analysis	79
6.4 Conclusions	82

## **Chapter 7 Neuronal calcium binding proteins (NECABs)**

7.1 Introduction	
7.1.1 Neuronal Ca <sup>2+</sup> signalling	85
7.1.2 NECABs proteins	86
7.2 Results and discussion	

7.2.1 Production of recombinant NECABs	88
7.2.2 Does the C-terminal region of NECAB1 induce the formation of high order oligomers?	92
7.3 Conclusions	93
<b><u>Chapter 8 Final discussion</u></b>	96
<b><u>References</u></b>	98
<b><u>Appendices</u></b>	
Papers	107

# *Chapter 1*

## *Introduction*

## 1.1 Ca<sup>2+</sup> homeostasis in cells and the model Ca<sup>2+</sup> sensor: Calmodulin

Ionized Calcium (Ca<sup>2+</sup>) represents the most common signal transduction element in cells ranging from bacteria to differentiated neuronal cells, coupling extracellular signals to specific intracellular responses and regulating communications between cells. The versatility of Ca<sup>2+</sup> as second messenger is reflected by the large number of processes in which the cation is involved, such as apoptosis, cell differentiation and proliferation, fertilization, transcription factors activation and many others [Clapham 1995].

Despite the involvement into such diverse processes, the cellular mechanisms through which external stimuli are converted into biological responses in the Ca<sup>2+</sup> pathway are highly conserved. The cytosolic free Ca<sup>2+</sup> concentration in resting cells is maintained in the region of 10<sup>-7</sup> M and rises to 10 times upon external excitation; unlike other signal molecules Ca<sup>2+</sup> cannot be metabolised to restore resting condition; moreover, prolonged stimulation by high intracellular Ca<sup>2+</sup> levels (up to 10<sup>-4</sup> M) leads to cellular death [Edel 2017]. Thus evolution required the development of a tight regulation machinery in Ca<sup>2+</sup> homeostasis, to decode variations in terms of speed, amplitude and spatio-temporal patterning of Ca<sup>2+</sup> levels into appropriate cellular response. This sophisticated system emerges from the use of a large molecular repertoire of signalling components, known as the Ca<sup>2+</sup> signalling toolkit. (Figure 1) [Dong 2006].

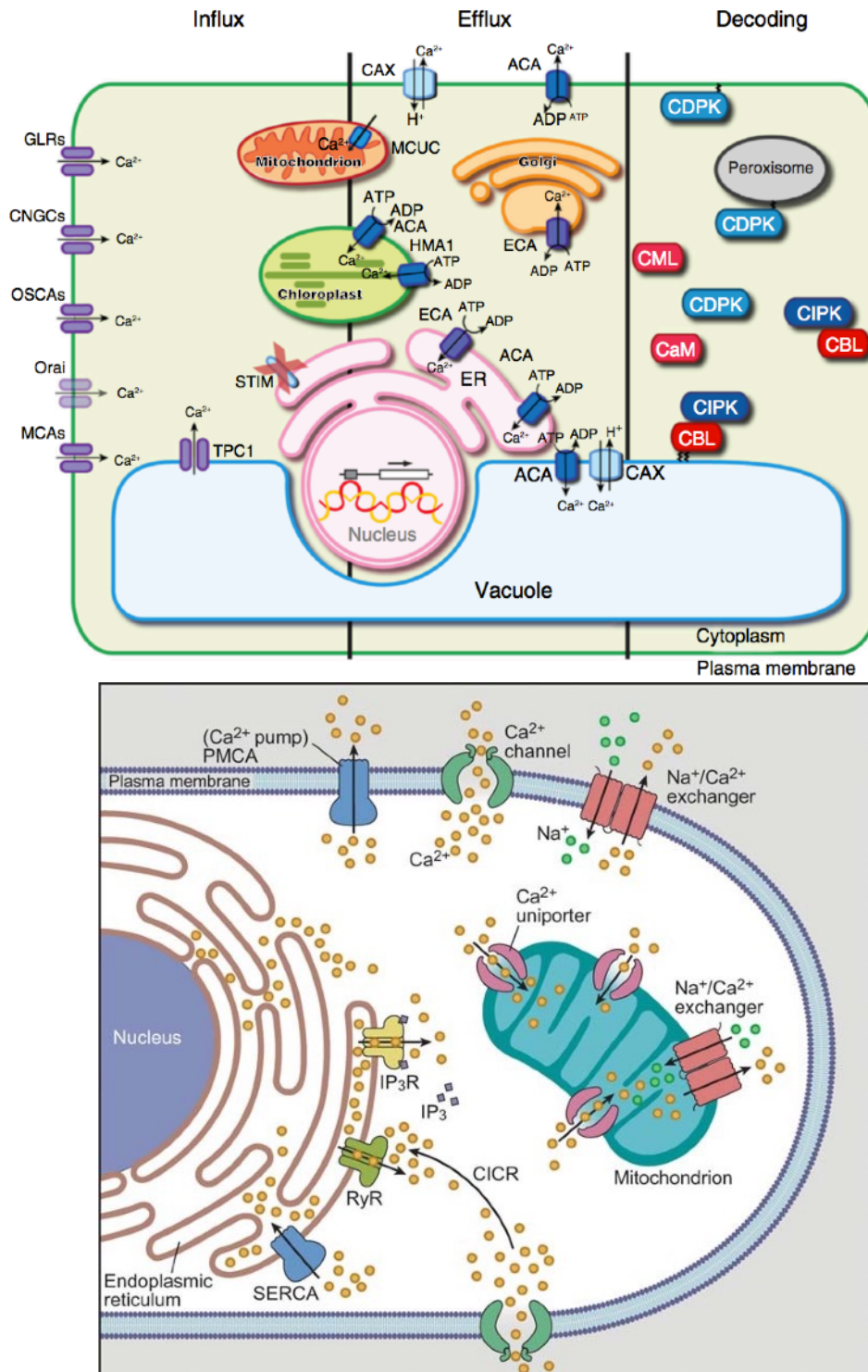
The generation of Ca<sup>2+</sup> signals, upon external stimuli, is achieved by the gradient dependent release of both external and internal Ca<sup>2+</sup> sources (ON mechanisms) such as the endoplasmic reticulum or, in the case of plants cells, the vacuole.

Ca<sup>2+</sup> influx is subsequently shaped and decoded by a large variety of Ca<sup>2+</sup> binding proteins (CBPs) which differ substantially among species. CBPs can generally be classified as either Ca<sup>2+</sup> buffer proteins, which help in controlling Ca<sup>2+</sup> levels and defining specific spatiotemporal distributions, or Ca<sup>2+</sup> sensor proteins which translate the chemical signal into diverse biological responses. Ca<sup>2+</sup> is then removed against gradient from the cytoplasm to the extracellular environment or back into subcellular compartments by various pumps and exchangers (OFF mechanisms) [Berridge 2000].

Ca<sup>2+</sup> sensors proteins represent the main subject of this PhD thesis. Among eukaryotes, calmodulin (CALcium MODULated proteIN) is a ubiquitous calcium



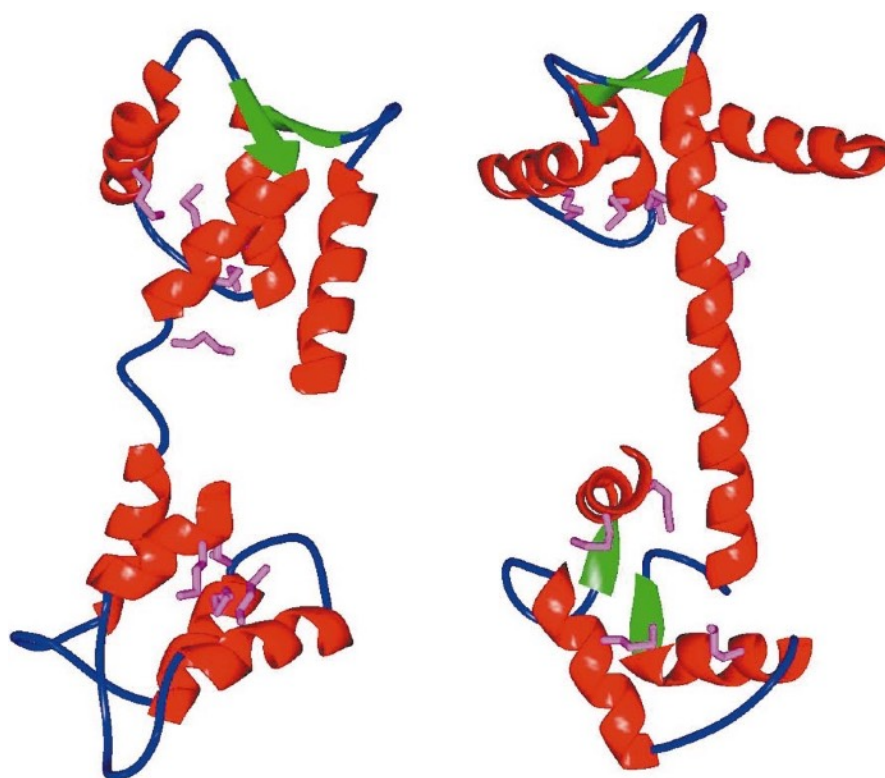
sensor). CaM is a small acidic protein that is characterised by the presence of four EF-hand domains and no other functional domains. The EF-hand domain is a highly



**Figure 1 [Edel 2017 and Dong 2006]:** Examples of the complexity and involved proteins in Ca<sup>2+</sup> homeostasis in plant (above) and normal human (below) cells.

conserved  $\text{Ca}^{2+}$ -binding domain that will be discussed in details in the following paragraph [Islam 2012].

The mechanism of action of CaM has been deeply studied among all species and it has been found that the protein is able to activate dozens of effector proteins such as kinases or phosphatases or membrane channels. As for the majority of  $\text{Ca}^{2+}$  sensors, the EF-hand domains of CaM are grouped into two pairs (lobes), separated by a central and flexible  $\alpha$ -helical structure. In the absence of  $\text{Ca}^{2+}$  ions, the N-terminal domain of the apo-CaM shows a “closed” conformation, whereas the C-terminal domain adopts a “semiopen” state in which a partially exposed hydrophobic region is accessible. This might allow the C-terminal domain of CaM to interact with some target proteins at resting levels of intracellular free  $\text{Ca}^{2+}$ . The binding of  $\text{Ca}^{2+}$  induces the change into a more “open” conformation, in which methionine rich hydrophobic regions of each domain are exposed to the solvent and available for target recognition [Chin 2000] (Figure 2).



**Figure 2 [Chin 2000]:**  $\text{Ca}^{2+}$ -induced conformational changes in calmodulin.  $\text{Ca}^{2+}$ -free (apo) CaM on the left and  $\text{Ca}^{2+}$ -bound (holo) on the right; N-terminal domains on top. Methionine side chains are shown in purple to denote the location of potential hydrophobic regions in each domain upon metal binding.

Other  $\text{Ca}^{2+}$  sensor proteins can be differently found in specific organisms:  $\text{Ca}^{2+}$ -dependent protein kinases (CDPKs), calcineurin B-like (CBL) proteins, and calmodulin-like proteins (CMLs) are specific of plants [Edel 2017], whereas Troponin C, parvalbumin, calcineurin and the S100 proteins are typical examples of mammals  $\text{Ca}^{2+}$ -sensors [Berridge 2000].

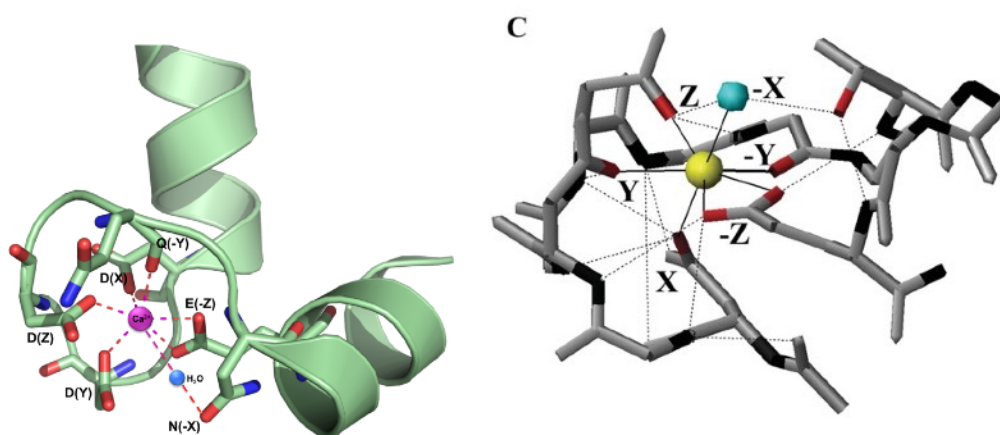
Although the large variety of  $\text{Ca}^{2+}$ -sensors in different organisms, the structural binding motif EF-hand has been evolutionary conserved among the majority of  $\text{Ca}^{2+}$  sensors, highlighting its importance and efficiency in  $\text{Ca}^{2+}$ -binding.

## 1.2 The EF-hand domain, a highly conserved $\text{Ca}^{2+}$ -binding motif

Nowadays, the EF-hand motif is among the five most common protein domains in animal cells, nevertheless, this domain has been identified in over a thousand proteins from different organisms, ranging from prokaryotes to eukaryotes throughout the entire plant, fungi and animal kingdoms. This highly conserved domain is composed by two  $\alpha$ -helices (E and F) connected by a 12 amino acid loop, that represents the necessary and sufficient  $\text{Ca}^{2+}$  coordination region of the domain. Indeed, it was demonstrated that the loop is able to bind the cation even in the absence of the helices E and F [Yiming 2004].

The chelation of  $\text{Ca}^{2+}$  within the loop requires the formation of seven coordinating groups, in which oxygen represents the atom of choice to provide the appropriate ionic interaction with the cation (Figure 3). In particular, three coordinating groups are represented by the oxygens in the side chains of the residues in position 1, 3 and 5 of the EF-hand loop [La Verde 2018]. The 4th oxygen required for the binding belongs to the carbonyl group of the amino acid backbone in position 7, whereas a water molecule forms a bridge between  $\text{Ca}^{2+}$  and the residue in the 9th position. Finally, position 12 is generally occupied by a glutamate due to its ability to form a double bidentate interaction with the chelated ion or in few cases (8% of the known EF-hand) by an aspartate which, however, interacts with  $\text{Ca}^{2+}$  only through a single electrostatic bound. Moreover, position 6, although not directly involved in the interaction, requires the presence of a Glycine (Gly6) to allow a  $90^\circ$  turn that enables the C-terminus ligands (residues 9 and 12) to embrace the correct position around  $\text{Ca}^{2+}$ . In addition, an extensive hydrogen bonding pattern, shown in Figure 3, contributes to the stabilisation of the protein-ion complex [Gifford 2007].

Variations in the above-mentioned sequence of the twelve residues in the EF-hand loop can be exploited to distinguish canonical and functional EF-hands from non-functional, non-canonical, pseudo EF-hand or EF-hand like loops [Zhou 2006]. When  $\text{Ca}^{2+}$  binding occurs, the ion is bound initially by residues 1, 3, 5 and 7 as this region can easily embrace the element due to its flexibility. Residues located at the C-



**Figure 3 [La Verde 2018 and Gifford 2007]:** **Left:** Coordination of  $\text{Ca}^{2+}$  in the canonical EF-hand loop 4 of *Arabidopsis thaliana* CaM7 (PDB: 5A2H). **Right:** Representation of the extensive stabilising hydrogen bonding network (dotted lines) found in the  $\text{Ca}^{2+}$  loaded loop.

terminal instead require a reorientation of the F helix by  $\sim 2\text{\AA}$  to fulfil the correct coordination; the movement of the helix causes the conformational changes observed in the  $\text{Ca}^{2+}$  bound state of CaM. The specificity of the EF-hand for  $\text{Ca}^{2+}$  over physiologically relevant monovalent cations such as  $\text{K}^+$  and  $\text{Na}^+$  with similar sizes is explained by its higher charge which can overcome the O-O repulsion within the loop [Reid 1980 and Marsden 1988]. On the other hand, CaM has been found to be capable of binding several divalent cations such as  $\text{Cd}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Mn}^{2+}$  etc., which however, do not represent physiological competitors in vivo due to their low abundance in cells [Ouyang 1998 and Habermann 1983]. A special mention should be given to  $\text{Mg}^{2+}$ , which can be found in resting cells at mM concentrations (0.5-2 mM in plant cells [Maathuis 2009] and 0.8-1.2 mM in mammal cells [Wolf 2003]). Therefore, CBPs must be able to discriminate  $\text{Ca}^{2+}$  against a higher excess of  $\text{Mg}^{2+}$ . Depending on the coordination geometry of  $\text{Mg}^{2+}$  into the EF-hand loop, the

existence of both  $\text{Ca}^{2+}$ -specific and  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -binding EF-hands has been described. In the first case,  $\text{Mg}^{2+}$  is chelated only by the N-terminal ligands of the loop (positions 1, 3, 5 and 7) and so induces none or limited effects on the conformational rearrangements. In contrast, in  $\text{Ca}^{2+}/\text{Mg}^{2+}$  sites the twelfth amino acid of the loop shows a  $120^\circ$  rotation of the  $\text{C}\alpha\text{-C}\beta$  that provides the bidentate ligand for  $\text{Ca}^{2+}$ , but can also allow a monodentate interaction with  $\text{Mg}^{2+}$  and thus activate the structural rearrangements essential in structural and/or physiological roles [Martin 1990].

The majority of the experiments in this PhD thesis has been performed by comparing the two cations or by evaluating the effect of  $\text{Mg}^{2+}$  on  $\text{Ca}^{2+}$  binding.

## *Chapter 2*

*Aim of the thesis*

This thesis examines EF-hand-containing CBPs from the biochemical and structural point of view focusing on: 1) the elucidation of their Ca<sup>2+</sup> sensing properties 2) the evaluation of the structural consequences of the metal binding 3) the structural details of CBP/target interaction.

Chapter 3, entitled “Methods”, provides the theoretical background of the experimental techniques exploited in this thesis, such as Nuclear Magnetic Resonance (NMR), Calorimetric techniques (ITC), Circular Dichroism (CD), Fluorescence spectroscopy, Electrophoretic mobility, Size Exclusion Chromatography (SEC), Homology modelling, Confocal laser scanning microscopy and immunoblotting; details on the purifications of the various proteins studied are also given.

Chapter 4 and 5 describe the characterization of metal and target binding properties of two plant calmodulin-like proteins (CMLs). In Chapter 4, we characterized CML7 from *Arabidopsis thaliana* in terms of Ca<sup>2+</sup> binding affinity, structural properties, and conformational perturbation upon metal (i.e., Ca<sup>2+</sup> and Mg<sup>2+</sup>) binding using spectroscopic and calorimetric techniques, and we determined its subcellular localization by confocal microscopy. Moreover, we studied the interaction of CML7 with the model CaM-binding peptide melittin. A similar approach was used in Chapter 5 to report an in-depth examination of the biophysical and structural properties of CML19 (the Centrin 2 of *A. thaliana*) by determining its detailed Ca<sup>2+</sup> binding properties. Moreover, we identified the specific CML19-binding site at the C-terminus of its physiological partner RAD4, the plant homologues of human XPC, and characterized the molecular association of a 17-mer peptide representing this site and CML19.

Chapter 6 contains the work of my abroad experience at the University of Manchester (special thanks to Prof. Lu Shin Wong and Silvia Fruncillo) and it describes the optimisation of a new method, based on X-ray fluorescence, in the determination of metal content in protein samples, overcoming many of the current limitations for the determination of metal/protein stoichiometry.

Finally, in Chapter 7 the preliminary analysis of a poorly characterised family of neuronal CBPs (NECABs) with an unusual domain composition is reported. Initial attempts to produce recombinant version of the three NECAB isoforms are reported; moreover, a preliminary biochemical characterization of isoform 1 is described.

Globally, these analyses have allowed significant steps forward in appreciating how such a vast array of CBP proteins can coexist, without apparent redundancy, and how they make a distinct contribution to cellular signaling, while being different from CaM and other Ca<sup>2+</sup> sensors and, as such, have both physiological and technical significance.



# *Chapter 3*

## *Methods*

### 3.1 DNA constructs

The cDNA sequence for CML7 (At1g05990) in Chapters 4 and 6 was purchased from the Arabidopsis Information Resource (TAIR) into a pUNI51 vector. The coding sequence was polymerase chain reaction (PCR)-amplified using the primers in table 1 carrying or not an N-terminal six histidine tag. The forward primers carried a NdeI restriction site (underlined), while the reverse primer contained a stop codon (**bold**) and a BamHI restriction site (underlined). For protein expression, the cDNA sequence was further cloned using standard laboratory protocols into pET21a vector using NdeI/BamHI restriction sites. For CML7 subcellular localization studies, (Chapter 4) the complete coding sequence (CDS) of the red fluorescent protein (RFP) was amplified by PCR using Phusion® High-Fidelity DNA Polymerase (New England Biolabs, USA) according to the manufacturer's protocol. The primers reported in the table were used to add a NotI site (underlined) to both ends of the PCR product. Following NotI digestion, RFP was inserted into the polylinker of the plant expression vector pGREEN0029::ter [Bonza 2013] under the control of the cauliflower mosaic virus (CaMV) 35S promoter (pGREEN0029::RFP). The complete CDS of CML7 was PCR amplified (see Table 1) to add XhoI and EcoRI restriction sites (underlined) at 5' and 3' respectively during amplification. The reverse primer also eliminates the stop codon at the 3' of CML7 CDS. Following enzymatic digestion, CML7 CDS was inserted into pGREEN0029::RFP for the in frame fusion at the C-terminus with RFP.

CML19 was produced in the laboratory before the beginning of my PhD as truncated variant (hereafter referred to as CML19) [La Verde (bis) 2018] missing the first 19 amino acids. The N-terminal region was removed as it was predicted to be intrinsically disordered by GeneSilico MetaDisorder Service analysis and thus to avoid experimental difficulties associated with the aggregation propensity of recombinant full length CML19, which was observed also for other centris [Thompson 2006 and Charbonnier 2007].

The cDNA sequence for NECAB1 and 2, Chapter 7, was purchased from DNAsu, in pENTR223 and pMK respectively and were amplified by PCR in order to add and NdeI recognition site in the forward primer and a XhoI site in the reverse (underlined); moreover the two proteins were his-tagged at the C-terminal and N-



500 mM imidazole was then applied and the fractions containing CML7 protein were pooled and extensively dialyzed against 50 mM Tris-HCl, 150 mM KCl, 0.5 mM DTT at pH 7.5. The final yield was about 5 mg/L culture.

In the case of non-tagged CML7, after removing cell debris by centrifugation, the supernatant was subject to heat treatment (6 min at 85 °C followed by 40 min on ice). The precipitate was removed by centrifugation (20000g for 40 min at 4°C) and the heat-treated lysate containing 5 mM CaCl<sub>2</sub> was loaded onto a phenyl-Sepharose column, previously equilibrated with 50 mM Tris-HCl, 0.5 mM CaCl<sub>2</sub>, 0.5 mM DTT pH 7.5 buffer. The column was then washed with the same buffer containing 150 mM KCl. CML7 was eluted in 50 mM Tris-HCl, 150 mM KCl pH 7.5, 2 mM ethylene glycol-bis(2-aminoethyl ether)- N,N,N',N'-tetraacetic acid (EGTA) buffer, which acts as a chelator of several cations. The purity of CML7 was estimated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis and the protein concentration was determined by the Bradford assay [Bradford 1976]. Final yield was about 20 mg/L culture.

The expression of CML19 truncated variant, studied in Chapter 5, was induced for 18 h at 25 °C by the addition of 0.4 mM IPTG to *E.coli* Rosetta DE3 cells. The bacterial pellet obtained by centrifugation (5000g for 15 min) of 1-L culture was resuspended in 5 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM imidazole containing 0.1 mg/mL lysozyme, 100 µg/mL DNAase (Sigma), 10 mM MgCl<sub>2</sub>, stirred at room temperature for 30 min, and then sonicated on ice. After centrifugation, the supernatant was loaded into a Ni-sepharose column previously pre-equilibrated with 5 mM Tris-HCl, 150 mM KCl and 10 mM imidazole, pH 7.5 buffer. A linear gradient from 10 to 500 mM imidazole was applied and CML19 protein was pooled and dialyzed against 5 mM Tris-HCl, 150 mM KCl, at pH 7.5 Protein concentration was estimated by Bradford assay, leading to a final yield of 25 mg/L culture.

<sup>15</sup>N and <sup>13</sup>C labeled samples for NMR spectroscopy in Chapters 4 and 5 were obtained using M9 minimal medium supplemented with 1g/L <sup>15</sup>NH<sub>4</sub>Cl and 4 g/L <sup>13</sup>C-glucose as the only sources of nitrogen and carbon. In M9 medium the final yield was about half compared to non labelled proteins.

Synthetic Melittin was obtained from Sigma ( $\geq 97\%$ , HPLC); whereas RAD4 peptide was synthesized by GenScript USA Inc. ( $> 90\%$ , HPLC). Peptide concentrations were determined using their predicted molar extinction coefficient ( $\epsilon_{280} = 5500$  and  $6990 \text{ M}^{-1} \text{ cm}^{-1}$  respectively, <http://web.expasy.org/protparam/>).

AtCaM1 from *A. thaliana* used in Chapters 4 and 6 was produced as described [Gut 2009]. Proteins specifically used in Chapter 6: Bilirubin oxidase from *Myrothecium verrucaria* (MvBOx); purchased from Amano Enzyme (Chipping Norton, UK). Horseradish peroxidase from *Armoracia rusticana* (HRP) and cytochrome from *Equus caballus* (CytC); purchased from Sigma Aldrich. Alcohol dehydrogenase 1 from *Saccharomyces cerevisiae* (ADH1); purchased from ThermoFisher Scientific (Loughborough, UK). Hemoglobin 1 from *Arabidopsis thaliana* (AHb1) and spore coat protein A from *Bacillus subtilis* CotA were recombinantly produced according to [Kumar 2016 and Duroo 2008]. Flavin prenyltransferase from *Escherichia coli* (UbiX) was donated by Prof. David Leys (University of Manchester).

NECAB1 and NECAB2DN56, in Chapter 7, were expressed in *E. coli* strain BL21(DE3) Codon Plus (Stratagene) induced with  $0.4 \text{ mM}$  IPTG to  $1 \text{ L}$  of bacterial culture grown in LB medium and incubating at  $24 \text{ }^\circ\text{C}$  for  $16 \text{ h}$ . The bacterial pellet was resuspended in  $50 \text{ mM}$  Tris-HCl,  $150 \text{ mM}$  KCl,  $0.1 \text{ mM}$  DTT and  $10 \text{ mM}$  imidazole, pH 7.5 and added with  $100 \text{ } \mu\text{g}/\text{mL}$  lysozyme,  $100 \text{ } \mu\text{g}/\text{mL}$  DNAase (Sigma),  $10 \text{ mM}$   $\text{MgCl}_2$ , stirred at room temperature for  $30 \text{ min}$ . Cell lysis was performed by sonication on ice and, after centrifugation, the supernatant was loaded into a Ni-sepharose column pre-equilibrated with  $5 \text{ mM}$  Tris-HCl,  $150 \text{ mM}$  KCl and  $10 \text{ mM}$  imidazole,  $0.1 \text{ mM}$  DTT, pH 7.5 buffer. Protein elution was achieved by a linear gradient from  $10$  to  $500 \text{ mM}$  imidazole. Purified proteins were pooled and dialyzed against  $50 \text{ mM}$  Tris-HCl,  $150 \text{ mM}$  KCl,  $0.1 \text{ mM}$  DTT, pH 7.5 buffer. Protein concentration was determined via their predicted molar extinction coefficient ( $\epsilon_{280} = 44983$  and  $36628 \text{ M}^{-1} \text{ cm}^{-1}$  respectively, <http://web.expasy.org/protparam/>) and led to a final yield of  $40$  and  $2 \text{ mg}/\text{L}$  respectively.

### 3.3 Nuclear Magnetic Resonance (NMR) spectroscopy

All NMR experiments in Chapters 4 and 5 were performed on a Bruker Avance III spectrometer (Bruker, Karlsruhe, Germany), operating at 600.13 MHz proton Larmor frequency, and equipped with a cryogenic probe.

$^1\text{H}$ - $^{15}\text{N}$  Heteronuclear Single Quantum Coherence (HSQC) experiments were recorded at 298 K on a  $^{15}\text{N}$  labeled sample, in the protein buffer (see Chapter 3.2) added with 7%  $\text{D}_2\text{O}$ . Protein was analyzed at 0.27 mM or 0.5 mM (CML7 and CML19 respectively) in the presence of EGTA (5 mM), or  $\text{MgCl}_2$  (5 mM EGTA + 10 mM  $\text{MgCl}_2$ ) or  $\text{CaCl}_2$  (5 mM).

For Chapter 4 experiments; the data matrix consisted of 2048 ( $\text{F}_2$ ,  $^1\text{H}$ )  $\times$  256 ( $\text{F}_1$ ,  $^{15}\text{N}$ ) complex data points, spectral windows of 8417.509 Hz ( $^1\text{H}$ )  $\times$  2189.445 Hz ( $^{15}\text{N}$ ), 8 transients, and 1.5 s relaxation delay. Standard triple resonance experiments, HNCACB, CBCA(CO)NH, and  $^1\text{H}$ - $^{15}\text{N}$  Nuclear Overhauser Effect Spectroscopy (NOESY) HSQC were recorded at 298 K on  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled samples at protein concentration of 500  $\mu\text{M}$  to achieve partial sequence-specific backbone atoms assignment of CML7 bound to  $\text{Ca}^{2+}$ . The HNCACB, and CBCA(CO)NH spectra were acquired with a data matrix consisting of 2 K ( $\text{F}_3$ ,  $^1\text{H}$ )  $\times$  90 ( $\text{F}_2$ ,  $^{15}\text{N}$ )  $\times$  120 ( $\text{F}_1$ ,  $^{13}\text{C}$ ) complex points, spectral windows of 8417.51 Hz ( $^1\text{H}$ )  $\times$  2189.42 Hz ( $^{15}\text{N}$ )  $\times$  11,318.16 Hz ( $^{13}\text{C}$ ), 8 transients, and 1.5 s relaxation delay. The  $^1\text{H}$ - $^{15}\text{N}$  NOESY HSQC spectrum was acquired with a data matrix consisting of 2K ( $\text{F}_3$ ,  $^1\text{H}$ )  $\times$  40 ( $\text{F}_2$ ,  $^{15}\text{N}$ )  $\times$  128 ( $\text{F}_1$ ,  $^1\text{H}$ ) complex points, spectral windows of 8417.51 Hz ( $^1\text{H}$ )  $\times$  2189.42 Hz ( $^{15}\text{N}$ )  $\times$  8417.51 Hz ( $^1\text{H}$ ), 16 transients, and 120 ms mixing time.

In Chapter 5 experiments a standard pulse sequence was used to record  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single-quantum coherence (HSQC) spectra, with pulsed field gradients for cancellation of artifacts and for the solvent signal suppression.  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were acquired with a data matrix consisting of 2 K ( $\text{F}_2$ ,  $^1\text{H}$ )  $\times$  256 ( $\text{F}_1$ ,  $^{15}\text{N}$ ) complex points, spectral windows of 8417.51 Hz ( $^1\text{H}$ )  $\times$  2128.63 Hz ( $^{15}\text{N}$ ), 8 transients, and 1.2 s relaxation delay. Standard triple resonance experiments, HNCA, HN(CO)CA, HNCACB and CBCA(CO)NH were recorded on  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled samples to achieve partial sequence-specific backbone atoms assignment of 0.5 mM CML19 bound to  $\text{Ca}^{2+}$ . The spectra were acquired with a data matrix consisting of 2 K ( $\text{F}_3$ ,  $^1\text{H}$ )  $\times$  90 ( $\text{F}_2$ ,  $^{15}\text{N}$ )  $\times$  120 ( $\text{F}_1$ ,  $^{13}\text{C}$ ) complex points, spectral windows of 8417.51 Hz ( $^1\text{H}$ )  $\times$  2128.63 Hz ( $^{15}\text{N}$ )  $\times$  11318.16 Hz ( $^{13}\text{C}$ ), 8 transients, and 1.5 s

relaxation delay. All spectra were processed using Topspin3.2 (Bruker, Karlsruhe, Germany).

### **3.4 Calorimetric studies**

Isothermal titration calorimetry (ITC) studies in Chapters 4 and 5 were performed on a TA Instrument nano-calorimeter at 25 °C with 250 rpm stirring speed and injections at 300s intervals.

In the analysis of CML7 metal binding properties 1.5  $\mu\text{L}$  injection-volume of 3 mM  $\text{CaCl}_2$  were made to a 200  $\mu\text{L}$  sample cell containing 200  $\mu\text{M}$  apo- or  $\text{Mg}^{2+}$ -bound CML7; whereas apo- or  $\text{Mg}^{2+}$ -bound CML19 at 120  $\mu\text{M}$  was titrated with 1  $\mu\text{L}$  injections of 4 mM  $\text{CaCl}_2$ .  $\text{Mg}^{2+}$ -saturated samples were obtained by incubation of apo-proteins with 10 mM  $\text{MgCl}_2$  for 15 min at 25 °C. Decalcified buffers were prepared using a Chelex-100 resin column. The residual buffer  $\text{Ca}^{2+}$  concentration was measured by BAPTA absorption assay according to manufacturer's instructions and ranged between 0.8 and 1.3  $\mu\text{M}$ ; all solutions were degassed prior to each experiment. In order to remove contaminating  $\text{Ca}^{2+}$ , the syringes and ITC sample cell were soaked with EDTA and washed with decalcified buffer before sample loading. Control experiments (injecting ligand solutions into the buffer) were conducted to evaluate the heat of dilution which was subtracted from the raw heat in the presence of protein. Data analysis were performed using Origin-based software and fitted using the sequential binding site model. The reported values represent the mean  $\pm$  S.E.M. of at least three independent titrations using two different proteins preparations.

### **3.5 Circular Dichroism (CD) spectroscopy**

All CD spectra were recorded on a Jasco J-1500 spectropolarimeter equipped with a thermostatically controlled Peltier type compartment. Five scans averaged spectra were recorded between 200 and 250 nm (far UV), or between 320 and 250 nm (near UV), using a 0.1 cm or 1 cm quartz cell respectively. Spectra were corrected for background signal by subtracting the spectrum of the buffer. 100  $\mu\text{M}$  (near UV) or 12  $\mu\text{M}$  (far UV) proteins samples, dissolved in the protein buffer, were analyzed at 25 °C after consecutive additions of 1 mM EGTA, 2 mM  $\text{MgCl}_2$  and 2 mM  $\text{CaCl}_2$ .

Spectra of protein/peptide complexes were obtained using a protein:peptide ratio of 1:1 in the presence of 2 mM EGTA or 2 mM  $\text{CaCl}_2$ .

Denaturation profiles were recorded in the same conditions as for far UV spectra by following CD signal at 222 nm between 4 and 96 °C (scan rate 90 °C/h).

### 3.6 Fluorescence spectroscopy

Fluorescence spectra were recorded on a Jasco FP8200 spectrofluorometer. 1-anilino-8-naphthalene sulfonate (ANS) fluorescence emission was used to evaluate the hydrophobicity of CML7 and CML19 upon Mg<sup>2+</sup> and/or Ca<sup>2+</sup> binding. Proteins were analysed at 1 μM final concentration and incubated 1 hour at room temperature with 15 μM ANS in the presence of 5 mM CaCl<sub>2</sub>, 5 mM EGTA, or 5 mM MgCl<sub>2</sub> + 0.5 mM EGTA. Spectra were recorded between 400 and 650 nm with excitation at 380 nm at room temperature with slit widths set to 5 mm.

For the CML7/melittin, AtCaM1/melittin and CML19/RAD4p interactions the single Trp residue of melittin or RAD4p was excited at 295 nm, and fluorescence emission spectra were recorded over the range 305–450 nm. RAD4p (1 μM) and Melittin (2 μM) were titrated with increasing concentration of the tested protein in the presence of 5 mM CaCl<sub>2</sub> or 5 mM EGTA. Titration experiments were performed by monitoring the blue-shift in wavelength ( $\lambda$ ) of the peptide emission peak and K<sub>d</sub> values were calculated according to a tight binding hypothesis as described [Astegno 2013 and 2017].

Guanidine-induced denaturation experiments were performed by monitoring the emission spectra from 290 to 550 nm upon excitation of both Tyr and Trp residues at 280 nm. Maximum emission peak wavelength ( $\lambda_{\text{max}}$ ) was plot against guanidine concentration [GdnHCl] and fit assuming a 2 state model. Samples were prepared using 1 μM protein concentration (NECAB1) and increasing concentration of guanidine chloride from 0 to 6 M, in the presence of 5 mM CaCl<sub>2</sub> or EGTA and incubated at room temperature overnight.

### 3.7 Limited proteolysis

0.5 mg/mL CML7, CML19 or NECAB1 were digested with trypsin (1:500 w/w) in the presence of 5 mM CaCl<sub>2</sub> or 5 mM EGTA at 25 °C. The reaction was followed on SDS-PAGE by withdrawing a 15 μL aliquot at various times (0, 1, 5, 10, 20, 40, 60, 120, and 180 min) and stopping the reaction by boiling for 3 min in reducing Laemmli buffer. Bands in denaturing gel were visualized by Coomassie Blue staining.



In Chapter 7, In order to isolate the proteolytic fragment, 200  $\mu\text{g}$  of NECAB1 were incubated with trypsin for 30 min. The reaction was stopped by adding a two-fold weight excess of soybean trypsin inhibitor and the proteolysis product was purified on a Superdex 200 HR 10/30 column.

### 3.8 Electrophoretic mobility shift assays

The  $\text{Ca}^{2+}$ -induced mobility shift of the various CBPs were assayed on a 15% (CML7 and CML19) or 12% (NECAB1 and NECAB2DN56) SDS-PAGE continuous gel after incubating 2  $\mu\text{g}$  of protein with 5 mM  $\text{CaCl}_2$  or 5 mM EGTA for 30 min at room temperature. CML7 was also tested in the presence or absence of reducing conditions (addition of DTT) to evaluate the formation of nonspecific disulphide bridges. Peptide-binding stoichiometry were analyzed by incubating 20  $\mu\text{M}$  of protein with increasing concentration of the tested peptides for 20 min at room temperature in 100 mM Tris-HCl pH 7.5 saturated with 5 mM  $\text{CaCl}_2$  or EGTA; samples were then separated under native conditions on a 12% continuous gel.

### 3.9 Size Exclusion Chromatography (SEC)

The Stokes radius ( $R_s$ ) of CML7, CML19 and NECAB1 in the presence of 5 mM EGTA, 5 mM  $\text{Mg}^{2+}$  (+ 0.5 mM EGTA) or 5 mM  $\text{Ca}^{2+}$  was estimated by size exclusion chromatography using a Superose 12 column (10/300 GL, GE Healthcare, Uppsala, Sweden) for CMLs and a Superdex 200 (10/300 GL, GE Healthcare, Uppsala, Sweden) for NECAB1 as described in [La Verde 2017]. The column was calibrated by running a combination of at least six standard proteins: thyroglobulin (669 kDa,  $R_s$  86 Å); apo-ferritin (443 kDa;  $R_s$  61 Å); bovine serum albumin (66 kDa,  $R_s$  35.5 Å), ovalbumin (43 kDa,  $R_s$  30.2 Å), carbonic anhydrase (29 kDa,  $R_s$  23.6 Å), myoglobin (17.8 kDa,  $R_s$  20.2 Å), and cytochrome c (12.4 kDa,  $R_s$  17 Å). Elution volume ( $V_e$ ) of the standards, the void volume and the total volume of the column were measured and  $R_s$  were determined from a calibration plot of  $\log R_s$  versus the partition coefficient  $K_{AV}$ . Reported values represent the means  $\pm$  S.E.M. of triplicate experiments.

### 3.10 Homology modelling

Homology model of the  $\text{Ca}^{2+}$ -bound CML7 from *A. thaliana* (Chapter 4) was generated using SWISS-MODEL Homology Modelling server

(swissmodel.expasy.org) [Arnold 2006 and Benkert 2011]. The template chosen for modelling procedure was the CaM from *Paramecium tetraurelia* (PDB entry: 1clm [Rao 1993]) which shares with CML7 the higher sequence identity (42.0%) among proteins whose structures have been solved so far. The Global Model Quality Estimation (GMQE) index was 0.77, and the Qualitative Model Energy Analysis (QMEAN) potential scoring function was 1.21, indicating a good reliability for the obtained model. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualisation and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) (<http://www.cgl.ucsf.edu/chimera>). Molecular surface was coloured with the Coulombic Surface colouring (colouring parameters:  $\mu = 4$ , thresholds  $\pm 12$  kcal mol<sup>-1</sup> at 298 K) implemented in Chimera (<http://www.cgl.ucsf.edu/chimera>).

### **3.11 A. thaliana protoplast isolation and transformation**

Mesophyll protoplasts were isolated and transiently transformed according to [Yoo, 2007]. Leaves from four week old *A. thaliana* Col-0 plants were collected and cut in 0.5–1 mm strips with fresh razor blades and put into multiwell cell culture in the presence of the following enzymatic plasmolysis mix: 20 mM KCl, 0.4 M D-Mannitol, 20 mM 2-(N-Morpholino) ethanesulfonic acid, 4-Morpholineethanesulfonic acid (MES), 1.25% R10 Cellulase, 0.3% R10 Macerozyme. Enzymes were pre-heated at 55°C for 10 min and 10 mM CaCl<sub>2</sub> and 0.1% bovine serum albumin (BSA) were added to the mix after the incubation. The enzymatic mix was filtered and added to the cut leaves. A 10 min vacuum favoured enzymes entering the leaves through the cuticle. Digestion was performed in the dark for about 3 h without shaking. Protoplasts were separated from the cell debris by filtering the mixture through a nylon sieve of 50  $\mu$ m pore size. Protoplasts were spun at 100g for 6 min in presence of 50 mM CaCl<sub>2</sub>, washed once in cold W5 washing solution (154 mM NaCl, 125 mM CaCl<sub>2</sub>, 5 mM KCl, 2 mM MES) and resuspended in the same solution. After incubation on ice for 30 min, protoplasts were resuspended in 600  $\mu$ L MMg (0.4 M D-Mannitol, 15 mM MgCl<sub>2</sub>, 4 mM MES pH 5.7) and centrifuged at 100g for 1 min. For the co-transformation with the free enhanced green fluorescent protein (EGFP), the DNA constructs (5–10  $\mu$ g of pGREEN0029::CML7-RFP and pGREAT::eGFP; [Teardo, 2017]) were mixed together before adding them to 100  $\mu$ L MMg isolated protoplasts and then an equal

amount of polyethylene glycol (PEG) solution (40% v/v PEG4000, 0.2 M D-Mannitol, 1mM CaCl<sub>2</sub>) was added, mixed gently and incubated at room temperature for 20 min. The protoplasts suspension was then diluted by adding a 4-fold volume of W5 to remove PEG. The supernatant was carefully removed after 3min centrifugation at 100g and the transformed protoplasts resuspended in 1 mL W5 solution and maintained at 23 °C in the dark. Protoplasts were microscopically analyzed 16–24 h after transformation.

### **3.12 Confocal laser scanning microscopy**

Confocal microscopy analysis was performed using a Nikon Eclipse Ti2 inverted microscope, equipped with a Nikon A1R+ laser scanning device (Nikon, <http://www.nikon.com/products/microscope-solutions/lineup/confocal/a1/index.htm>). EGFP and chlorophyll were excited with the 488 nm laser and the emission was collected at 500–550 nm and 663–738 nm, respectively. RFP was excited with the 561 nm laser and the emission was collected at 570–620 nm. The pin-hole was set to 1.2 airy unit. Images were acquired by a CFI Plan Apo Lambda 60X Oil (N.A. 1.4) objective and were analyzed using Fiji software (<https://fiji.sc>).

### **3.13 His-tag immunoblotting**

Samples were subjected to SDS-PAGE (12%) and transferred to nitrocellulose paper in transfer buffer at 70 V for 2 h. Non-specific protein binding sites were blocked by incubating the membrane for 1h in PBS buffer containing 5% (w/v) non-fat dried milk. Histidine tagged proteins were then revealed by incubating with monoclonal anti-polyHistidine peroxidase conjugate antibody (Sigma, Catalog number: A7058) according to manufacturer's instructions.

# Chapter 4

*Calmodulin like protein 7 (CML7)*  
*from Arabidopsis thaliana*

## 4.1 INTRODUCTION

### **4.1.1 CaM-like proteins (CMLs) are a unique class of Ca<sup>2+</sup> sensors in plants**

Ca<sup>2+</sup> sensor proteins in plants can be classified into three major groups: 1) Ca<sup>2+</sup>-dependent protein kinases (CDPKs), 2) calcineurin B-like (CBL) proteins and 3) calmodulins and calmodulin-like proteins (CaMs/CMLs).

It has been found that plants genome encode for a high number of EF-hand contains proteins (e.g. 250 in rice and *Arabidopsis thaliana* versus 132 in *Drosophila melanogaster* and 83 in human), which reflects the importance of these proteins in plant growth and development processes [Perochon 2011]. CMLs, which are the main subject of this thesis, are uniquely found in the plant kingdom. It has been proposed that CMLs expansion was driven by selective pressures to process environmental signals as plants adapted to land environments during evolution [Zhu 2015].

The number of predicted CMLs in different plant species can vary substantially, for example: 19 CMLs have been found in *Lotus japonicus* [Liao 2017], 32 in *Oryza sativa* [Boonburapong 2007], 36 in woodland strawberry [Zhang 2016], 52 in tomato [Munir 2016] and 79 in Chinese cabbage [Nie 2017].

In *Arabidopsis*, a well-known model plant, seven genes are believed to encode for four different CaMs, whereas 50 CML encoding genes have been identified. *Arabidopsis* CML proteins differ between each other in terms of length (from 83 to 330 amino acids), sequence identity with the canonical CaM2 (between 16% and 75%) and are characterised by possessing exclusively EF-hand domains (from 1 to 6) [Bender 2013].

*Arabidopsis* CaMs and CMLs have been previously divided into nine groups by [McCormack 2003] depending on their sequence similarity with the canonical CaM2; subsequently a more accurate division was made by Zhu and colleagues [Zhu 2015] into eleven different groups, taking into account also the number of predicted EF-hand domains and the protein length as illustrated in Table 2.

Various approaches aimed at understanding the CMLs function are beginning to reveal that these proteins are not likely to have redundant functions, but rather play central and highly specific roles in plant development as well as in response to biotic and abiotic stress.

Name	Accession number	Number of amino acids	Predicted EF-hand	% identity with AtCaM2	Group (Zhu 2015)	Group (McCormack 2003)
CaM1	At5g37780	149	4	97	I a	I
CaM2	At2g41110	149	4	100	I a	I
CaM3	At3g56800	149	4	100	I a	I
CaM4	At1g66410	149	4	97	I a	I
CaM5	At2g27030	149	4	100	I a	I
CaM6	At5g21274	149	4	99	I a	I
CaM7	At3g43810	149	4	99	I a	I
CML1	At3g59450	148	1	22	IV	VI
CML2	At4g12860	152	4	38	VII	VI
CML3	At3g07490	153	4	40	VII	VI
CML4	At3g59440	195	4	40	VII	VI
CML5	At2g43290	215	4	42	VII	VI
CML6	At4g03290	154	4	44	VII	VI
CML7	At1g05990	150	4	44	VII	VI
CML8	At4g14640	151	4	73	I b	II
CML9	At3g51920	151	4	50	I b	II
CML10	At2g41090	191	4	65	I c	II
CML11	At3g22930	173	4	75	I b	II
CML12	At2g41100	324	6	62	I c	II
CML13	At1g12310	148	3	50	I c	III
CML14	At1g62820	148	3	50	I c	III
CML15	At1g18530	157	4	40	II	IV
CML16	At3g25600	161	4	40	II	IV
CML17	At1g32250	166	4	44	II	IV
CML18	At3g03000	165	4	43	II	IV
CML19	At4g37010	167	4	42	III	V
CML20	At3g50360	169	4	45	III	V
CML21	At4g26470	248	4	28	VIII	V
CML22	At3g24110	229	4	24	VIII	V
CML23	At1g66400	157	4	41	V	VIII
CML24	At5g37770	161	4	40	V	VIII
CML25	At1g24620	186	4	44	V	VIII
CML26	At1g73630	163	4	38	V	VIII
CML27	At1g18210	170	4	40	V	VIII
CML28	At3g03430	83	2	35	VIII	VIII
CML29	At5g17480	83	2	33	VIII	VIII
CML30	At2g15680	187	4	35	V	VIII
CML31	At2g36180	144	4	38	V	VIII
CML32	At5g17470	146	4	33	V	VIII
CML33	At3g03400	137	4	37	V	VIII
CML34	At3g03410	131	4	36	V	VIII
CML35	At2g41410	216	4	34	V	IX
CML36	At3g10190	209	4	37	V	IX
CML37	At5g42380	185	4	34	VI	VI
CML38	At1g76650	177	4	29	VI	VI
CML39	At1g76640	159	4	26	VI	VI
CML40	At3g01830	146	2	23	VI	VI
CML41	At3g50770	205	3	36	VI	VI
CML42	At4g20780	191	3	35	VII	VII
CML43	At5g44460	181	3	34	VII	VII
CML44	At1g21550	155	3	30	VII	VII
CML45	At3g29000	194	3	30	IV	VII
CML46	At5g39670	204	3	29	IV	VII
CML47	At3g47480	183	2	30	IV	VII
CML48	At2g27480	186	2	16	VIII	VII
CML49	At3g10300	330	2	23	VIII	VII
CML50	At5g04170	315	2	23	VIII	VII

**Table 2:** *Arabidopsis thaliana* CaMs and CMLs groups.

Different stimuli can activate a different set or specific CMLs and thus induce the activation of the proper cellular response: abiotic stresses such as cold and salt treatment induce the activation of CML24 and CML8 respectively [Polisensky 1996; Delk 2005 and Park 2010], moreover, CML18 affects the  $\text{Na}^+/\text{K}^+$  selectivity of the  $\text{Na}^+/\text{H}^+$  exchanger 1 (AtNHX1) reducing the  $\text{Na}^+/\text{H}^+$  exchange speed, suggesting the presence of a  $\text{Ca}^{2+}$ -pH-dependent signaling involved in mediating plant responses to salt stress [Yamaguchi 2005], CML19 (that will be further discuss in the following Chapter) takes part in the DNA repair system induced by UV treatment [Liang 2006]. Response to biotic stresses are also decoded by different CMLs as observed by the role of CMLs 42 and 37 in activating herbivore resistance [Vadassery 2012 and Scholz 2014] or by the accelerated hypersensitive response (HR, a localized programmed cell death at tissue surrounding the infection site that limits pathogen invasion) produced by CML43, following *Pseudomonas syringae* infection [Chiasson 2005]. In addition, mutations of CML24 were also found to inhibit pathogen-induced NO generation and the HR [Ma 2008]. CML9 was observed to be involved in both abiotic and biotic stresses, playing important roles in the osmotic stress tolerance via abscisic acid (ABA)-mediated pathway [Magna 2008] but also in plant defence as its expression is rapidly induced by phytopathogenic bacteria, flagellin and salicylic acid [Leba 2012].

However, the physiological roles of CMLs is not limited to stress responses, but also involved in a large variety of developmental processes. For instance, CML 23 and 24 regulate plant flowering by inhibiting the expression of the central repressor: Flowering Locus C [Tsai 2007 and Delk 2005]. CML42 knockout showed aberrant trichomes with increased branching and its interaction with the kinesin-interacting  $\text{Ca}^{2+}$ -binding protein (KIC), a protein known to function in trichome development, was demonstrated [Dobney 2009]. CML25 mediates the  $\text{Ca}^{2+}$  regulation of  $\text{K}^+$  influx during pollen germination and tube elongation [Wang 2015] and CML20 represents one of the two centrins in Arabidopsis, taking part in microtubule organisation (see paragraph 5.1) [Azimzadeh 2008].

Although increasing information is now available that links CMLs function with defined biological processes, knowledge is limited on (i) the metal-binding and structural properties of this plant-specific class of  $\text{Ca}^{2+}$  sensors, and (ii) the identification of specific CMLs targets and the characterization of the CML-target interactions.

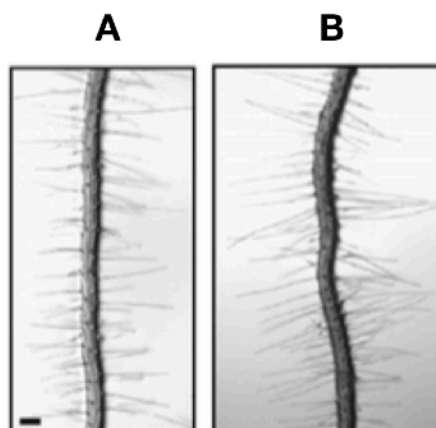




occupied by an Asp residue, is replaced by an Asn in CML7. EF-hand 2 shows an Asp to Asn substitution in position 3, while position 5 (in which an Asn is usually present) is occupied by an Asp residue. Similarly to EF-hand 2 of CML7, also in EF-hand 3 Asp to Asn substitutions in positions 3 and 5 are present. The residues responsible for  $\text{Ca}^{2+}$  coordination are fully conserved in EF-hand 4. In analogy to what reported for CaMs, it is likely that the binding of  $\text{Ca}^{2+}$  to each globular domain in CML7 changes the inter-helical angles in the EF-hand motifs, inducing exposure of hydrophobic sites that usually involves Met residues and favours the interactions with downstream partners. CML7 primary sequence contains eight Met residues; two of them are located in the N-lobe (Met12 and Met45), two in the central region (Met70 and Met80), and three in the C-lobe (Met121, Met141 and Met142). Only five of them are conserved in other CaMs [Trande 2019].

The high similarity with CaM2 might raise questions regarding the role of CML7: is it functionally redundant or does it play specific functions in physiological plant processes? Importantly, functional analysis performed by Won et al. [Won 2009] showed that the coding gene of CML7 (At1g05990) possesses root hair-specific cis-element in its promoters and so it is believed to be specifically expressed only in roots. In the work of Won and colleagues, further information has been collected about CML7. In particular, T-DNA insertion mutants allowed the evaluation of CML7-deficient phenotypes, which, as observed in Figure 5, showed an increase in root hair length, suggesting a negative regulation function of CML7 in root development.

However, molecular information on CML7 such as the mechanism of action, the  $\text{Ca}^{2+}$  sensing properties and its physiological targets has not been elucidated yet. This Chapter focuses on the biochemical characterisation of CML7, with the aim to evaluate the metal and target binding properties of the protein.

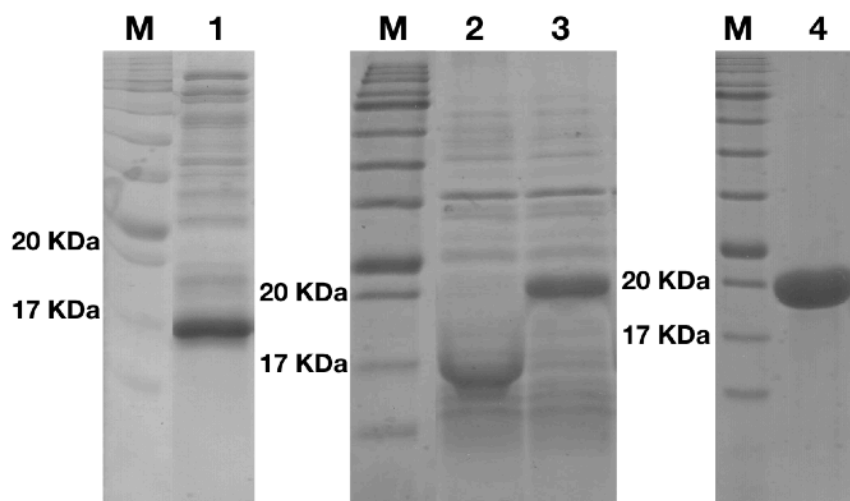


**Figure 5 [Won 2009]:** Root images of wild-type (A) and CML7-deficient (B) Arabidopsis lines.

## 4.2 RESULTS AND DISCUSSION

### 4.2.1 CML7 production

Initial attempts to purify CML7 from *E. coli* cells was performed by Nickel affinity chromatography exploiting the presence of a N-terminal 6His-tag. However, as observed in the SDS-PAGE analysis (Figure 6, lane 1) the final yield and purity of the protein were very low. Thus, we decided to apply another purification procedure involving heat treatment of the supernatant, since CMLs usually are heat-stable proteins, followed by hydrophobic interaction chromatography. For this procedure we cloned and expressed the gene of CML7 in *E. coli* cells without the hexa-histidine tag. As shown in Figure 6, lanes 2 and 3, after the heat treatment CML7 is in the soluble fraction and shows an increased electrophoretic mobility in the presence of  $\text{Ca}^{2+}$ , which is an hallmark of  $\text{Ca}^{2+}$  sensor proteins. Next, the heat-treated lysate containing  $\text{CaCl}_2$  was loaded onto a phenyl-Sepharose column.  $\text{Ca}^{2+}$ -dependent hydrophobic binding to phenyl-sepharose is typical of classic  $\text{Ca}^{2+}$  sensors, and thus represents a further clue indicating that CML7 has  $\text{Ca}^{2+}$  sensor properties. Moreover,

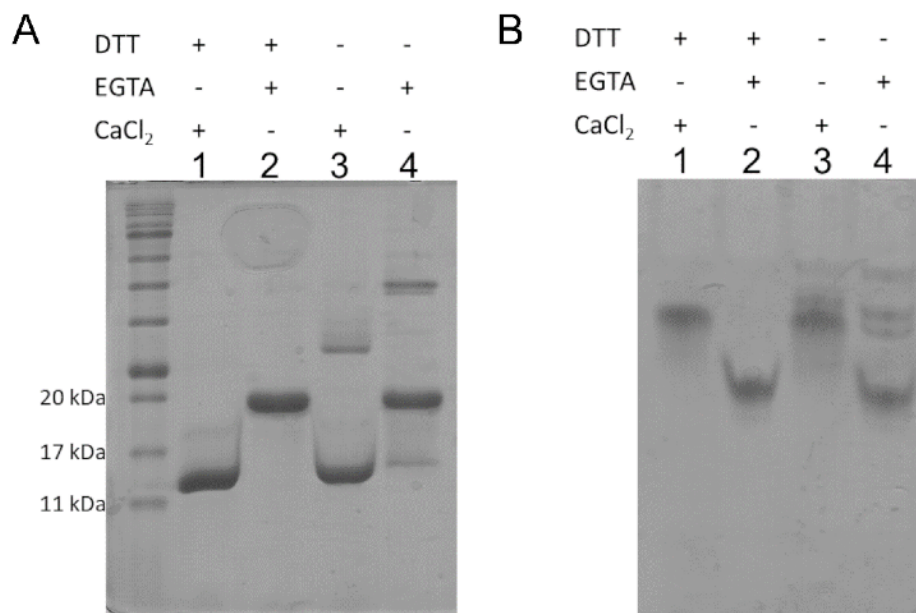


**Figure 6:** SDS-PAGE analysis of CML7 at different steps of the purification process. Lane M represents the molecular weight markers. Lane 1 is a representative eluted fraction of Ni-affinity column. Lanes 2 and 3 represent the soluble fraction after heat treatment in the presence of  $\text{Ca}^{2+}$  (lane 2) or EGTA (lane 3). Lane 4 is a representative eluted fraction from phenyl-sepharose column.

as we can observe in the SDS-PAGE analysis of eluted fractions (Figure 6 lane 4), recombinant CML7 was highly pure.

We then further evaluated CML7 electrophoretic mobility shift in the presence of  $\text{Ca}^{2+}$ , in both native and denaturing conditions (Figure 7). As previously mentioned and highlighted in the following Chapters, altered electrophoretic mobility in the presence of  $\text{Ca}^{2+}$  ions is a typical hallmark of  $\text{Ca}^{2+}$  sensors proteins. Although the molecular mechanism of such different migration remains unclear, a work by Burgess et al. [Burgess 1980] revealed that prolonged boiling intervals of  $\text{Ca}^{2+}$  saturated CaM resulted in the appearance of slower intermediates bands, proposed as partially saturated conformations, and the decrease in the staining intensity of the fastest migrating one. Thus it can be speculated that the increased mobility of  $\text{Ca}^{2+}$  saturated CaMs and CMLs is the results of incomplete protein denaturation by high temperature due to their high thermal stability in the  $\text{Ca}^{2+}$  bound conformation (see also paragraph 4.2.2).

CML7 shows in this regards the typical behaviour of a  $\text{Ca}^{2+}$  sensor, indeed, the presence of the cation produces the appearance of a faster migrating band at  $\sim 13$  KDa compared to the  $\sim 20$  KDa when EGTA is present (Figure 7A, Lanes 1 and 2

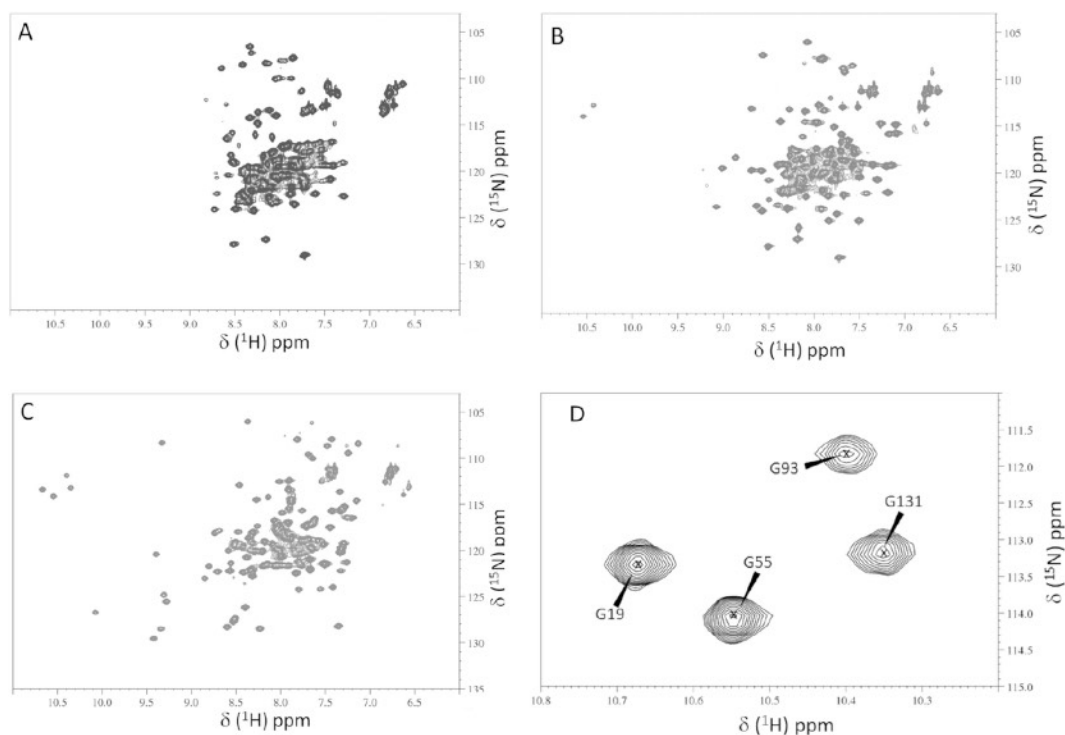


**Figure 7:** Purified CML7 analysed on a 15% SDS-PAGE (A) and a 12% native-PAGE (B) under reducing (lanes 1 and 2) and non-reducing (lanes 3 and 4) conditions, in the presence of 5 mM  $\text{CaCl}_2$  (lanes 1 and 3) or 5 mM EGTA (lanes 2 and 4).

respectively). Moreover, both apo- and Ca<sup>2+</sup>-bound CML7 show a single band on gel electrophoresis in the presence of DTT (Figures 7A and 7B, lanes 1 and 2), while additional bands were present in the absence of DTT for both forms, independently of the presence of Ca<sup>2+</sup> (lanes 3 and 4). This behaviour is likely due to the presence of oligomers through formation of intermolecular disulphide bridges via Cys oxidation.

#### **4.2.2 Apo-, Mg<sup>2+</sup>- and Ca<sup>2+</sup>-bound states of CML7 exist in distinct structural conformations**

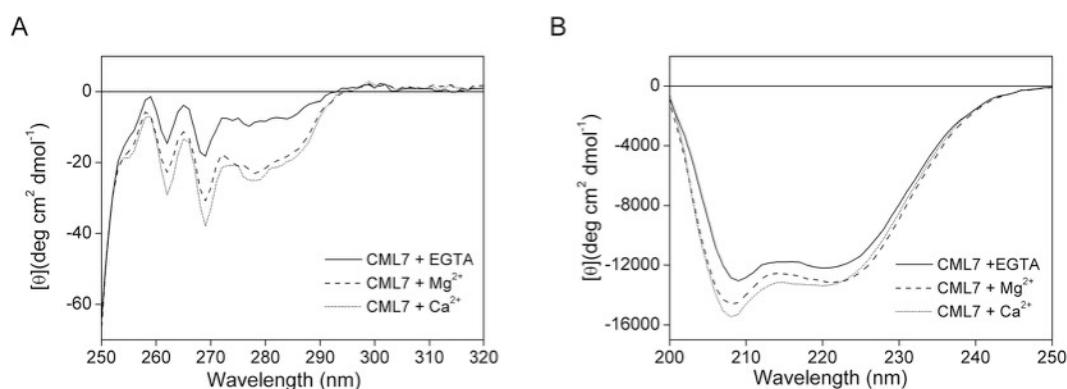
A typical hallmark of Ca<sup>2+</sup> sensors is the ability to experience a reversible conformational change upon Ca<sup>2+</sup> binding that leads to the exposure of hydrophobic regions involved in the binding and regulation of downstream targets. To evaluate these variations, we exploited two-dimensional NMR spectra (<sup>1</sup>H-<sup>15</sup>N HSQC) of <sup>15</sup>N-labeled CML7 resulting from Mg<sup>2+</sup> and Ca<sup>2+</sup> binding to the protein (Figure 8). In each spectrum, the peaks are representative of protons in the main and side chain amides, and their dispersion is a good probe of protein conformation. As shown in Figure 8, the apo-, Mg<sup>2+</sup>- and Ca<sup>2+</sup>-bound states of CML7 exist in distinct structural conformations. Upon addition of either Mg<sup>2+</sup> or Ca<sup>2+</sup> ions to the apo-protein a better dispersion of NMR signals was observed, and new peaks appeared, including down field shifted peaks between 10 and 11 ppm in the <sup>1</sup>H dimension and 111–115 ppm in the <sup>15</sup>N dimension, which are indicative of metal binding in the EF-hand motifs. In particular, the <sup>1</sup>H-<sup>15</sup>N HSQC spectra indicate that the Mg<sup>2+</sup>-bound conformation of CML7 is different from both its apo- and Ca<sup>2+</sup>-bound states. The broadening of the existing resonances line and the fewer number of peaks observed in the spectrum of Mg<sup>2+</sup>-CML7, compared to that of the Ca<sup>2+</sup>-CML7, suggest that Mg<sup>2+</sup> and Ca<sup>2+</sup> bind to CML7 with different affinities and that the Mg<sup>2+</sup>-bound form of CML7 may still possess a degree of flexibility. Despite this partial conformational instability, the dispersion of resonances observed in the two-dimensional spectrum of Mg<sup>2+</sup>-CML7 is distinct from the narrow range of <sup>1</sup>H chemical shifts observed in the apo-CML7 spectrum and suggests the presence of a more folded conformation of the protein when bound to Mg<sup>2+</sup>. The presence of Ca<sup>2+</sup> promoted a further change in chemical shifts that describes the transition between a closed apo conformation and an open Ca<sup>2+</sup>-bound state, analogous to that observed with both



**Figure 8:** Two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectra of apo- (A), Mg<sup>2+</sup>-bound (B), Ca<sup>2+</sup>-bound (C) <sup>15</sup>N-CML7 and zoom of the Gly<sub>6</sub> down field peaks of Ca<sup>2+</sup>-bound CML7 (D). The assignment of the Gly of the EF-hands is indicated by the position number of the residue.

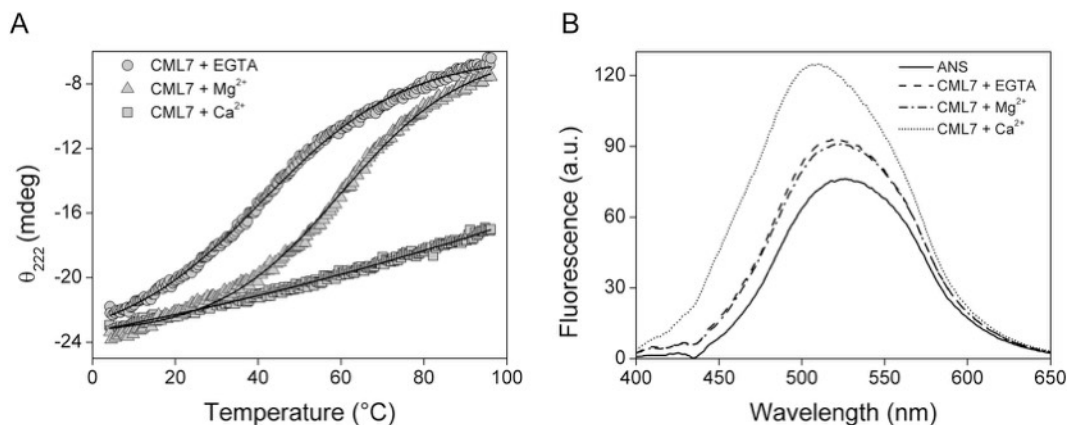
mammalian and plant CaMs [Vallone 2016, Dobney 2009, Astegno 2017 and Ohki 1997]. Many EF-hand proteins, including CML7, are characterized by a hydrogen bond network in metal-bound loops that causes the appearance of down field-shifted peaks in the <sup>1</sup>H-<sup>15</sup>N HSQC spectra, which belong to the backbone amides of glycine residues present at position 6 (Gly6) of the EF-hand loops when bound to metal. The extent of this down field shift enables these resonances to act as indicators for the Ca<sup>2+</sup>-bound state of a given EF-hand, providing indication on the stoichiometry of ion binding. Thus, by standard triple resonance NMR experiments acquired on <sup>15</sup>N,<sup>13</sup>C-labeled CML7 in its Ca<sup>2+</sup> bound state, we assigned the four resonances at positions 10.68, 10.55, 10.40 and 10.35 ppm (<sup>1</sup>H dimension) to the Gly6 of the EF-hands, Gly19 (EF-1), Gly55 (EF-2), Gly93 (EF-3), and Gly131 (EF-4), respectively (Figure 8 D). The presence of four low-field resonances assigned to the Gly6 of the four EF-hands strongly suggests that CML7 can bind four Ca<sup>2+</sup> ions.

NMR spectral features of CML7 were consistent with our CD experiments. As observed in Figure 9 A, the near UV CD fingerprint of CML7, which reflects protein tertiary structure through local arrangements of aromatic residues, showed very well resolved bands at 269 and 262 nm arising from Phe side chains and a broad minimum for Tyr at around 280 nm. No peaks corresponding to Trp residue are present as CML7 lacks Trp residues. Binding of  $Mg^{2+}$  to apo-CML7 produced significant changes in the near UV CD spectrum, with a decrease in the negative ellipticity both in the Phe and Tyr bands (Figure 9 A, dashed line). Addition of  $Ca^{2+}$  refined the spectrum, slightly increasing the signal intensity, consistently with further changes in the tertiary structure upon  $Ca^{2+}$  binding (Figure 9 A, dotted line). Consistently, CD spectra in the far UV region indicated that CML7 possesses significant  $\alpha$ -helical structures as seen by the characteristic negative ellipticity peaks at 222 nm and 208 nm in both the apo- and metal-bound forms (Figure 9 B). Upon  $Ca^{2+}$  binding, CML7 showed a change of the spectrum characterised by an increase in intensity and variations in shape visible as a change in the  $\theta_{222}/\theta_{208}$  ratio (0.94 for apo- CML7 and 0.86 for  $Ca^{2+}$ -bound CML7), which has been ascribed mainly to helix reorientation within the tertiary structure, and not to a change in  $\alpha$ -helical content [Zhang 1995 and Gagne 1994]. Thus, NMR and CD data clearly indicated that both  $Ca^{2+}$  and  $Mg^{2+}$  ions induce conformational changes, affecting the secondary and tertiary structure of CML7.



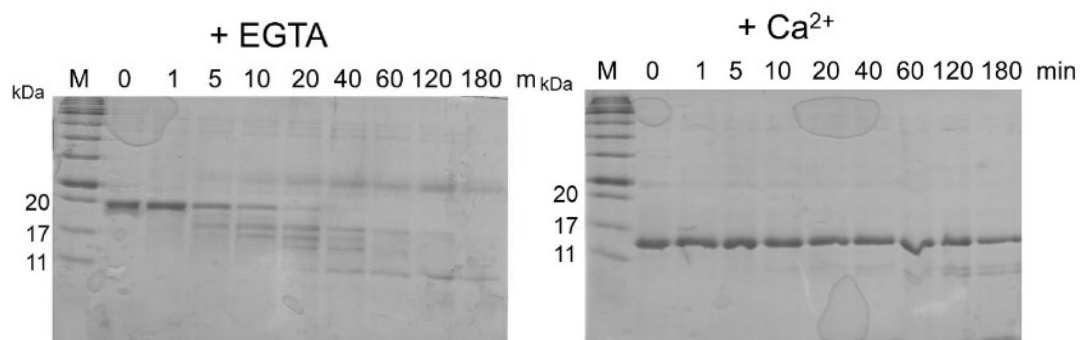
**Figure 9:** Near (A) and far (B) UV CD spectra of apo-CML7 (solid line) and after sequential additions of 2 mM  $MgCl_2$  (dashed line) and 2 mM  $CaCl_2$  (dotted line).

The metal binding also resulted in an increase of the thermal stability of the protein (Figures 10 A and 11). The thermal unfolding profiles of apo- and metal-bound CML7 were compared by following the CD signal at 222 nm as a function of



**Figure 10:** CML7 thermal stability and ANS fluorescence. (A) Thermal denaturation profiles of CML7 in the presence of EGTA (circle),  $MgCl_2$  (triangle) and  $CaCl_2$  (square). (B) ANS fluorescence spectra of CML7 upon incubation with EGTA (dashed line),  $MgCl_2$  (dot dashed line) and  $CaCl_2$  (dotted line). The fluorescence spectrum of ANS alone is also shown in solid line.

temperature. In the sample containing CML7 and EGTA the melting temperature ( $T_m$ ) of CML7 was  $40.9 \pm 0.8^\circ C$  whereas the  $T_m$  in the presence of  $Mg^{2+}$  was  $60.0 \pm 0.4^\circ C$ . In the presence of  $Ca^{2+}$ , no estimate of  $T_m$  was possible as the denaturation curve increases monotonically with increasing temperature up to  $\sim 95^\circ C$ , and even at  $95^\circ C$  the protein maintains a partially folded state. Along with this, proteolysis by trypsin revealed that the  $Ca^{2+}$ -bound form of CML7 was more stable as the protein band corresponding to intact CML7 was still present even after 3 h of incubation with trypsin compared to the apo-form of CML7 which disappeared after 10 min (Figure 11). This data suggests that the binding of  $Ca^{2+}$  or  $Mg^{2+}$  strongly



**Figure 11:** SDS-PAGE analysis of CML7 digestion by trypsin in the presence of EGTA or  $CaCl_2$ , after incubation of CML7 with trypsin 1:500 (w/w) for several time intervals. Lane M represents molecular weight markers.

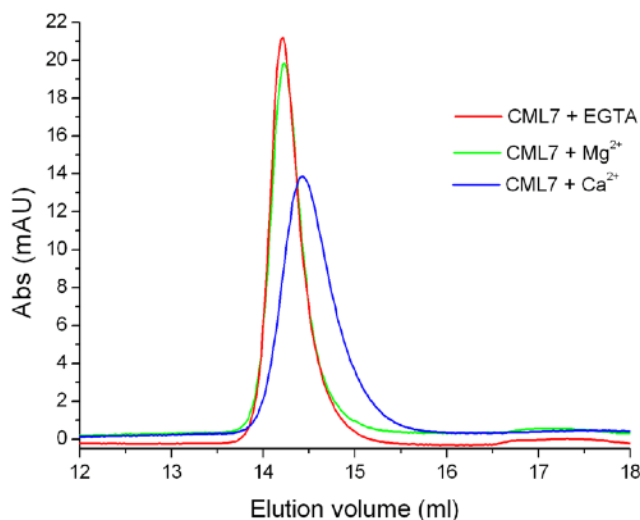
stabilises CML7 by leading to a rearrangement in both the secondary and the tertiary structures.

However, for each metal the stabilised final conformation is distinct.  $\text{Ca}^{2+}$ -induced conformational changes in  $\text{Ca}^{2+}$  sensors are often accompanied by exposure of Met-rich hydrophobic surfaces which represent a crucial step in target recognition. A specific conformational probe for such effect is the hydrophobic fluorophore ANS, since its emitted fluorescence is altered after binding to hydrophobic surfaces. Using ANS probe, we demonstrated that the binding of  $\text{Ca}^{2+}$  induces a domain opening in CML7 leading to the exposure of hydrophobic regions, whereas  $\text{Mg}^{2+}$  does not. As shown in Figure 10 B, upon addition of  $\text{Ca}^{2+}$  to CML7 a  $21 \pm 2$  nm blue-shift in the emission peak of ANS fluorescence (from 525 nm to 504 nm) was observed along with a  $1.6 \pm 0.1$  fold increase in the maximal emission intensity, indicating that the hydrophobic regions of the recombinant CML7 were exposed to solvent, and therefore accessible to the ANS probe. On the contrary, the addition of  $\text{Mg}^{2+}$  to CML7 did not cause a significant change in the ANS fluorescence spectrum compared to that observed in the apo form. However, few variations in the emission spectra were observed for either apo-CML7 or  $\text{Mg}^{2+}$ -bound CML7 compared to the spectrum of ANS alone (blue shift of  $7 \pm 1$  and  $6 \pm 2$  nm, respectively and  $1.2 \pm 0.1$  fold increase in intensity), suggesting that some hydrophobic exposure is present even in the absence of  $\text{Ca}^{2+}$ .

In addition, SEC studies were conducted to determine the Stokes radii ( $R_s$ ) of apo-,  $\text{Ca}^{2+}$ -bound and  $\text{Mg}^{2+}$ -bound CML7. SEC allows the evaluation of variations in the hydrated shape of proteins, thus monitoring  $\text{Ca}^{2+}$ -induced conformational changes in  $\text{Ca}^{2+}$  sensors. CML7 was found to have a  $R_s$  of  $22.61 \pm 0.15$  Å in the presence of EGTA. Considering the molecular weight of the protein (17.78 KDa), the relatively high  $R_s$  can be explained by the an elongated shape, as observed for other  $\text{Ca}^{2+}$  sensor proteins [Astegno 2016, La Verde 2017, Wiech 1996 and Sorensen 2001]. Addition of  $\text{Ca}^{2+}$  to CML7 caused a decrease in its  $R_s$  ( $21.46 \pm 0.12$  Å), which is likely the result of the conformational change induced upon  $\text{Ca}^{2+}$  binding. In contrast,  $\text{Mg}^{2+}$  binding did not induce a variation in  $R_s$  ( $22.50 \pm 0.11$  Å) and, therefore, in the hydration shell structure compared to apo-protein.

Thus, we showed that, in accordance with a putative function as a  $\text{Ca}^{2+}$  sensor, CML7 undergoes a large structural rearrangement leading to the solvent exposure of hydrophobic regions only upon the binding of  $\text{Ca}^{2+}$ .





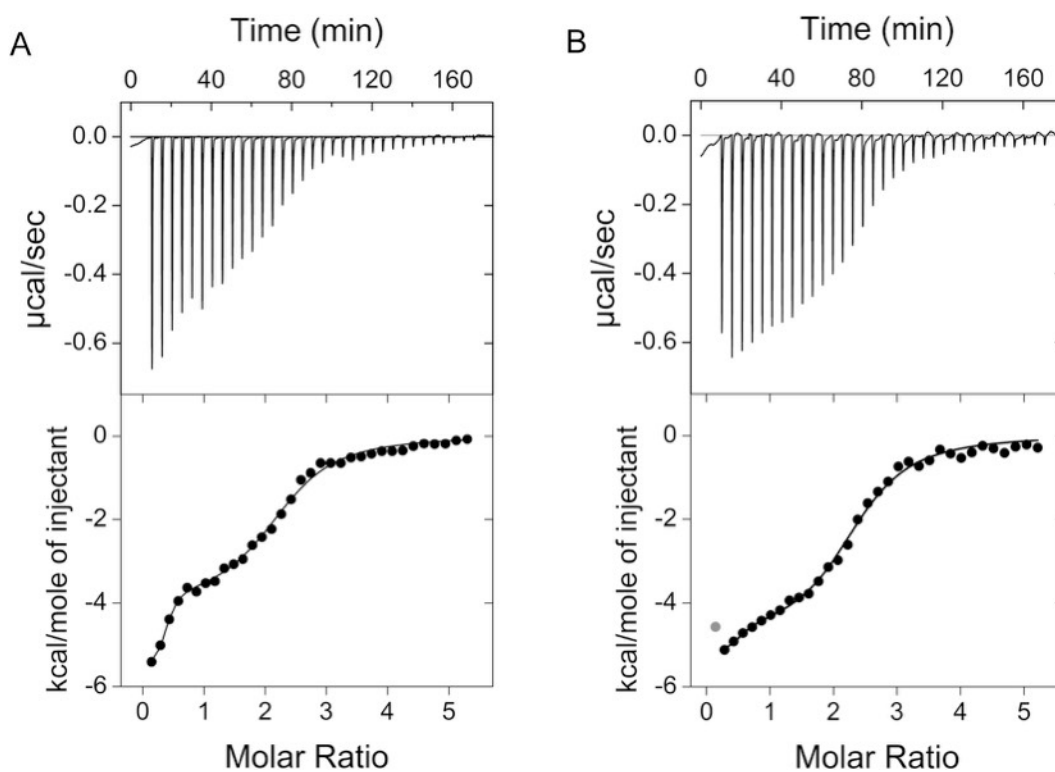
**Figure 12:** Size exclusion chromatography analysis of CML7 on a Superose 12 column 10/300GL. Elution profiles were recorded in the presence of EGTA (red line), MgCl<sub>2</sub> (green line) or CaCl<sub>2</sub> (blue line).

#### 4.2.3 CML7 has two high affinity Ca<sup>2+</sup> sites in the C-lobe and two weaker in the N-lobe

We examined Ca<sup>2+</sup> binding to CML7 by ITC, starting from conditions of decalcified protein and buffer (Figure 13). Data were fitted using a sequential binding model with fixed stoichiometry of Ca<sup>2+</sup> binding to four sites (as deduced by NMR experiments). The values obtained for the four macroscopic binding constants showed that the Ca<sup>2+</sup> binding sites in CML7 are grouped into two pairs of different strength (Table 3): a first set of sites that comprises two high-affinity sites ( $K_{d1} = 0.9 \pm 0.2 \mu\text{M}$ ,  $K_{d2} = 0.5 \pm 0.1 \mu\text{M}$ , respectively) and a second set that likely represents the binding of two Ca<sup>2+</sup> ions with lower affinity ( $K_{d3} = 14.7 \pm 0.8 \mu\text{M}$ ,  $K_{d4} = 29.4 \pm 4.7 \mu\text{M}$ ).

Ca<sup>2+</sup> binding to CML7 followed an exothermic reaction, displaying a negative enthalpy change.

To learn more about the pathways of Ca<sup>2+</sup> binding to CML7, Ca<sup>2+</sup> was titrated into a solution of apo-CML7 and two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC spectra were recorded at each step. As shown in Figure 14, the Gly6 peaks of the C-terminal lobe (Gly93 and Gly131) appeared simultaneously at Ca<sup>2+</sup>:protein ratios lower than those of the N-terminal Gly6 (Gly19 and Gly55) which developed together only at higher Ca<sup>2+</sup> concentration. These observations indicate that the high affinity sites are located in the C-lobe of CML7 and allowed us to assign the dissociation constants determined



**Figure 13:** Representative raw trace (top) of the calorimetric titration of  $\text{CaCl}_2$  into 200  $\mu\text{M}$  apo-CML7 (A) and 200  $\mu\text{M}$  CML7 pre-incubated with 10 mM  $\text{MgCl}_2$  (B), and integrated binding isotherm (bottom). Curve fitting was performed using a four-site sequential binding model. The ligand dilution blank experiments ( $\text{Ca}^{2+}$  ions titrated into buffer) were subtracted from the binding isotherm obtained in the presence of protein.

	No $\text{Mg}^{2+}$			+ 10 mM $\text{Mg}^{2+}$		
	$K_a$ ( $\text{M}^{-1}$ )	$K_d$ ( $\mu\text{M}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )	$K_a$ ( $\text{M}^{-1}$ )	$K_d$ ( $\mu\text{M}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )
Site 1	$1.1\text{E}6 \pm 2.9\text{E}5$	$0.9 \pm 0.2$	$-6.6\text{E}3 \pm 0.8\text{E}3$	$2.6\text{E}5 \pm 6.3\text{E}4$	$3.8 \pm 0.9$	$-5.9\text{E}3 \pm 0.9\text{E}3$
Site 2	$2.2\text{E}6 \pm 4.1\text{E}5$	$0.5 \pm 0.1$	$-1.6\text{E}3 \pm 0.4\text{E}3$	$3.2\text{E}5 \pm 6.2\text{E}4$	$3.1 \pm 0.6$	$-1.9\text{E}3 \pm 0.3\text{E}3$
Site 3	$6.8\text{E}4 \pm 3.9\text{E}3$	$14.7 \pm 0.8$	$-3.1\text{E}3 \pm 0.7\text{E}3$	$4.9\text{E}4 \pm 7.3\text{E}3$	$20.4 \pm 3.0$	$-2.6\text{E}3 \pm 0.6\text{E}3$
Site 4	$3.4\text{E}4 \pm 5.4\text{E}3$	$29.4 \pm 4.7$	$-1.5\text{E}3 \pm 0.4\text{E}3$	$2.6\text{E}4 \pm 8.4\text{E}3$	$38.5 \pm 6.9$	$-0.8\text{E}3 \pm 0.1\text{E}3$

**Table 3:** Thermodynamic parameters for  $\text{Ca}^{2+}$  binding to CML7 in the absence (left) or in the presence (right) of  $\text{Mg}^{2+}$ . Sites 1 and 2 are located at the C-terminal lobe and sites 3 and 4 at the N-terminal, as explained in the main text.

from ITC Kd1 and Kd2 to the C-terminal domain (higher affinity Ca<sup>2+</sup>-binding sites), and Kd3 and Kd4 to the N-terminal EF-hands (lower affinity Ca<sup>2+</sup>-binding sites).

The paired nature of these EF-hands and the fact that two of the sites have higher Ca<sup>2+</sup> affinity than the other two enabled us to estimate the cooperativity within each EF-hand pair. The cooperativity in a two-site system (EF-hand pair) can be quantified in terms of free energy, with the term  $-\Delta\Delta G$  reflecting the increase in affinity of the second EF-hand as a result of Ca<sup>2+</sup> binding to the first [Linse 1991 and 1995]. If the affinity for the second ion is greater than the first, this term is positive, indicating positive cooperativity. As shown in Table 4, each pair of CML7 EF-hands displayed positive cooperativity. Thus, we can conclude that the two sites of the C-terminal domain in CML7 bind Ca<sup>2+</sup> most strongly in a cooperative manner.

	$\Delta G^{1,2}$ (KJ mol <sup>-1</sup> ) <sup>a</sup>	$-\Delta\Delta G_{\eta=1}^{1,2}$ (KJ mol <sup>-1</sup> ) <sup>b</sup>	$\Delta G^{3,4}$ (KJ mol <sup>-1</sup> ) <sup>c</sup>	$-\Delta\Delta G_{\eta=1}^{3,4}$ (KJ mol <sup>-1</sup> ) <sup>d</sup>
<b>Apo-CML7</b>	-70.7 ± 0.8	5.2 ± 0.8	-53.4 ± 0.4	1.7 ± 0.4
<b>+10 mM Mg<sup>2+</sup></b>	-62.3 ± 0.8	4.0 ± 0.8	-52.0 ± 0.9	1.9 ± 0.9

**Table 4:** Free energy of binding of the first two ( $\Delta G^{1,2}$ ) and last two ( $\Delta G^{3,4}$ ) Ca<sup>2+</sup> ions to apo-CML7 and Mg<sup>2+</sup>-bound CML7 and the lower limit of the cooperativity within each pair ( $-\Delta\Delta G_{\eta=1}$ ). The presence of the term  $\eta=1$ , indicates the case of equally strong sites [Linse 1991 and 1995 and Gifford 2007 and 2013]

<sup>a</sup>  $\Delta G^{1,2} = -RT \ln(K_{a1}K_{a2})$

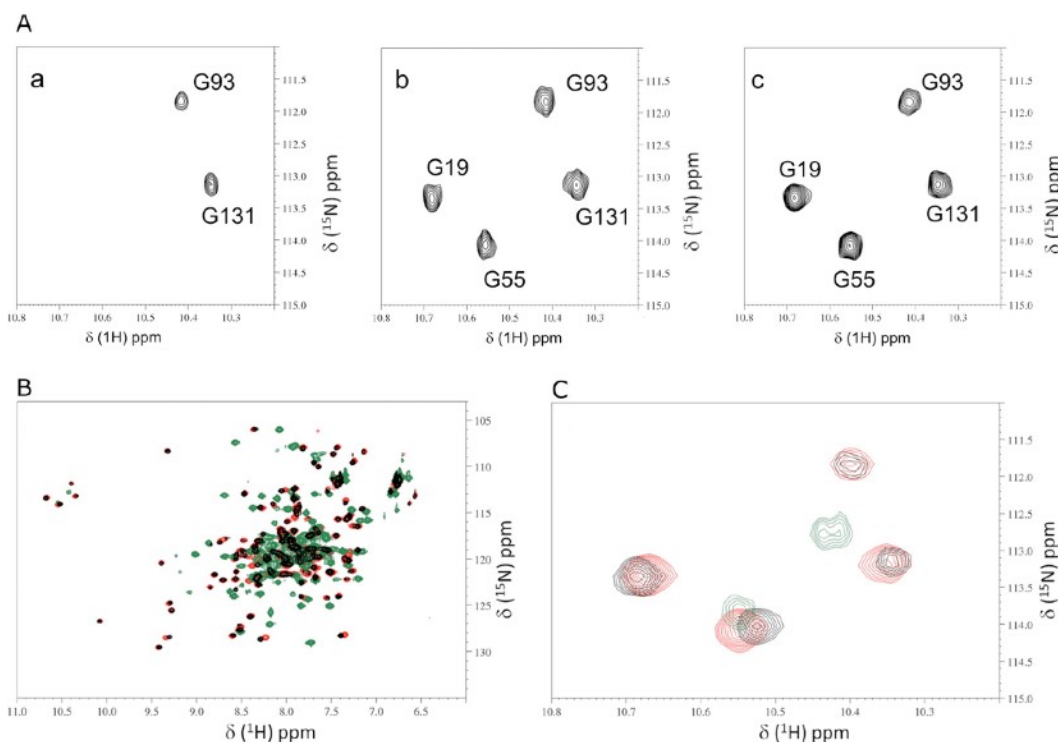
<sup>b</sup>  $-\Delta\Delta G_{\eta=1}^{1,2} = RT \ln(4K_{a2}/K_{a1})$

<sup>c</sup>  $\Delta G^{3,4} = -RT \ln(K_{a3}K_{a4})$

<sup>d</sup>  $-\Delta\Delta G_{\eta=1}^{3,4} = RT \ln(4K_{a4}/K_{a3})$

We next investigated the influence of Mg<sup>2+</sup> on the Ca<sup>2+</sup>-binding thermodynamic parameters as well as association constants of CML7 by ITC. Ca<sup>2+</sup> titrations were performed upon incubation of CML7 with 10 mM MgCl<sub>2</sub> (Figure 13, Table 3 and 4). Although the four Ca<sup>2+</sup>-binding sites still bind Ca<sup>2+</sup> with positive cooperativity, the presence of 10 mM Mg<sup>2+</sup> slightly decreased the Ca<sup>2+</sup> association constants of sites 1 and 2 (~up to 6-fold), while it did not significantly modify the Ca<sup>2+</sup> association constants of the lower affinity sites (sites 3 and 4) of CML7. Importantly, Ca<sup>2+</sup>

addition to  $Mg^{2+}$ -bound CML7 leads to the disappearance of the peaks associated with  $Mg^{2+}$  bound form in  $^1H$ - $^{15}N$  HSQC spectrum and the concomitant appearance of the peaks characteristic of the  $Ca^{2+}$ -bound CML7; the resulting spectrum is nearly superimposable with that of  $Ca^{2+}$ -CML7 (Figure 14). This provides evidence of  $Mg^{2+}$  displacement by  $Ca^{2+}$ , and that CML7 attains the same conformation after binding  $Ca^{2+}$ , in both the presence and absence of  $Mg^{2+}$ .



**Figure 14:** NMR analysis of the  $Ca^{2+}$  binding to apo- and  $Mg^{2+}$ -bound CML7. (A) Down field region of selected  $^1H$ - $^{15}N$  HSQC NMR spectra of CML7 recorded as a function of increasing  $Ca^{2+}$  concentration. The molar ratio of  $Ca^{2+}$  : protein was 2 (a), 4.5 (b), 10 (c). (B) Superimposition of  $^1H$ - $^{15}N$  HSQC spectra recorded on  $^{15}N$ -CML7 after addition of 5 mM  $CaCl_2$  (red), 20 mM  $MgCl_2$  (green), or 20 mM  $MgCl_2$  + 20 mM  $CaCl_2$  (black). (C) Zoom of the Gly down field peaks.

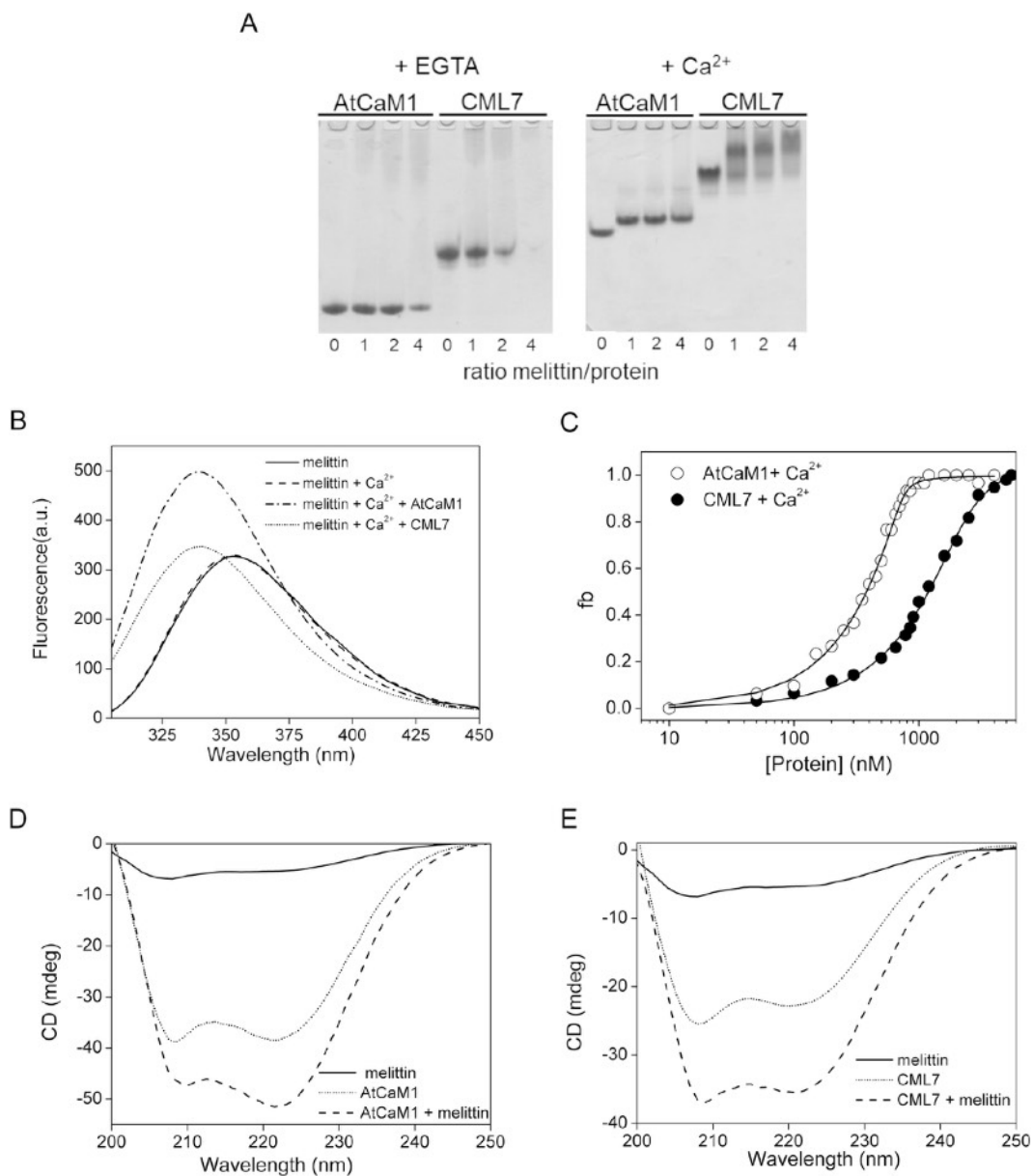
#### 4.2.4 CML7 binds to the model target peptide melittin in a Ca<sup>2+</sup>-dependent manner

In order to demonstrate that CML7 behaves as a member of CaM family in terms of targets interaction and since CML7 physiological target/s in plants have not been identified yet, we examined CML7 for interaction with melittin. Melittin is an amphipathic model peptide (GIGAVLKVLTTGLPALISWIKRKRQQ) that is used to study the interaction of downstream targets with proteins of the CaM family. It interacts with Ca<sup>2+</sup>-saturated bovine CaM to form a 1:1 complex with nanomolar affinity [Comte 1983], assuming an  $\alpha$ -helix conformation upon binding [Seeholzer 1986].

In the experiments, a canonical plant CaM available in our lab (AtCaM1 from *A. thaliana*) was used as a positive control.

Initially, the interaction of CML7 or AtCaM1 with melittin was analysed by native PAGE in the presence of increasing amounts of the peptide (Figure 15 A). Nondenaturing gel band shift electrophoresis directly confirmed that melittin forms a 1:1 complex with AtCaM1 only in the presence of Ca<sup>2+</sup>. Melittin also interacts with CML7, forming a complex in the presence of Ca<sup>2+</sup>. Upon addition of the peptide, the single band reflecting pure CML7 disappeared and a new protein band of lower mobility appeared, indicating the formation of a new complex. At 1:1 protein:peptide ratio, almost all the free CML7 disappeared, suggesting the formation of a 1:1 peptide/CML7 complex. At higher concentrations of melittin, a smearing of the band representing the complex was observed, likely indicating that additional binding occurs (up to 4 mol of melittin/mol of CML7).

The formation of protein/peptide complexes was also studied by fluorescence spectroscopy. Because neither CML7 or AtCaM1 have any Trp residues and melittin possesses only one Trp (Trp19), we exploited the single Trp of the peptide as a probe to monitor binding of the peptide to the protein since, upon complex formation, Trp usually experiences changes in its spectral properties. For melittin alone, as expected for a solvent exposed Trp residue, the maximum fluorescence emission wavelength was 354 nm and no changes in the fluorescence spectrum were observed after the addition of CaCl<sub>2</sub>. On the contrary, the fluorescence was blue shifted to 339 nm and 338 nm upon binding to AtCaM1 and CML7, respectively (in the presence of saturating concentrations of Ca<sup>2+</sup>) (Figure 15 B), which is indicative of



**Figure 15:** Interaction of melittin with AtCaM1 and CML7. (A) Native PAGE in the presence of  $\text{CaCl}_2$  or EGTA. (B) Complex formation monitored by Trp fluorescence. Emission spectra, recorded in the presence of  $\text{CaCl}_2$ , of melittin (dashed line), melittin/AtCaM1 (dash dotted line) and melittin/CML7 complex (dotted line). The spectrum of melittin before addition of  $\text{CaCl}_2$  (solid line) is also shown. (C) Fluorescence titration of AtCaM1 (open circles) or CML7 (filled circles) with melittin in the presence of  $\text{CaCl}_2$ . (D–E) Far UV CD spectra changes in secondary structure for AtCaM1/ melittin (D) and CML7/melittin complex (E) in the presence of  $\text{CaCl}_2$ . Melittin alone (solid line), AtCaM1 or CML7 alone (dotted line), and protein/peptide complex at 1:1 molar ratio (dashed line).

an augmented hydrophobicity of the Trp environment, suggesting that the interaction between proteins and melittin occurs via the insertion of the melittin Trp inside an hydrophobic pocket. The weaker intensity of the Trp fluorescence emission when melittin is bound to CML7 compared to AtCaM1 may be due to increased quenching either by the surrounding amino acid side chains or from a greater exposure to solvent. Moreover, the microenvironment of the Trp residue of melittin could be perturbed in a different way in the two complexes. However, it might also suggest that the Trp is less rigidly held in the binding pocket of CML7, therefore we cannot exclude the presence of conformational changes which could potentially alter the fluorescence properties of melittin in a different manner in the two complexes.

To estimate the apparent affinity between proteins and melittin, titration experiments in Trp fluorescence were performed by adding increasing amount of AtCaM1/CML7 to  $\text{Ca}^{2+}$  or EGTA saturated melittin (Figure 15 C, paragraph 3.6; EGTA saturated experiments are not shown). In the presence of  $\text{Ca}^{2+}$ , AtCaM1 binds melittin with high affinity ( $K_d$  of  $8 \pm 2$  nM) and with 1:1 stoichiometry, similarly to previously reported data for bovine CaM/melittin interaction [Comte 1983]. Binding of melittin to  $\text{Ca}^{2+}$ -bound CML7 also occurs with high affinity ( $K_d$  of  $0.3 \pm 0.1$   $\mu\text{M}$ ), even if lower compared to AtCaM1.

Furthermore, to examine if the binding of melittin to AtCaM1 or CML7 is associated with changes in secondary structure we acquired far-UV CD spectra of the complexes. As shown in Figures 15 D and E, CD signals of both AtCaM1 and CML7 were increased following the addition of melittin. In aqueous solution, free melittin has a small  $\alpha$ -helical content, as observed for monomeric peptides that are 20-25 residues in length [Smith 1998 and Kuhlman 1997]. As described in paragraph 4.2.2., the spectrum of AtCaM1 and CML7 showed the typical two minima peaks at 208 and 222 nm of a protein with a significant  $\alpha$ -helical content [Kelly 2005]. Addition of equimolar melittin yielded a spectrum that was more negative than the sum of each taken separately, suggesting that the  $\alpha$ -helical content of the  $\text{Ca}^{2+}$ -protein/peptide complex increases. The higher value of CD signal observed is attributable to induced helicity in melittin, since  $\text{Ca}^{2+}$ -CaM itself does not gain secondary structure following binding to target peptide [Ikura 1992 and Meador 1992]. Moreover, in the case of AtCaM1, the shape of the spectrum representing the complex was also different from the individual components. In the presence of melittin and  $\text{Ca}^{2+}$ , the  $\theta_{222}/\theta_{208}$  ratio for AtCaM1 approaches the value of  $\sim 1.08$

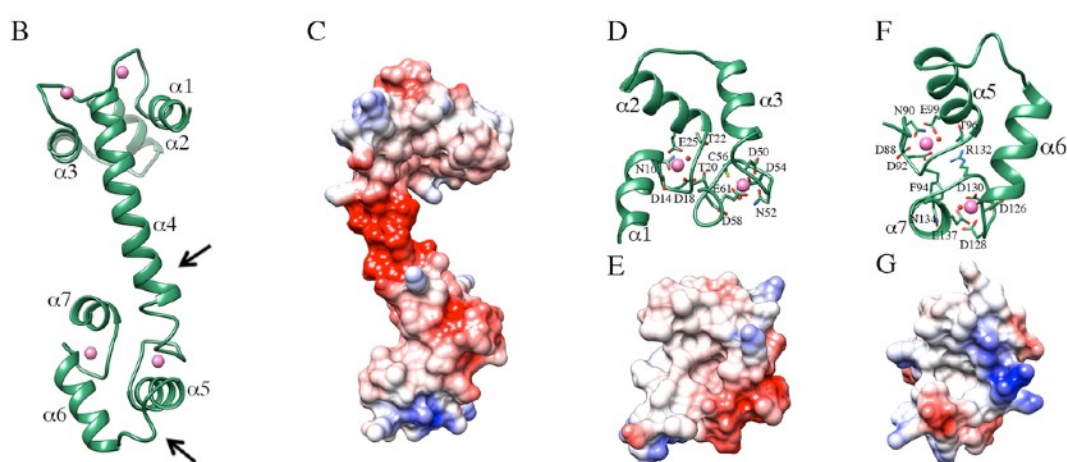
observed for coiled-coil motifs.[Lau 1984 and Zhou 1992]. Thus, it is possible that a tight interaction between  $\alpha$ -helices, which in certain regions may coil together, occurs when AtCaM1 binds melittin. Importantly, in the presence of EGTA, CD spectra were characterized by only a modest change upon addition of melittin to CML7 (data not shown), suggesting that the interaction takes place mainly in the presence of  $\text{Ca}^{2+}$ .

#### 4.2.5 Homology modelling of CML7

In collaboration with Dr. Adele di Matteo at the Institute of Molecular Biology and Pathology, National Research Council, Department of Biochemical Sciences A. Rossi Fanelli, Sapienza University of Rome, a homology model of the CML7 from *A. thaliana* in the  $\text{Ca}^{2+}$  bound form was generated using the structure of the CaM from *Paramecium tetraurelia* (PDB entry: 1clm [Rao 1993]) as template. The model assumes the dumbbell-like structure typical of CaM proteins formed by two structurally similar globular domains (the N-term lobe and the C-term lobe), containing a pair of  $\text{Ca}^{2+}$ -binding motifs connected by a central  $\alpha$ -helix (Figures 16 B and C). In analogy to what reported for CaMs, it is likely that the binding of  $\text{Ca}^{2+}$  to each globular domain in CML7 changes the inter-helical angles in the EF-hand motifs, inducing exposure of hydrophobic sites that favour the interactions with downstream partners. As shown in Figures 16 E and G, the Coulombic surface of both N-terminal and C-terminal lobes of the model of CML7 in the  $\text{Ca}^{2+}$ -bound form revealed the presence of hydrophobic patches in which Met residues likely play a key role in the interaction with targets [Tidow 2013]. CML7 primary sequence contains eight Met residues; the starting one (Met1), two located in the N-lobe (Met12 and Met45), two in the central region (Met70 and Met80), and three in the C-lobe (Met121, Met141 and Met142). Only five of them are conserved in other CaMs. In addition to the exposure of hydrophobic residues upon  $\text{Ca}^{2+}$  binding, the ability of CaM to recognize different interacting molecules is assured by the presence of a central linker helix connecting the two globular domains [Tidow 2013 and Spyraakis 2011]. This helix has been described as “magical” because it is highly flexible, allowing CaM to interact with a wide range of targets in a remarkable variety of conformations [Ikura 1992 and Wilson 2000]. For this reason, while the CML7



model can be considered highly accurate in the  $\text{Ca}^{2+}$  binding domains, the region corresponding to the central helix may be less accurate. Moreover, CML7 has an amino acid insertion (Asp79) in the C-terminal end of the central helix (Figure 16 B) that may lead the linker to adopt different orientations compared to typical CaMs. In CML7, this helix appears to contain a negatively charged region that is longer than that present in CaM proteins (Figure 16 C). This region is quite variable among CMLs and this diversity may contribute to different modes of target recognition and/or binding topology that distinguish CMLs from one another.

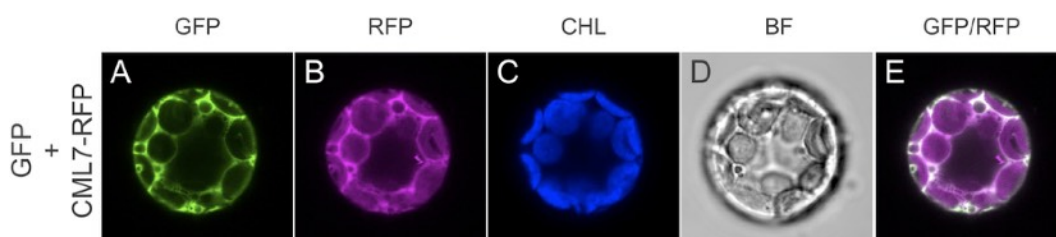


**Figure 16:** Molecular modelling of CML7. (B) Cartoon representation of the CML7 molecular modelling;  $\text{Ca}^{2+}$  is represented as a pink sphere; arrows represent region of amino acid insertion in the CML7 sequence. (C) Coulombic surface colouring of the CML7 molecular surface. (D–F) Close view of the N-lobe and the C-lobe respectively. Residues involved in  $\text{Ca}^{2+}$  coordination are represented in stick, water molecules (red spheres) were transferred from the template (E–G) Coulombing surface coloring of the CML7 N-lobe and the C-lobe.

#### 4.2.6 CML7 subcellular localisation

Determining protein localization is a necessary step if we aim to understand the function of CML7 in a physiological context and examine functional redundancy between CaM and CML isoforms. Interesting, prediction tools (i.e. Plant-mPLoc) suggested different localization including cytoplasm, plasma membrane and vacuole. Therefore, in collaboration with Dr Cristina Bonza and Prof. Alex Costa (University

of Milan), we determined the subcellular localization of CML7 in plant cells. In particular, we fused the RFP at the C-terminus end of CML7 (see paragraphs 3.1 and 3.11) and placed it under the control of the CaMV35S promoter. *A. thaliana* protoplasts cotransformed with the CML7-RFP fusion and the free EGFP showed green fluorescence with typical cytoplasmic localization (Figure 17 A). The red fluorescence from the CML7-RFP fusion was also restricted to cytoplasm (Figure 17 B), perfectly overlapping with the green fluorescence (Figure 17 E), in contrast with previous prediction.



**Figure 17:** Subcellular localization of CML7-RFP in *A. thaliana* mesophyll protoplasts. The CML7-RFP fusion protein was transiently coexpressed in protoplasts with free EGFP. Fluorescence was analyzed by confocal microscopy. (A) EGFP fluorescence (green) from a single focal plane of a representative protoplast. (B) RFP fluorescence (magenta) from the same protoplast as shown in (A). (C) Chlorophyll autofluorescence of chloroplasts. (D) Bright field image. (E) Overlay image of A and B.

### 4.3 CONCLUSIONS

In this Chapter the biochemical and biophysical characterization of CML7 was described, providing significant information about the structure modulation induced by  $\text{Ca}^{2+}$ . Apo-CML7 was found to have an extensive helical content and adopted a closed conformation with none or limited exposure of hydrophobic surfaces. The transition into an open and more stable conformation is induced by the presence of  $\text{Ca}^{2+}$  resulting in the reorientation of the helical content and the formation of amphipathic regions at the C and N terms, which are supposed to take part in target recognition. CML7 possesses four EF-hand motifs and the ability of all of them to bind  $\text{Ca}^{2+}$  was confirmed; moreover, CML7 was shown to possess two high affinity  $\text{Ca}^{2+}$  sites in the C-lobe and two weaker sites in the N-lobe. This high/moderate  $\text{Ca}^{2+}$  affinity appears to be an emerging pattern in CMLs [La Verde 2018 and Ogunrinde

2017]; the growing idea is that CMLs might possess different  $\text{Ca}^{2+}$  binding affinities to specifically sense a broad range of cytosolic free  $\text{Ca}^{2+}$  variations and, for most CMLs, they might be considerably different from the universal  $\text{Ca}^{2+}$  sensor CaM [Gifford 2013 and Bender 2014]. Importantly, positive cooperativity of  $\text{Ca}^{2+}$  binding is observed within each globular domain of CML7, in line with the paired nature of these plant CML7 EF-hands, and as frequently observed in EF-hand containing proteins [Linse 1991 and 1995].

We also studied the effect of  $\text{Mg}^{2+}$  on the binding of  $\text{Ca}^{2+}$  and the conformation of CML7. The interaction between  $\text{Mg}^{2+}$  and CML7 is crucial for understanding how CML7 specifically responds to micromolar variations in free cytosolic  $\text{Ca}^{2+}$  concentrations in the presence of  $\sim 1000$ -fold excess of the chemically similar  $\text{Mg}^{2+}$  ions. In the presence of  $\text{Mg}^{2+}$  the order of  $\text{Ca}^{2+}$  binding to CML7 remained unchanged and similar positive cooperativity was also observed. The modest effect of  $\text{Mg}^{2+}$  on  $\text{Ca}^{2+}$  affinity of CML7 and the null effect in inducing exposure of hydrophobic patches suggest that the  $\text{Mg}^{2+}$ -bound protein and the  $\text{Mg}^{2+}$ -bound loops likely remain in a “closed” conformation and allow the EF-hands of CML7 to be classified as  $\text{Ca}^{2+}$ -specific sites.

Overall, the obtained results indicate that CML7 belongs to the group of classical  $\text{Ca}^{2+}$  sensors instead of  $\text{Ca}^{2+}$ -buffering proteins, which experience only minor structural variations and bind very tightly to  $\text{Ca}^{2+}$ . Working as a  $\text{Ca}^{2+}$  sensor, CML7 is therefore expected to interact with and/or regulate specific target proteins. Melittin, a small peptide of 26 residues, is widely used to study protein–peptide interactions in  $\text{Ca}^{2+}$ -binding proteins, including CaM, and can provide insight into target recognition mode of CML7 [Scaloni 1998 and Weljie 2000]. Our results showed that CML7 has a high affinity interaction with melittin, in a  $\text{Ca}^{2+}$ -dependent manner. Binding of melittin to  $\text{Ca}^{2+}$ -AtCaM1 and  $\text{Ca}^{2+}$ -CML7 occurred in an  $\alpha$ -helical conformation, and the interaction involved the Trp residue of the peptide which becomes solvent shielded. Therefore, we can envisage that CML7 may be involved in  $\text{Ca}^{2+}$  mediated protein-protein interactions even if its physiological partner/s is still unidentified.

The binding of melittin definitely places CML7 in the group of  $\text{Ca}^{2+}$  sensors, but more insightful information will require the identification and examination of its physiological targets. Moreover, although important clues can be obtained by the analysis of CML complexes with short target peptides, these complexes do not always perfectly reflect the biophysical and structural properties of the full-length

protein. Therefore, in the absence of physiologically significant functional information, the structural and biochemical data should be interpreted with caution. An involvement of CML7 in root development has been proposed, even if the downstream targets of CML7 have not yet been identified. Despite the large number of molecular interactions that take place with members of the CaM superfamily, specific differences in the structure of physiological targets are likely to ensure specificity of interactions. However, it is worth noting that the structural properties and the mechanism of action of CML7 and AtCaM1, as predicted from our results, seem to be very similar. Thus, it is unclear how they can have non-overlapping physiological functions. Clearly, for Ca<sup>2+</sup> sensors it is crucial to consider the Ca<sup>2+</sup> binding properties in connection with the target. Moreover, factors other than Ca<sup>2+</sup> binding, such as relative abundance of Ca<sup>2+</sup> sensor and target, tissue expression, and subcellular localization as well phosphorylation may have significant implications for target regulation mechanism. However, only the identification of the physiological targets of CML7 will allow us to obtain deeper insights into the cellular events that this protein participates in and to determine its selectivity and mechanism of action.

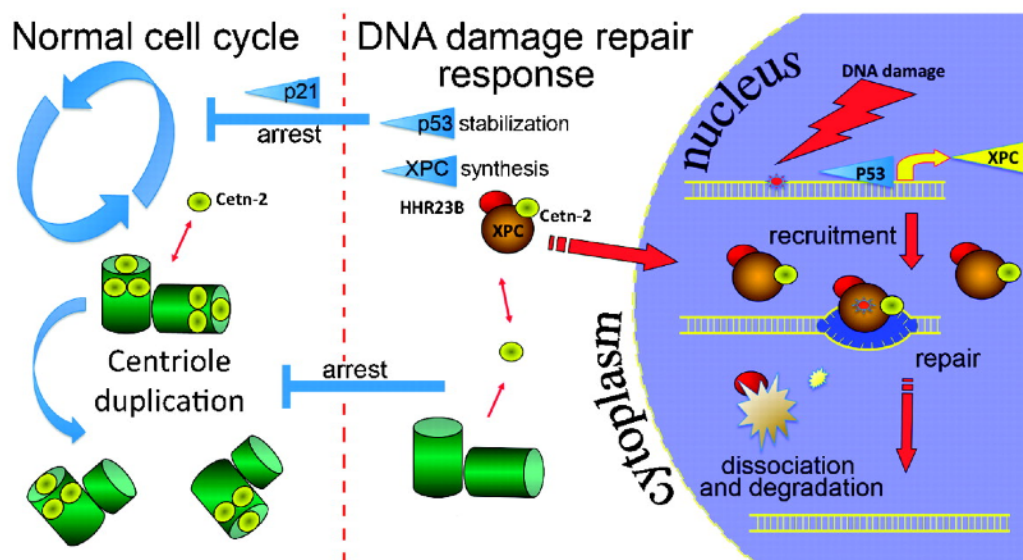
# Chapter 5

## *Calmodulin Like Protein 19 (CML19) from Arabidopsis thaliana*

## 5.1 INTRODUCTION

### 5.1.1 The role of centrin 2 in the Nucleotide Excision repair system

The Nucleotide Excision repair system (NER) is a well studied DNA repair mechanism. Nowadays, two sub-pathways are known for this system: the transcription-coupled NER and the global genome repair (GG-NER) [Dantas 2011]. The first one acts rapidly to remove lesions that block transcription, whereas the second one acts slowly but genome wide in the removal of bulky DNA lesions such as UV-induced thymidine dimers or cisplatin-DNA adducts to guarantee correct cell replication [Batty 2002 and Lindahl 1993]. Briefly, excess of DNA damages interrupts the normal cell cycle progression and induce the expression of XPC, which in complex with hRad23B and centrin 2 colocalizes at the nucleus and recognizes DNA helix distortions[Sugasawa, 1998], in parallel due to depletion of cytoplasmic centrin 2 from the centrosome system, necessary for cell replication, XPC induce the arrest of cell duplication cycle (Figure 18) [Batty 2000, Araki 2001 and Popescu 2003]. Upon DNA binding, the XPC/hRad23B/centrin2 complex recruits the DNA helicases XPD and XPB, which unwind the DNA, allowing the assembly of the repair system machinery [Ng 2003, Fitch 2003 and van der Spek



**Figure 18 [Acu 2010]:** Schematic representation of the role of Human Centrin 2 and XPC in the DNA damage response and centrosome homeostasis.

1996]. Thus the recruitment and formation of the XPC/hRad23B/centrin2 complex represent a crucial and rate-limiting step for the correct GG-NER progression.

Centrins are small, acidic, highly conserved proteins required for centrosome homeostasis [Salisbury 2002] and XPC/hRad23B/centrin2 complex stabilization. They belong to the EF-hand superfamily of calcium ( $\text{Ca}^{2+}$ )-binding proteins and were firstly discovered in green algae, associated with basal bodies and contractile striated fibers [Huang 1988; Salisbury 1984 and Weber 1994]. Subsequently, centrins have been found in many other organisms, including yeast, plants, and humans [Baum 1986, Madeddu 1996 and Del Vecchio 1997]. The discovery of new centrin targets is supporting their involvement in various biological processes, such as DNA repair (XPC/RAD4) [Liang 2006, Thompson 2006 and Marteiijn 2014], centrosome duplication (SFI1)[Kilmartin 2003] and mRNA export (SAC3) [Grecu 2014]. Centrins are closely related to CaMs in terms of domain composition (EF-hands),  $\text{Ca}^{2+}$  sensing properties and conformational changes upon calcium binding. Usually centrins possess an N-terminal ~ 20 amino acids disordered region, which is involved in protein self-assembly and subcellular localization; moreover, the C-terminal region, which is mainly involved in target recognition, is characterised by the presence of a conserved Tyr or Phe [Friedberg 2006 and Hu 2004].

### 5.1.2 Centrins in *Arabidopsis thaliana*, the role of CML19

*Arabidopsis thaliana* possesses two centrin isoforms, i.e., centrin 1 and 2 (respectively CML20 and CML19). Centrin 1 is suggested to be involved in the organization of microtubule centers by interacting with TON1, a component of plant cell cytoskeleton [Azimzadeh 2008], whereas CML19 is the plant homologue of the human centrin 2. CML19 is predicted to have four functional EF-hand in the primary sequence and share 42% sequence identity with CaM2 and 56% with its human homologue. As observed for other centrins, CML19 possesses an N-terminal region that shows a high degree of disorder as predicted by the GeneSilico MetaDisorder Service analysis; moreover, the distal EF-hand harbors one Tyr residue and two Phe residues, which might be involved in target recognition.

As for human centrin 2, CML19 function is related to Nucleotide Excision Repair (NER) via the interaction with RAD4 protein, the plant homologue of human XPC [Molinier 2004]. CML19/RAD4 interaction has been observed to be strongly  $\text{Ca}^{2+}$

dependent and to act in the C-terminal region of CML19; moreover, it has been shown that UV-C treatment induced centrin translocation from the cytoplasm into the nucleus, where it is supposed to take part in the NER complex [Liang 2006].

While the physiological role of CML19 is well understood, the molecular mechanism and the centrin binding site on RAD4 target are still missing [La Verde (bis) 2018]; to this aim, this Chapter of my thesis reports an examination of the biophysical and structural properties of CML19 and the identification of the specific CML19-binding site at the C-terminus of its physiological partner RAD4.

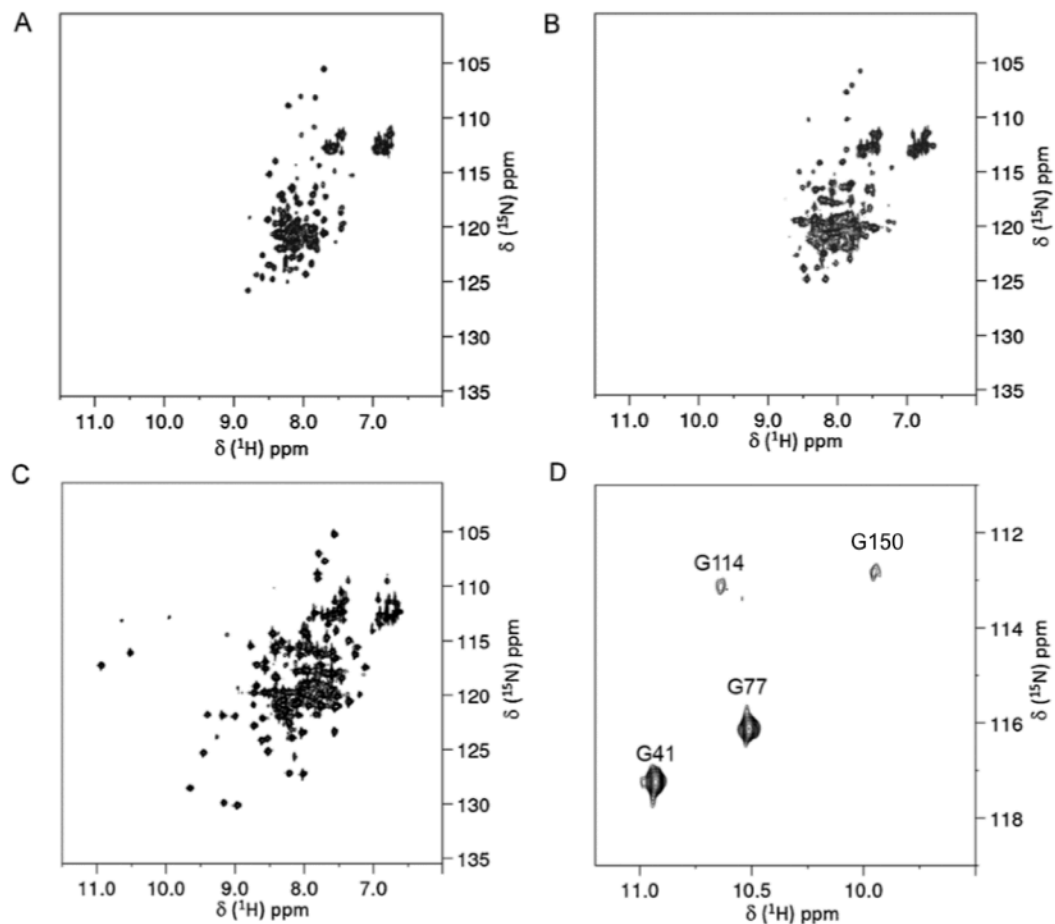
## **5.2 RESULTS AND DISCUSSION**

### **5.2.1 Ca<sup>2+</sup>-induced conformational changes**

To avoid experimental difficulties associated with the aggregation propensity of recombinant full length CML19, which was observed also for other centrin [Thompson 2006 and Charbonnier 2007], we focused our studies on a truncated variant of CML19 missing the first 19 amino acids (hereafter referred to as CML19).

The conformational changes of CML19 upon metal binding were first investigated via NMR spectroscopy which is a very informative technique to evaluate secondary and tertiary structure rearrangements of CBPs upon metal binding. Figure 19 shows the comparison of the two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC spectra of CML19 in the presence of saturating amounts of EGTA, Mg<sup>2+</sup> and Ca<sup>2+</sup>. The NMR spectrum of the apo-protein (Figure 19 A) is characterized by the presence of few peaks, many of which are broadened beyond detection, and a narrow range of chemical shift dispersion, especially in the proton dimension, indicating that, in the absence of Ca<sup>2+</sup>, the protein is characterized by conformational heterogeneity. The addition of Mg<sup>2+</sup> to apo-CML19 (Figure 19 B) did not alter substantially the HSQC spectrum and no peaks in the low field-shifted <sup>1</sup>H region are visible, likely suggesting the absence Mg<sup>2+</sup> binding to the protein. On the contrary, when saturating amounts of Ca<sup>2+</sup> were added to the apo-protein (Figure 19 C), the NMR spectrum changed dramatically, showing an increased number of well-dispersed peaks, meaning that the protein undergoes a conformational rearrangement with a predominant ordered structure. Four peaks appeared in the low field-shifted <sup>1</sup>H region at about 10 ppm in

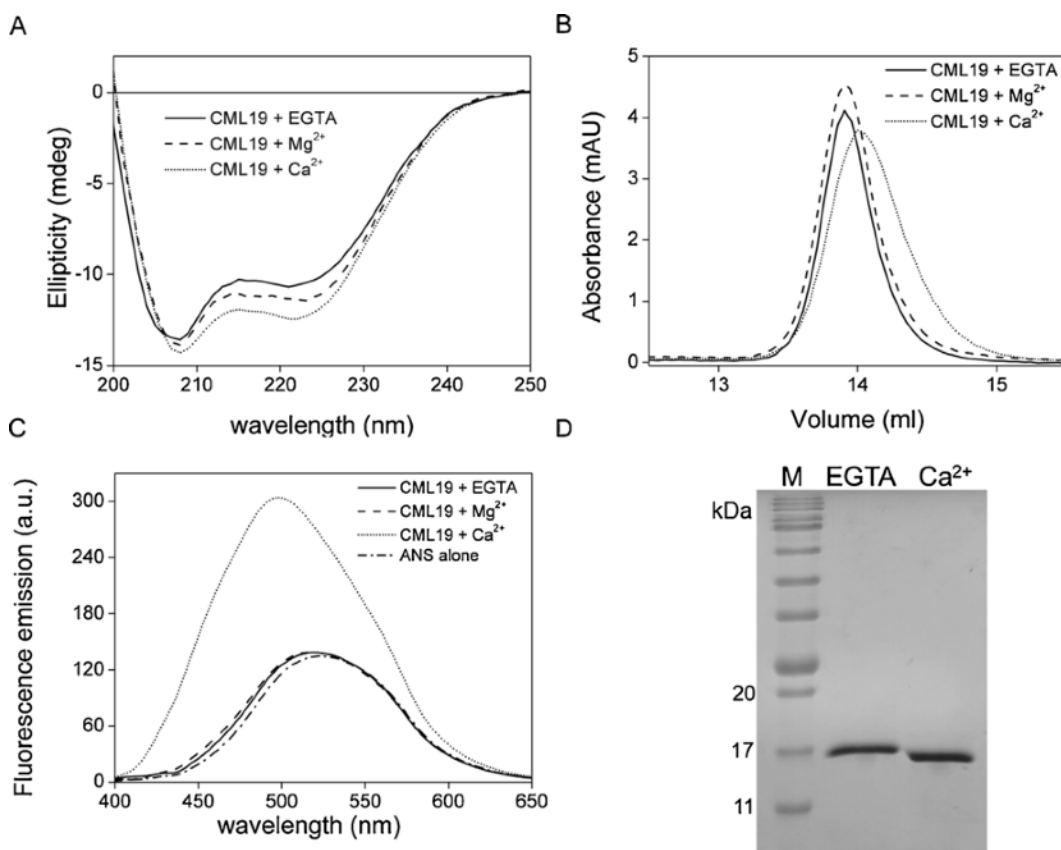




**Figure 19:** Two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of  $^{15}\text{N}$ -CML19 in its apo-state (A),  $\text{Mg}^{2+}$ - (B),  $\text{Ca}^{2+}$ - (C) bound forms, and the zoom of the glycine downfield peaks of CML19 in complex with  $\text{Ca}^{2+}$  ions (D). The assignment of the glycines of the EF-hands is indicated by the position number of the residue.

the presence of  $\text{Ca}^{2+}$  (Figure 19 D). These signals were specifically assigned to the corresponding Gly6 by using standard triple resonance NMR experiments acquired on  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled samples. In particular, the peaks at 10.9, 10.5, 10.63 and 9.95 ppm were assigned to the Gly6 of the four EF-hands: Gly41 (EF-1), Gly77 (EF-2), Gly114 (EF-3), and Gly150 (EF-4) respectively (Figure 19 D).

Far-UV CD spectra of CML19 showed the typical shape and amplitude of a protein with a high amount of  $\alpha$ -helical structure, (minima at 208 and 222 nm, Figure 20 A).  $\text{Ca}^{2+}$  binding caused an increase in molar ellipticity, leading to a switch in the  $\theta_{222}/\theta_{208}$  ratio from 0.79 in the apo-form to 0.87 in the  $\text{Ca}^{2+}$ -bound form, which indicates helical reorientation within the tertiary structure [Zhang 1995 and Gagne 1994]. The analysis of the  $\theta_{222}/\theta_{208}$  ratio also indicated that the shape of the spectrum



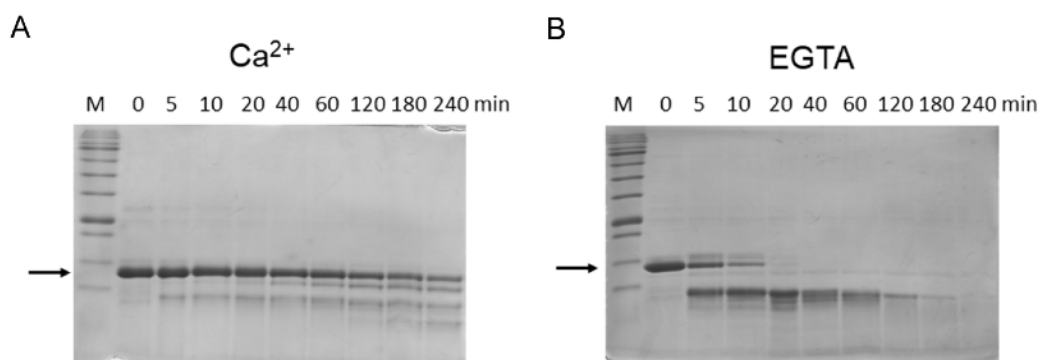
**Figure 20:** Metal-induced conformational changes to CML19. (A) Far-UV CD spectra of CML19 in the presence of EGTA (solid line), MgCl<sub>2</sub> (dashed line), and CaCl<sub>2</sub> (dotted line). (B) Size exclusion chromatography analysis of CML19 on a Superose 12 column 10/300GL. Elution profiles were recorded in the presence of EGTA (solid line), MgCl<sub>2</sub> (dashed line) or CaCl<sub>2</sub> (dotted line). (C) ANS fluorescence spectra of CML19 in the presence of EGTA (solid line), MgCl<sub>2</sub> (dashed line), and CaCl<sub>2</sub> (dotted line). The spectrum of ANS alone (dash dotted line) is also shown. (D) SDS-PAGE mobility shift of CML19 in the presence of EGTA or CaCl<sub>2</sub>.

corresponding to the Ca<sup>2+</sup>-bound is distinguished from that of the Mg<sup>2+</sup>-bound form. In fact, the spectra exhibited a more similar shape for apo- and Mg<sup>2+</sup>-bound CML19, as quantitatively demonstrated by nearly the same  $\theta_{222}/\theta_{208}$  ratio (0.79 versus 0.81, respectively).

These CD results are consistent with NMR spectra, as both analyses indicated that Mg<sup>2+</sup> did not induce significant conformational changes, and the apo- and Mg<sup>2+</sup>-bound CML19 species possess highly similar structural properties. It is likely that the minor changes observed on both NMR and CD spectra upon Mg<sup>2+</sup> addition to the

protein are the result of non-specific binding of  $Mg^{2+}$  ions, perhaps to sites that differ from the main  $Ca^{2+}$  binding sites.

Improved protein stability in the presence of  $Ca^{2+}$  was tested for CML19 by limited trypsin proteolysis. In Figure 21 we can observe that the protein underwent complete digestion in few minutes in the absence of  $Ca^{2+}$ , whereas in its presence CML19 was more stable to proteolysis (up to 60 min).



**Figure 21:** SDS-PAGE analysis of CML19 digestion by trypsin in the presence of EGTA or  $CaCl_2$ , after incubation with trypsin 1:500 (w/w) for several time intervals. Lane M represents molecular weight markers.

SEC experiments revealed that recombinant CML19 in the presence of EGTA is characterised by a relatively high  $R_s$  ( $26.06 \pm 0.02 \text{ \AA}$ ; CML19 MW= 18.09 KDa), which, as previously discussed (paragraph 4.2.2), is likely because of the elongated shape of  $Ca^{2+}$  sensor proteins. Addition of  $Mg^{2+}$  did not induce variations in the protein  $R_s$  ( $26.00 \pm 0.09 \text{ \AA}$ ), whereas  $Ca^{2+}$  binding to CML19 caused a decrease in its  $R_s$  to  $25.51 \pm 0.02 \text{ \AA}$ , confirming the conformational changes observed with other techniques (Figure 20 B).

Moreover, recombinant CML19 shifted to a faster-migrating conformation in the  $Ca^{2+}$ -bound form during SDS-PAGE (Figure 20 D).

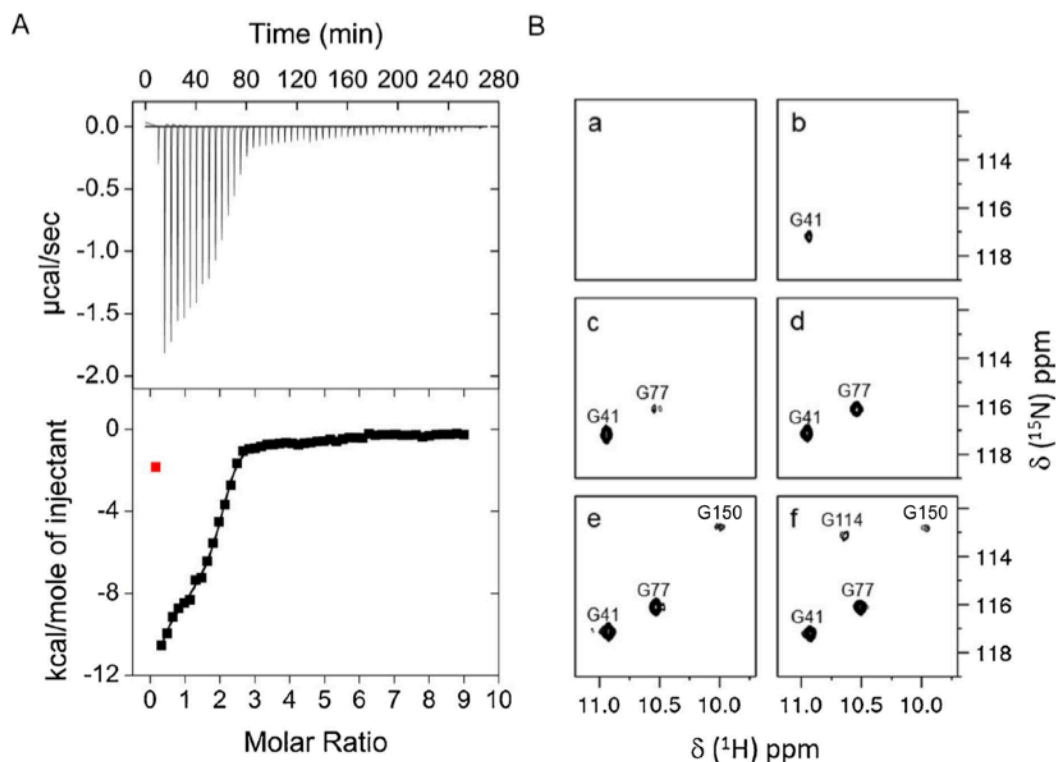
The  $Ca^{2+}$ -induced exposure of hydrophobic surfaces which is a crucial event for target interaction was tested using the the ANS fluorophore (Figure 20 C). Apo-CML19 does not possess any degrees of hydrophobicity, as observed by comparing the emission spectra of ANS alone with that of CML19 in the presence of EGTA. When  $Ca^{2+}$  binds to CML19, a  $\sim 20 \text{ nm}$  blue shift in the ANS maximum emission (max  $\sim 500 \text{ nm}$ ) occurred with a 2.2-fold increase in fluorescence intensity in comparison to the spectrum of the ANS-CML19 complex in the presence of EGTA. In contrast, binding of  $Mg^{2+}$  to the apo-protein led to an emission spectrum that was approximately identical to that of ANS-CML19 in the presence of EGTA,

therefore indicating a  $\text{Ca}^{2+}$  and not  $\text{Mg}^{2+}$ -dependent exposure of hydrophobic patches.

As we can deduct from the above analysis,  $\text{Mg}^{2+}$  seems to have no evident effect on the protein, leading to the conclusion that CML19 is a  $\text{Ca}^{2+}$  signaling selective protein that is not able to bind  $\text{Mg}^{2+}$  within its EF-hand domains.

### 5.2.2 Calcium binding to CML19

We combined ITC and NMR analysis to investigate the intrinsic  $\text{Ca}^{2+}$  binding properties of CML19.



**Figure 22:**  $\text{Ca}^{2+}$  binding to CML19. (A) Representative raw trace minus baseline (top) of the calorimetric titration of  $\text{CaCl}_2$  into apo-CML19 and integrated binding isotherm (bottom). Curve fitting was performed using a four-site sequential binding model. (B) Downfield region of  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of CML19 recorded as a function of increasing  $\text{Ca}^{2+}$  concentration. The molar ratio of  $\text{Ca}^{2+}$  /protein was 0 (a), 0.1 (b), 0.25 (c), 0.67(d), 8 (e), 16 (f).

A representative ITC isotherm is shown in Figure 22 A and the thermodynamic parameters of the interaction are listed in Table 5. Titration of  $\text{CaCl}_2$  into apo-

CML19 results in an exothermic curve demonstrating the binding of four Ca<sup>2+</sup> ions to CML19, consistently with the sequence-based prediction of PROSITE-ProRule annotation [de Castro 2006]. The four Ca<sup>2+</sup> binding sites were clustered into two groups, one containing two higher affinity sites (Kd1 = 3.5 ± 0.8 μM, Kd2 = 2.4 ± 0.4 μM) and a second group containing two sites with lower affinity (Kd3 = 71 ± 19 μM, Kd4 = 270 ± 58 μM).

	K <sub>A</sub> (M <sup>-1</sup> )	ΔH (kcal mol <sup>-1</sup> )
Site 1	2.9 E5 ± 7.1 E4	-1.2 E4 ± 1.9 E3
Site 2	4.1 E5 ± 6.8 E4	-4.4 E3 ± 0.6 E3
Site 3	1.4 E4 ± 3.8 E3	-2.1 E3 ± 0.9 E3
Site 4	3.7 E3 ± 0.8 E3	-1.9 E3 ± 0.7 E3

**Table 5:** Thermodynamic parameters for Ca<sup>2+</sup> binding to apo-CML19. Sites 1 and 2 are located at the N-terminal lobe and sites 3 and 4 at the C-terminal, as explained in the main text.

Furthermore, the calculation of cooperativity within the two groups (Table 6) suggested the presence of positive cooperativity only in the high affinity sites as the calculated  $-\Delta\Delta G$  in sites 3 and 4 was close to be zero.

	ΔG <sup>1,2</sup> (KJ mol <sup>-1</sup> ) <sup>a</sup>	-ΔΔG <sub>η=1<sup>1,2</sup></sub> (KJ mol <sup>-1</sup> ) <sup>b</sup>	ΔG <sup>3,4</sup> (KJ mol <sup>-1</sup> ) <sup>c</sup>	-ΔΔG <sub>η=1<sup>3,4</sup></sub> (KJ mol <sup>-1</sup> ) <sup>d</sup>
<b>Apo-CML19</b>	-63.2 ± 0.4	4.3 ± 0.4	-44.0 ± 0.5	0.1 ± 0.5

**Table 6:** Free energy of binding of the first two (ΔG<sup>1,2</sup>) and last two (ΔG<sup>3,4</sup>) Ca<sup>2+</sup> ions to apo-CML19 and the lower limit of the cooperativity within each pair (-ΔΔG<sub>η=1</sub>), as described in Chapter 4.2.3.

Next, we monitored the downfield shift <sup>1</sup>H resonances of Gly6 of the EF-hand binding loops via Ca<sup>2+</sup> titration of apo-CML19 through the acquisition of <sup>1</sup>H-<sup>15</sup>N HSQC spectra. As shown in Figure 22 B, in the Ca<sup>2+</sup> titration of CML19 the NH peaks of the Ca<sup>2+</sup>-bound N-terminal Gly6 residues of CML19 (Gly41 and Gly77) develop firstly, while the other two glycine peaks (Gly114 and Gly150) appear at higher Ca<sup>2+</sup> concentrations and allowed us to assign the dissociation constants to specific lobes of CML19: Kd1 and Kd2 to the N-terminal lobe (higher affinity Ca<sup>2+</sup>-binding sites), while Kd3 and Kd4 to the C-terminal EF-hands (lower affinity Ca<sup>2+</sup>-binding sites).

### 5.2.3 Interaction with RAD4 peptide

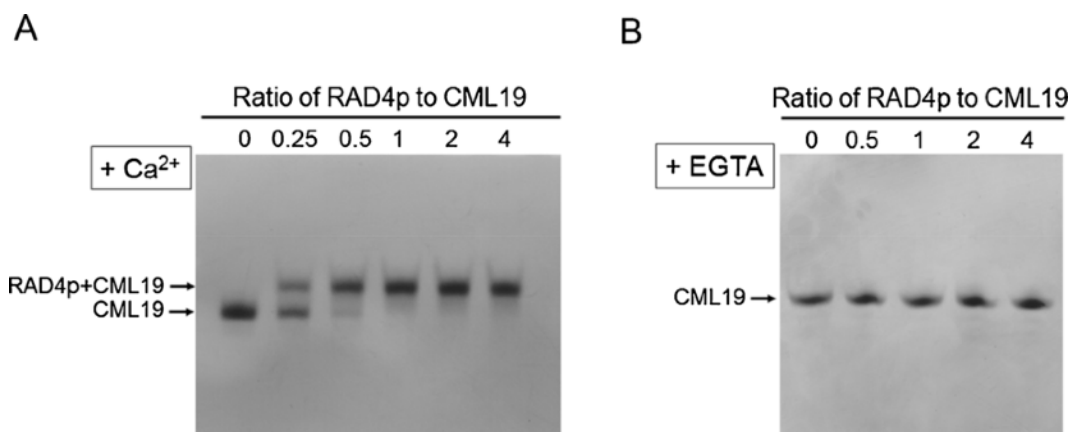
Although the involvement of CML19 in the RAD4 plant recognition system for functional NER has been described [Liang 2006], the structural and biophysical data for this interaction was almost non-existent and there was no clear indication about which specific recognition sites on both RAD4 (865 residues) and CML19 were involved in the interaction. Generally, the binding sites in the centrin targets possess some common features: they are relatively small (15–30 residues long), able to adopt, upon centrin binding, an amphipathic  $\alpha$ -helical structure, and they generally possess a hydrophobic W<sup>1</sup>-L<sup>4</sup>-L<sup>8</sup> triad (Figure 23). The human centrin 2 binding site on XPC has been well characterised (Ca<sup>2+</sup> dependent interaction,  $K_a = 2.2 \times 10^8 \text{ M}^{-1}$ , stoichiometry 1:1) [Charbonnier 2007 and Popescu 2003] and identified as the hydrophobic W<sup>848</sup>-L<sup>851</sup>-L<sup>855</sup> triad. Therefore, we have used bioinformatics methods (i.e. the CALMODULIN TARGET DATABASE [Yap 2000]) to analyse RAD4 sequence and search for a segment that could work as a potential centrin binding site. We found a putative consensus sequence with the hydrophobic triad and sequence homology to the HsCen2 binding region of human counterpart XPC comprising the region Arg<sup>762</sup>-Asn<sup>778</sup> on RAD4; thus, the corresponding peptide (RAD4p,

	1	4	8
<b>AtRad4</b>	<b>R</b> WY <b>Q</b> LLSS <b>I</b> LTRERLKN		
<b>XPC</b>	N <b>W</b> KLLAK <b>G</b> LLIRERLKR		
<b>Kar1</b>	KKRELIESK <b>W</b> H <b>R</b> LL <b>F</b> H <b>D</b> KK		
<b>Melittin</b>	QQRKRKI <b>W</b> S <b>I</b> LAP <b>L</b> GTTLVKLVAGIG		
<b>hSfi1</b>	TV <b>W</b> A <b>Q</b> LAR <b>H</b> L <b>V</b> SHQHHL <b>D</b> AR		

**Figure 23:** Sequences of centrin-binding motifs. The bulky hydrophobic side chains anchoring the peptide into the hydrophobic pocket of the protein are in bold. AtRAD4 is the target of CML19, XPC and hSfi1 are targets of human centrans, whereas Kar1 is a target of yeast centrans and melittin is a model target peptide (Chapter 4.2.4). Notably, the sequences of hSfi1 and melittin are disposed in a reverse orientation (8,4,1) as compared to the centrin-binding motif in XPC.

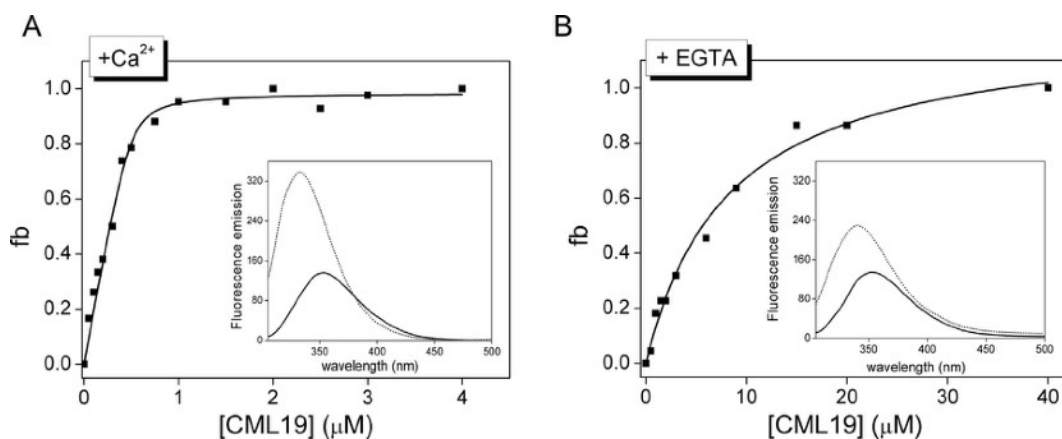
R<sup>762</sup>**WYQLLSSILTRERLKN**<sup>778</sup>) was synthesized and used for experiments.

The interaction was firstly examined by native PAGE evaluating mixtures of CML19 and increasing amounts of RAD4p. When RAD4p and CML19 were mixed in the



**Figure 24:** Non-denaturing-polyacrylamide gel mobility shift characterization of the interaction RAD4p/CML19 in the presence of (A) 5 mM CaCl<sub>2</sub> and (B) 5 mM EGTA. 20 μM CML19 was incubated with increasing concentrations of RAD4p (5–80 μM).

presence of Ca<sup>2+</sup>, an extra band appeared on the native gel likely representing the CML19/RAD4p complex. Notably, at a 1:1 protein/peptide molar ratio, the band of free CML19 disappeared (Figure 24 A), suggesting the formation of a well-defined 1:1 complex. Under the same experimental conditions this band was absent in the EGTA-containing sample, indicating that the binding is primarily Ca<sup>2+</sup> dependent (Figure 24 B).



**Figure 25:** Representative fluorescence titration of 1 μM RAD4p with CML19 in the presence of (A) 5 mM CaCl<sub>2</sub> and (B) 5 mM EGTA. fb represents the fraction bound. Insets, Trp fluorescence emission spectra of RAD4p alone (solid line) and RAD4p/CML19 complex (dotted line) in the presence of (A) 5 mM CaCl<sub>2</sub> and (B) 5 mM EGTA.

Because CML19 lacks Trp residues in its primary sequence, the interaction between CML19 and RAD4p was further confirmed by following changes in the fluorescence

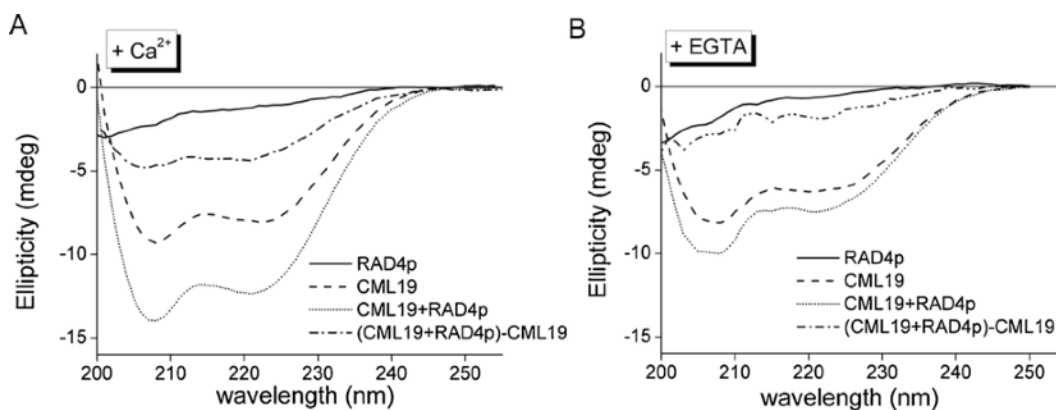
emission of Trp763 of RAD4p upon titration of CML19, giving information on the stoichiometry, binding mode and strength (Figure 25). In the presence of  $\text{Ca}^{2+}$  the fluorescence spectrum of the Trp residue in the peptide changed substantially upon complex formation, as observed by the significant blue shift (20 nm, from 353 to 333 nm) and a 2.5-fold increase in intensity (Figure 25 A, inset). Solvent-exposed Trp side chains in a polar environment typically possess a fluorescence emission maximum at 350 nm, and therefore these data clearly indicate that RAD4p interacts with CML19 and that the interaction leads to the embedding of Trp in a non-polar environment, similarly to the interaction of HsCen2 to XPC.

The fluorescence titration curve revealed a fast increase in fluorescence intensity that confirmed binding with a 1:1 stoichiometry and gave a  $K_d$  of  $54 \pm 14$  nM (Figure 25 A). RAD4p displayed a different behavior when interacting with CML19 in the presence of EGTA, as the fluorescence increase was smaller (1.6-fold increase) and less blue shifted (12 nm, from 353 to 341 nm) (Figure 25 B, inset). Importantly, the fluorescence response in the CML19 titration performed in the presence of EGTA yielded a  $K_d$  value of  $8 \pm 1$   $\mu\text{M}$ , thus indicating an affinity for CML19/RAD4p complex that is about 150-fold lower than in the presence of  $\text{Ca}^{2+}$  (Figure 25 B). The enhancement of the peptide binding affinity by about two orders of magnitude in the presence of  $\text{Ca}^{2+}$  clearly shows that the cation activates CML19 for the binding to the target.

The ability of CML19 to bind RAD4p was also analyzed by far-UV CD spectroscopy in order to describe the conformational properties of the CML19/RAD4p interaction and evaluate the folding of RAD4p into an amphipathic helix when binding to  $\text{Ca}^{2+}$ -saturated CML19. Under both  $\text{Ca}^{2+}$ -depleted and  $\text{Ca}^{2+}$ -saturating conditions, free RAD4p in the aqueous solution has a spectrum with an ellipticity close to zero and a negative peak at  $\sim 203$  nm, reflecting that RAD4p has random coil type secondary structure by itself. (Figure 26 A, solid line),  $\text{Ca}^{2+}$ -bound CML19 exhibited the spectra of a well-folded protein with high  $\alpha$ -helical content (Figure 26 A, dashed line). The addition of RAD4p to CML19 at 1:1 molar ratio in the presence of  $\text{Ca}^{2+}$  caused a significant increase in the intensity of the dichroic signal ( $\sim 50\%$ ) with conservation of the relative intensity of the different bands (Figure 26 A, dotted line). This contribution likely comes from the peptide rearrangements from a random coil to  $\alpha$ -helical structure following interaction with CML19. CD spectra taken in the presence of EGTA led to different results, with a modest change in the



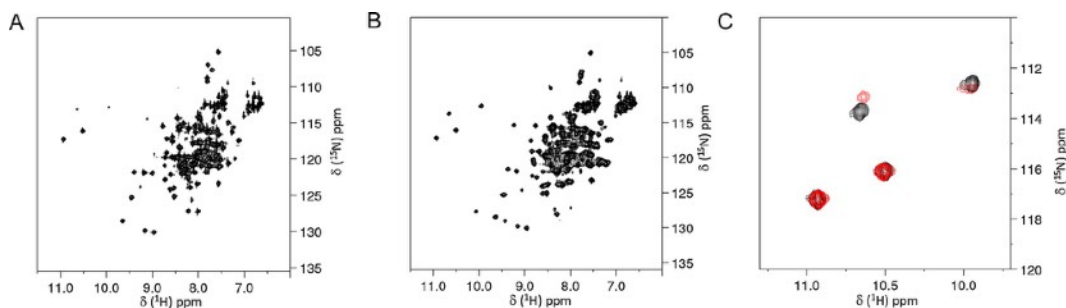
far-UV CD signal upon addition of RAD4p to CML19 (Figure 26 B), confirming that CML19 binds to RAD4 primary in a  $\text{Ca}^{2+}$  dependent mode to form a 1:1 complex.



**Figure 26:** Far-UV CD spectra changes in secondary structure for CML19/RAD4p complex in the presence of  $\text{CaCl}_2$  (A) and EGTA (B). Far-UV CD spectra of 5  $\mu\text{M}$  RAD4p alone (black line), 5  $\mu\text{M}$  CML19 alone (dashed line), and RAD4p/CML19 complex (dotted line) at 1:1 molar ratio. The CD spectrum resulting from subtraction of the spectrum of CML19/RAD4p complex from that of CML19 alone was also shown (dash dotted line).

Information on the structural and dynamic properties of the complex formed by  $\text{Ca}^{2+}$ -bound CML19 and RAD4p were also obtained by NMR experiments. Upon addition of the peptide to the  $\text{Ca}^{2+}$ -bound CML19 (Figure 27), the NMR spectrum showed important variations as some peaks experienced chemical shift variation while new peaks appeared. The RAD4p binding-induced changes in NMR spectrum confirmed that, in agreement with previous experiments, the interaction between the protein and peptide results in well-organized, unique, and stable structure. Notably, detailed analysis of the NMR spectra revealed that the resonances belonging to glycines Gly41 (EF-1) and Gly77 (EF-2) of CML19 remain unperturbed upon protein-peptide interaction, meaning that the CML19 N-terminal domain is not critically involved in RAD4p binding. On the contrary, the other two peaks corresponding to the Gly114 of the EF-3 and to the Gly150 of EF-4, are characterized by a clear chemical shift variation and an increase in intensity. These observations strongly suggest that the changes in the NMR spectrum originate from modifications in the C-terminal part of CML19. This finding is in accordance with

the notion that one of the main centrin specific binding features is related to the principal role played by the C-terminal domain in target binding.



**Figure 27:** Two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectra of Ca<sup>2+</sup>-CML19 in the absence (A) and in the presence of RAD4p at a 1:1 molar ratio (B). (C) Zoom of the glycinamide downfield peaks of the overlaid <sup>1</sup>H-<sup>15</sup>N HSQC spectra recorded on Ca<sup>2+</sup>-CML19 before (red) and after the addition of RAD4p (black).

## 5.4 CONCLUSIONS

The purpose of this Chapter was to provide a detailed molecular description of the binding of Ca<sup>2+</sup> to CML19 and subsequent protein conformational changes. In addition, we underline the significance of the interaction between CML19 and its physiological binding partner RAD4 by identifying the CML19 binding site on RAD4.

Similar to algae centrans SdCen and CrCen [Hu 2004 and Radu 2010], CML19 can specifically bind four Ca<sup>2+</sup> ions, two with high affinity in the N-lobe and two with moderate/low affinity in the C-lobe. Upon Ca<sup>2+</sup> binding, CML19 undergoes to structural rearrangements peculiar of the Ca<sup>2+</sup> sensors behaviour of, acquiring a more stable and ordered conformation that results in the exposure of a hydrophobic regions that is essential for target protein recognition [Ikura, 1992; Meador, 1992; Hoeflich, 2002 and Ishida, 2006].

CML19 is known to bind RAD4 during GG-NER process. Considering the direct interaction between CML19 and RAD4, we identified the CML19-binding site located at the C-terminus of RAD4. Centrans are known to interact with their target through the recognition of the well conserved binding motifs W<sup>1</sup>-L<sup>4</sup>-L<sup>8</sup>. The importance of the three residues in the interaction with human centrin have been

studied by Charbonnier and coworkers [Charbonnier, 2007], by single aminoacids mutations. Upon W<sup>1</sup> to A substitution the interaction was significantly reduced, the L<sup>4</sup> to A substitution had a less dramatic effect, while L<sup>8</sup> to A substitution did not cause changes. In RAD4 the final Leu is substituted by a chemically equivalent Ile; whereas in the yeast centring binding region of Kar1, L<sup>8</sup> is by the a negatively charged residue (Asp) (Figure 23), suggesting that the hydrophobic dyad W1-L4 represents the core of the centrin's binding motif.

RAD4p/CML19 complex is formed specifically in the presence of Ca<sup>2+</sup>, resulting in the formation of a high affinity 1:1 complex in which the W<sup>1</sup> of the target becomes highly buried in a hydrophobic pocket, as observed by variations in the Trp fluorescence emission.

Detailed analysis of the NMR spectra of Ca<sup>2+</sup>-saturated CML19 before and upon addition of RAD4p, revealed that the most significant variations in the chemical shift peaks intensity are seen for the residues belonging to the C-terminal domain of Ca<sup>2+</sup>-CML19, while resonances unperturbed by peptide binding are attributable mainly to the N-domain. Our speculation is consistent with the notion that the C-terminal domain of centrin acts as an anchor point for target proteins in the formation of many centrin-target complexes [Thompson 2006, Martinez-Sanz 2006, Hu 2003, Hu 2004, Azimzadeh 2009, La Verde 2018], even though the macroscopic binding affinity for Ca<sup>2+</sup> of this domain measured *in vitro* is often outside the range for a normal Ca<sup>2+</sup> sensors [Hu 2004, La Verde 2018]. The binding of the peptide mainly to the C-lobe could also allow the N-terminal domain of CML19, comprising two high affinity Ca<sup>2+</sup> binding sites, to interact with other target proteins (or even with a distal region of RAD4), as already suggested for yeast and green algal centrin's [Hu 2004, Radu 2010, Sheehan 2006].

Deeper insights into the molecular basis for CML19 activity would be given by studying the interaction with different types of natural targets, as well as the experimental determination of the Ca<sup>2+</sup>-CML19/RAD4p complex structure, which still remain an urgent priority in the field.

# Chapter 6

*A method for metal/protein  
stoichiometry determination using  
Thin-Film Energy Dispersive X-  
ray Fluorescence*

## **6.1. INTRODUCTION**

### **6.1.1 The commonly used techniques for metallic cofactor identification and quantification and their limitations**

Ca<sup>2+</sup> binding proteins (CBP) belong to the larger superfamily of metalloproteins, which are believed to comprehend from one-quarter to one-third of all existing proteins. The expanding study of this protein superfamily is rising a demand for convenient methods to identify and quantify the inorganic species present in proteins [Waldron 2009].

Generally, for the determination of metal cofactors the most commonly used techniques include inductively coupled plasma mass spectrometry (ICPMS), optical emission spectrometry (ICP-OES) and atomic absorption spectroscopy (AAS) [Rio-Echevarria 2019 and Garicia 2006].

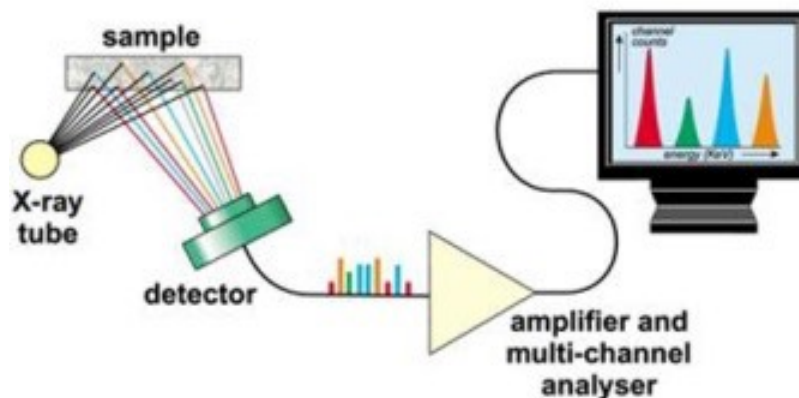
Since CBPs are complex systems with multiple metal-binding sites, cooperativity and a wide range of metal affinity, the unequivocally identification of metal binding stoichiometry can be very tricky. ITC and NMR are ideal techniques to understand the intricacies of Ca<sup>2+</sup>-binding proteins (as we observed in Chapters 4 and 5), however these techniques are time and sample consuming and/or require expensive labelling.

To overcome these limitations, in collaboration with the University of Manchester (Prof. Lu Shin Wong group) we applied the Energy Dispersive X-ray Fluorescence technology (EDXRF) to determine the metal/protein stoichiometry in CBPs.

### **6.1.2 XRF-based techniques**

X-ray Fluorescence technology (XRF) is a widely used technique for the elemental analysis of inorganic materials and is based on the detection of X-rays emitted from sample atoms upon excitation with X-rays of higher energy (Figure 28) [Margui 2014]. The emitted X-rays are of energies that are characteristic for each element, allowing the identification and quantification of the atomic composition of the sample. XRF is suitable for a wide range of samples (i.e. solid, liquid etc.) and does not require any sample pretreatment prior to analysis. Previous researches exploited Total-reflection XRF (TXRF) for the qualitative and quantitative analysis of metal

cofactors in proteins [Szoboszlai 2009, Mertens 2001 and Wittershagen 1998], however, due to its complex equipment, this technique is not easily available [Garman 2005]. On the other hand, energy dispersive X-ray fluorescence (EDXRF) has relatively higher detection limits compared to TXRF but, due to its simpler design, requires lower costs and maintains the characteristics of X-ray spectroscopy. In this Chapter, we explored EDXRF technology to develop a suitable method for the determination of metal content in protein samples [Fruncillo 2019]. By adapting the general principles of protein thin film TXRF, we report herein a generally applicable method that employs EDXRF for the rapid, convenient and affordable quantitation of inorganic cofactors in proteins and for accurate determination of element/protein stoichiometry within a protein complex. The method is demonstrated on a range of proteins containing Cu, Fe, Zn, Ca and the biologically relevant non-metal P. It includes examples of proteins of prokaryotic and eukaryotic origin, multimeric proteins, as well as proteins bearing post-translational modifications.



**Figure 28:** Schematic representation of the apparatus for measuring the EDXRF spectrum. Emitted X-ray from excited samples are characteristic for each element, thus allowing a precise identification and quantification.

## 6.2 SPECIFIC METHODS

All X-ray fluorescence measurements were performed using a benchtop MiniPal 4 EDXRF instrument from PANalytical (now Malvern Panalytical, Royston, UK). The instrument was equipped with an X-ray source with a rhodium anode, a high-resolution silicon drift detector, three metal filters (see Table 7) and a computer-controlled multichannel analyser system. Data collection and analysis was performed with a MiniPal 4 software ver. 3.0 (Malvern Panalytical, Royston, UK), which employs the AXIL algorithm for peak deconvolution [Vekemans 1994]. Each element was identified by the peak corresponding to its known  $K\alpha$  energy. All samples were analysed under a helium atmosphere and the EDXRF spectrum was acquired for 1600 s. Different instrument settings were used to maximise the sensitivity towards each element as reported in Table 7.

Setting reference	Element	Filter material	Filter thickness ( $\mu\text{m}$ )	Voltage (kV)	Current ( $\mu\text{A}$ )	$K\alpha$ Energy (keV)
1	Zn	Ag	100	30	300	8.639
2	Cu	Al	200	20	450	8.048
	Fe					6.404
3	Ca	Al	50	12	750	3.692
4	P	None	-	5.5	1000	2.014

**Table 7:** Preferred instrumental setting used for each element in EDXRF analysis.

All sample proteins were dissolved or buffer exchanged by centrifugal ultrafiltration in 50 mM TRIS-acetate buffer pH 7.6. Protein concentrations were determined according to the best method for each protein as reported in Table 8. Removal of residual  $\text{Ca}^{2+}$  in CBPs was achieved by saturating with a 5-fold molar excess of  $\text{CaCl}_2$  over protein, followed by a 15-min incubation at room temperature and extensive buffer exchanges (by centrifugal ultrafiltration) with 50 mM TRIS-acetate pH 7.6 previously decalcified as reported in Chapter 3.4. 8  $\mu\text{L}$  of analyte solution was pipetted on to an EDXRF sample cup covered with Mylar film (3.6  $\mu\text{m}$  thickness) and dried for 3 h at room temperature under a fume hood, resulting in a sample film approximately of 5 mm in diameter (Figure 29). Calibration curves were created using at least eight different concentrations, prepared in triplicate, for each

Protein	Approximate MW <sup>a</sup> (kDa)	Quaternary structure	Quantification method <sup>b</sup>	Molar extinction coefficient (mM <sup>-1</sup> cm <sup>-1</sup> )
<b>CotA</b>	61	Monomer	A	$\epsilon_{280} = 88.24$
<b>MvBOx</b>	60	Monomer	A	$\epsilon_{280} = 95.23$
<b>HRP</b>	43	Monomer	B	$\epsilon_{403} = 100.00$
<b>CytC</b>	12	Monomer	A	$\epsilon_{280} = 11.46$
<b>AHb1</b>	18	Dimer	B	$\epsilon_{413} = 73.80$
<b>ADH1</b>	37	Tetramer	A	$\epsilon_{280} = 48.86$
<b>CaM1</b>	17	Monomer	C	n.u.c
<b>CML7</b>	17	Monomer	C	n.u.
<b>CML19</b>	19	Monomer	C	n.u.
<b>UbiX</b>	25	Monomer	A	$\epsilon_{280} = 16.96$

**Table 8:** Main characteristics and best quantification methods of proteins analysed by EDXRF

<sup>a</sup> Refers to the molecular weight (MW) of the monomer.

<sup>b</sup> (A) UV-vis absorbance at 280 nm, (B) UV-vis absorbance of the Soret band and (C) Bradford assay.

<sup>c</sup> n.u. not used for quantification.

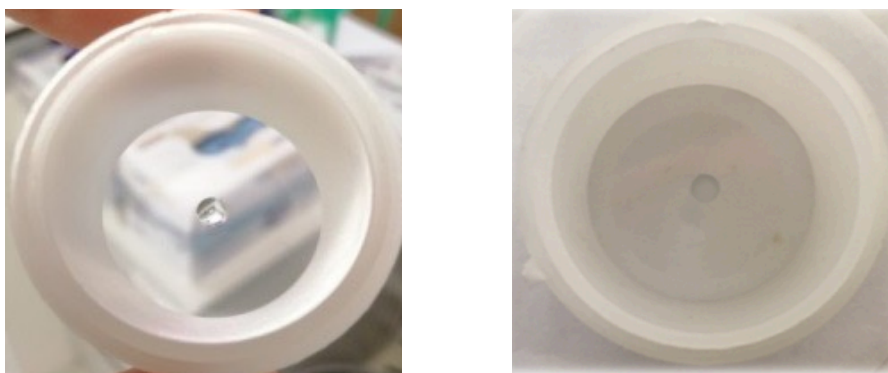
element (from 0 to 317 ppm for Cu, 279 ppm for Fe, 98 ppm for Zn, 80 ppm for Ca, and 93 ppm for P). Average values of the area under the K $\alpha$  peaks were correlated to the element concentrations. Limit of detection (LOD) and limit of quantification (LOQ) were calculated respectively as 3 times the standard deviation of the blank (number of samples, n = 10) divided by the slope of the calibration curve or as 10 times the standard deviation of the blank (n = 10), divided by the slope of the calibration curve [Shrivastava 2011]. In the analysis of protein samples, the concentration of each element was calculated from the area under the peak of the K $\alpha$  line using the calibration curve. For each protein, at least 6 independent



samples at different concentrations were analysed and the stoichiometry calculated using the following equation:

$$\text{Stoichiometry} = \frac{[\text{element}]}{[\text{protein}] \times \text{MW}_{\text{element}} \times 1000}$$

Where [element] is the concentration of the relevant element calculated using the calibration curve (in ppm), [protein] is the protein concentration (in M), determined as reported in Table 8, and  $\text{MW}_{\text{Element}}$  is the molecular weight of the element (in Da). For each protein, the final stoichiometry was calculated by averaging the result from individual samples of differing concentrations.



**Figure 29:** Formation of the thin film from a sample drop

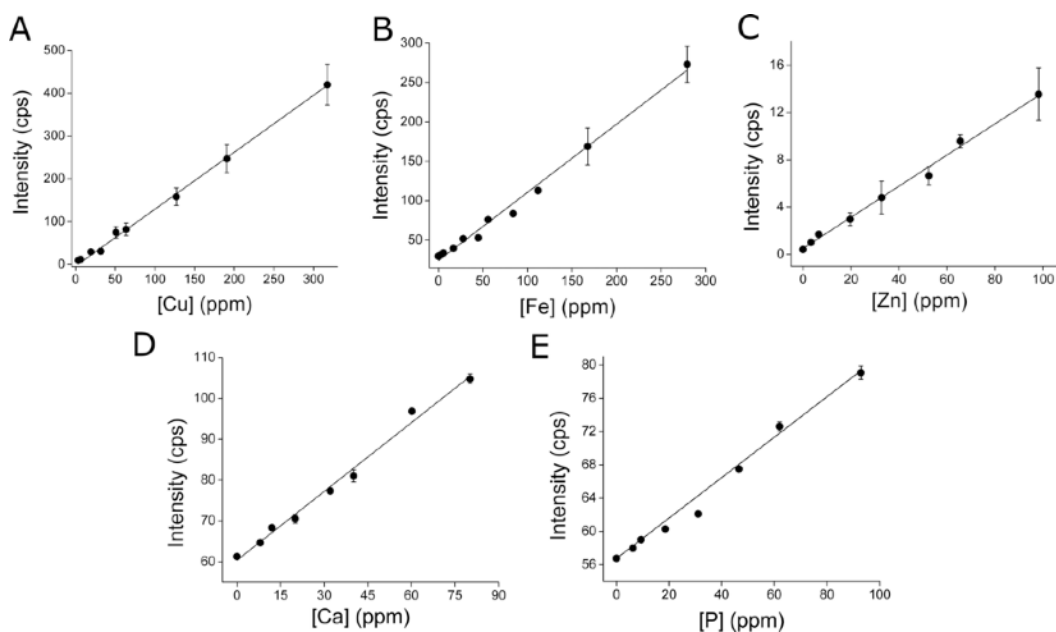
## **6.3 RESULTS AND DISCUSSION**

### **6.3.1 Experimental design and instrument calibration**

It was envisaged that by employing thin-film analysis it would be possible to minimize the amount of sample material required and avoid distortions of the spectra associated with thick samples that could lead to less accurate quantification [Markowicz 2011]. To further maximize the signal-to-noise ratio, the analyte solutions were made up in a TRIS-acetate buffer, which had previously been reported to give good results with TXRF [Szoboszlai 2009]. This buffer is particularly suitable as it contains no elements with an atomic number >11, thus

avoiding the introduction of any peaks that may overlap with the signal of interest and minimizing any intra-element absorption or enhancement effects.

Calibration curves were obtained, by producing thin-films from 8  $\mu\text{L}$  droplet standards for each element from salt solutions of known concentrations. EDXRF spectra were then recorded, and the area under the peak for the element of interest were plot against the original concentration of the solution (Figure 30). These plots gave good linear correlations as showed by the  $R^2$  values in Table 9, indicating that film thickness or small variations in the size of the film had a negligible effect on the quantification [Markowicz 2011]. Since very low concentrations of each element were analyzed, the spectra were recorded for 1600 s in order to maximize the limit of detection (LOD) while maintaining a reasonable duration for the proposed analysis. The LOD and limit of quantification (LOQ) values were found to be in the low ppm range (nmol-pmol quantities) for all the tested elements (Table 9). The light elements Ca and P emit photons of relatively low energy that are poorly transmitted through air [Rio-Echevarria 2019]; thus, to mitigate this effect, the sample chamber was purged with helium prior to recording the spectra. Even so, these elements showed the highest LOD and LOQ values. Moreover, P, Ca, and Fe generally showed the



**Figure 30:** Calibration plots of EDXRF peak area against element concentration for (A) Cu at 0–317 ppm, (B) Fe at 0–279 ppm, (C) Zn at 0–98 ppm, (D) Ca at 0–80 ppm, and (E) P at 0–93 ppm. Error bars represent the standard error of the mean (SEM).

highest levels of background signal compared to all the other tested elements (also evidenced by the high baseline in the calibration fit), likely due to their presence as common contaminants in Mylar films and other materials commonly used during EDXRF analysis such as polypropylene and polyester films. Nevertheless, these data indicate that the LOD and LOQ are sufficiently low to allow for the proposed protein analysis.

**Table 9:** Best Linear Fit Calibrations for Each Element

element	linear best fit equation <sup>a</sup>	R <sup>2</sup> <sup>b</sup>	LOD (ppm)	LOQ (ppm)
Cu	$I_{\text{Cu}} = 1.31[\text{Cu}] + 0.71$	0.997	1.1	3.5
Fe	$I_{\text{Fe}} = 0.86[\text{Fe}] + 24.16$	0.991	0.9	3.2
Zn	$I_{\text{Zn}} = 0.13[\text{Zn}] + 0.50$	0.993	0.8	2.7
Ca	$I_{\text{Ca}} = 0.56[\text{Ca}] + 60.39$	0.989	4.4	14.6
P	$I_{\text{P}} = 0.25[\text{P}] + 56.21$	0.986	5.4	18.8

<sup>a</sup> $I_{\text{element}}$  represents the signal intensity in the EDXRF spectrum for the element (in cps); [element] is the concentration of the element (in ppm). <sup>b</sup>R<sup>2</sup> is the correlation coefficient for a linear regression.

### 6.3.2 Protein analysis

Elemental analysis was undertaken on ten proteins containing Cu, Fe, Zn, Ca, and the nonmetallic element P (Table 10). Thin films were prepared from protein solutions of known concentration, and their EDXRF spectra were recorded (Figure 31). Protein quantification represents a crucial step in the stoichiometry determination, thus it was accurately performed according to the specific protein characteristics (Table 8). By applying the calibration line equations in Table 9, we were able to determine the concentration of the metal in the original analyte solution. Subsequently, using the previously measured protein concentration (expressed as the concentration of the monomer), the element/protein stoichiometry could then be calculated according to the formula in paragraph 6.2.

All the tested proteins were analysed at different concentrations in the  $\mu\text{M}$  range and the obtained element/protein ratios were in good agreement with the values obtained through other techniques (CaM1 [Astegno 2016]; CML7 [Trande 2019]; CML19 [La Verde (bis) 2018]; CytC [Bushnell 1990]; AHb1 [Kumar 2016]; HRP [Gajhede 1997]; CotA [Duraio 2008]; MvBOx [Cracknell 2011]; ADH1 [Raj 2014]; UbiX [White 2015]). For robust stoichiometry determination, it was found that the

Full name	Abbreviation	Element	Organism
Calmodulin 1	CaM1	Ca	<i>Arabidopsis thaliana</i>
Calmodulin Like 7	CML7		<i>Arabidopsis thaliana</i>
Calmodulin Like 19	CML19		<i>Arabidopsis thaliana</i>
Horseradish peroxidase	HRP		<i>Armoracia rusticana</i>
Spore coat protein A	CotA	Cu	<i>Bacillus subtilis</i>
Bilirubin oxidase	MvBOx		<i>Myrothecium verrucaria</i>
Cytochrome C	CytC	Fe	<i>Equus caballus</i>
Hemoglobin 1	Ahb1		<i>Arabidopsis thaliana</i>
Horseradish peroxidase	HRP		<i>Armoracia rusticana</i>
Alcohol dehydrogenase 1	ADH1	Zn	<i>Saccharomyces cerevisiae</i>
Flavin prenyltransferase	UbiX	P	<i>Escherichia coli</i>

**Table 10:** List of protein analysed by EDXRF method

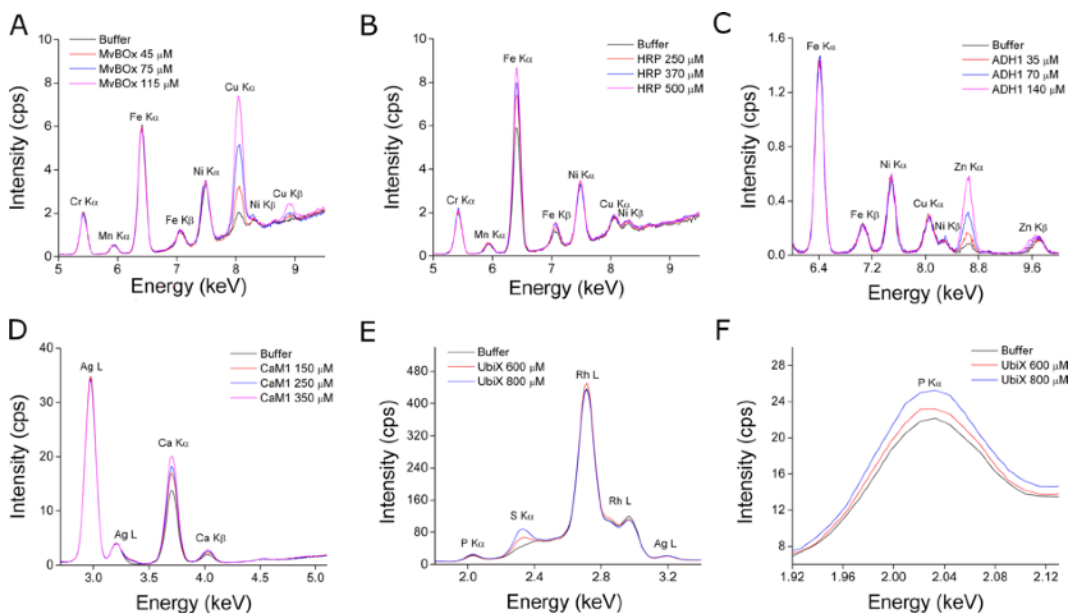
concentration of the metal should ideally lie between the LOQ and the highest calibrant concentration used in the construction of the calibration curve. Thus, heavier elements and proteins with multiple binding sites required lower amounts of protein for successful analysis, ranging from 35-140  $\mu\text{M}$  in the case of two Zn binding sites or four Cu binding sites up to 600-800  $\mu\text{M}$  in the case of a single P atom (Table 11). The stoichiometry values obtained by this method are not dependent on the quaternary structure of the protein since the calculations rely only on the protein monomer concentration. For example, AHb1 forms a homodimer [Kumar 2016 and Mukhi 2013] with one Fe atom per monomer, and this method gives a metal/monomer ratio of 1:1, while ADH1 forms a homotetramer [Raj 2014] with two Zn atoms per monomer, giving a metal/monomer ratio of 2:1. The analysis of individual elements was also not affected by the presence of other elements or post-translational modifications. For example, commercial batches of eukaryotic proteins CytC and HRP include varying degrees of acetylation and glycosylation, respectively [Bushnell 1990, Gajhede 1997 and Kalpage 2019], but analyses of both proteins gave the expected Fe stoichiometry of 1:1. Simultaneous quantification of multiple elements from a single protein sample is also possible, as in the case of

Abbreviation	Element	Range of protein concentration tested ( $\mu\text{M}$ )	Expected element/protein <sup>a</sup>	Obtained element/protein ( $\pm$ S.E.M.) <sup>b</sup>
CaM1	Ca	140-400	4	$4.01 \pm 0.20$
CML7			4	$3.85 \pm 0.08$
CML19			4	$3.60 \pm 0.12$
HRP		250-500	2	$1.85 \pm 0.12$
CotA	Cu	35-120	4	$3.97 \pm 0.10$
MvBOx			4	$4.13 \pm 0.12$
CytC	Fe	200-500	1	$1.03 \pm 0.05$
Ahb1			1	$0.98 \pm 0.04$
HRP			1	$0.99 \pm 0.05$
ADH1	Zn	35-140	2	$2.18 \pm 0.11$
UbiX	P	600-800	1	$1.30 \pm 0.20$

**Table 11:** Range of protein concentration tested, depending on the number of binding site and tested element, and measured element/protein stoichiometry. <sup>a</sup>Expected stoichiometry obtained by several techniques as reported in the main text. <sup>b</sup>E/P represents the element/protein stoichiometry. The mean values  $\pm$  SEM from at least six independent samples are presented.

HRP, which contains both a catalytic Fe and two structural Ca atoms. This method can also be used for the quantification of nonmetal elements, provided the photon energies are within the detectable range, as in the case of UbiX. The protein harbors a single riboflavin-5'-phosphate (flavin mononucleotide, FMN) cofactor, which itself contains a single P atom. Here, although the S.E.M. was relatively large due to the low sensitivity of the instrument toward light elements (Figure 31 E and F) the EDXRF-determined P/protein ratio was consistent with previous literature [White 2015]. On the other hand, elements such as Fe and Cu gave appreciable differences in spectral intensities even with differences in the concentration of the analyte solutions as low as 5  $\mu\text{M}$ . Nevertheless, this method can be used to confirm or exclude the presence of cofactor binding in proteins of interest. In the case of CytC, this protein has two possible phosphorylation sites [Kalpage 2019] but no signal for P was detectable above the background even at the highest protein concentration

that was analyzed, thus indicating that the CytC used in this study was in fact not phosphorylated.



**Figure 31:** Illustrative EDXRF spectra of protein samples at different concentrations. Overlaid spectra (A) of Cu-dependent MvBOx, (B) of Fe- dependent HRP, (C) of Zn-dependent ADH1, (D) of Ca-dependent CaM1, (E) of P-containing UbiX, and (F) showing a magnification of the P peak from the spectra shown in (E). The Rh signal observed in (E) originates from the Rh anode of the instrument.

## 6.4 CONCLUSIONS

In this Chapter, a simple and accurate thin-film EDXRF method has been described for the determination of metal stoichiometry within protein complexes. The method gives LOD values between 1 and 6 ppm and LOQ values between 3 and 20 ppm with a minimal consumption of sample (pmol) and does not require pre-digestion prior to analysis. This approach was demonstrated with a range of proteins containing Cu, Fe, Zn, and Ca as well as with one protein containing the nonmetal P. Though only a selected range of biologically relevant elements was investigated using existing EDXRF equipment, in principle, this type of spectroscopy allows the quantification of any element between Na and U, provided the appropriate instrument setup is available. It is generalizable across proteins of different molecular

weights, oligomerization states, and types and numbers of bound metals as well as post-translational modifications (e.g., glycosylation and acetylation). The quantitation of multiple elements within a single protein is also possible. By combining the fact that only minimal sample preparation is required and EDXRF instruments are relatively widely available, this method represents a competitive alternative to other methods for the quantitation of elements in proteins. It is intended that the method reported here can be employed for the analysis of novel synthetic proteins containing transition metals such as Rh or Ru<sup>26</sup> the validation of currently uncharacterized metalloproteins where the metal binding capability is only predicted through sequence-based bioinformatic analysis, or the quality assurance of recombinantly produced metalloproteins.

In particular for CBPs this method offers many advantages in the determination of Ca/protein stoichiometry compared to ITC and NMR. LOD and LOQ values were 4.4 ppm and 16.4 ppm respectively, corresponding to a molar concentration of 11  $\mu$ M and 41  $\mu$ M, implicating that buffer decalcification prior the analysis was not required as residual Ca was beyond detection [Dkhar 2014]. Purified protein preparation only required few washes steps as reported in paragraph 6.2 thus overcoming the delicate decalcification process needed for ITC analysis or the <sup>13</sup>C/<sup>15</sup>N labelling in NMR experiment. Finally, data analysis only required the design of a linear fit equation without the necessity of specific softwares, simplifying results calculation (e.g. Origin).

Thus, in the study of CBP proteins, our method represents a quick and easy procedure to support stoichiometry determination data, requiring lower amount of samples and experimental procedure compared to the currently used methodologies.

# Chapter 7

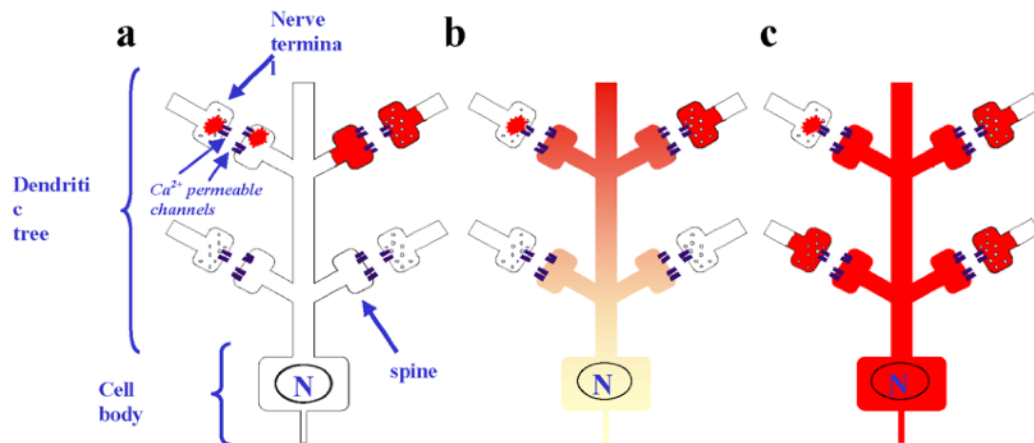
## *Neuronal calcium binding proteins (NCCABs)*



## 7.1 INTRODUCTION

### 7.1.1 Neuronal $\text{Ca}^{2+}$ signaling

$\text{Ca}^{2+}$  is essential for the regulation of physiological processes in neuronal and glial cells. In neurons, variation in the  $\text{Ca}^{2+}$  concentrations ranges from highly localized and transient to longer lasting and global inducing alteration in cell function that can persist from fractions of second to days or years (for example, in learning and memory), thus a wide variety of different  $\text{Ca}^{2+}$  signals can be generated in mature neurons, compared to other cell types (Figure 32).



**Figure 32 [Burgoyne 2007]:** Diversity of  $\text{Ca}^{2+}$  signals in neurons. Representation of three types of  $\text{Ca}^{2+}$  signal are shown schematically. a) Local, that are generated and remain in the pre-synaptic nerve terminal or the post-synaptic dendritic spine. b) Propagated into part of the dendritic tree close to the site of synaptic inputs but that dissipates before reaching the cell body. c) A global  $\text{Ca}^{2+}$  signal that would occur following more extensive stimulation.

At the molecular level,  $\text{Ca}^{2+}$  affects neurotransmitter release, ion channels and gene expression, resulting in controlling neuronal growth, development, survival and death. Aside from calmodulin, neuronal cells display their own set of  $\text{Ca}^{2+}$  sensors (NCS); NCS show poor sequence identity with CaM (generally lower than 20%) and even if they usually share the presence of four EF-hand domains in the primary sequence only two or three of them are actually functional  $\text{Ca}^{2+}$  binding sites [Burgoyne 2007].

Nowadays, NCS have been classified in 5 classes (A-E), depending on their amino acid sequence: Class A is composed by Frequentin proteins (NCS-1); class B by Visin-like proteins (VILIPS), class C by Recoverins, class D by GC-activating proteins (GCAPs) and finally class E include Kv-channel interacting proteins (KChIPs) [Burgoyne 2001].

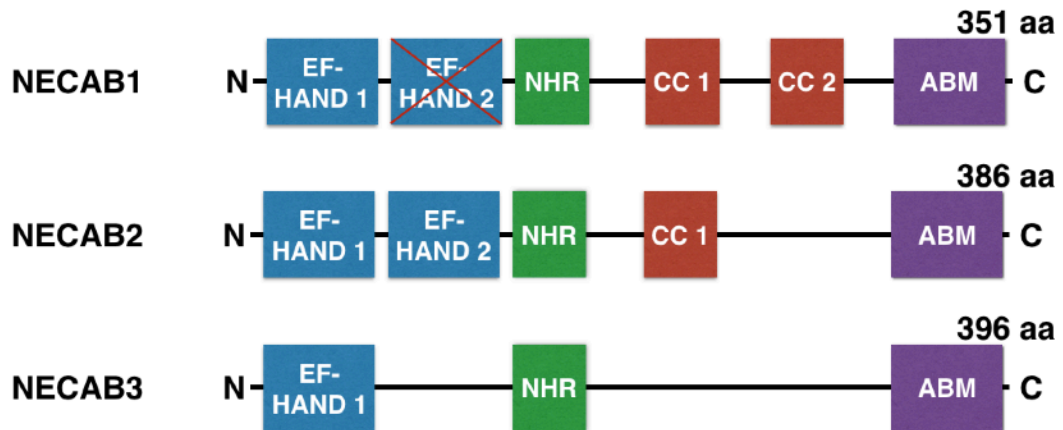
The majority of NCS and neuronal calcium binding proteins have been already deeply characterised, however, many neuronal EF-hand containing proteins, which are not believed to act as Ca<sup>2+</sup> sensors, are known only in terms of their primary sequence and no biochemical information is available about their mechanism of action or Ca<sup>2+</sup> binding properties. In particular, we found particularly interesting a family of putative NCS identified by Sugita et al [Sugita 2002] which has been termed NECABs (NEuronal CALcium Binding proteins).

### 7.1.2 NECABs proteins

NECABs family is composed by three members (NECAB1, NECAB2 and NECAB3) which share three main common structural features (Figure 33): the presence of one or two EF-hand domains at the N-terminus, a central unique and highly conserved region (named NECAB homology region (NHR)) of unknown function and a C-terminal antibiotic biosynthesis monooxygenase domain (ABM) [Sugita 2002]. The ABM domain is a bacterial domain, never found in mammals, that catalyses the oxidations of different aromatic intermediates and is present in proteins involved in a diverse range of biological processes, including metabolism, transcription, translation and biosynthesis of secondary metabolites. From the structural point of view, the ABM domain crystallizes as a homodimer in which the core of the barrel is mainly hydrophobic [Kendrew, 1997; Sciara, 2003; Adams, 2005; Wu, 2005; Lee, 2008; Henrick, 2012; Chim, 2010; Lemieux, 2005 and Wada, 2004]. In addition to these feature, bioinformatic tools (i.e. COILS [Lupas 1991]) revealed the presence of two and one predicted coiled coiled regions (CC) in NECAB1 and NECAB2 respectively.

Notably, NECAB1 and NECAB2 mRNAs have been detected only in brain, whereas NECAB3 have been found at similar levels in brain, heart, skeletal muscle and, at lower level in pancreas [Sugita 2002]. Even if the role of this subclass of proteins is still unclear and poorly characterized, few studies have highlighted the possible

involvement of NECAB proteins in neurodegenerative diseases and have identified some of their putative interactors.



**Figure 33:** Predicted domain composition of NECABs proteins. EF-hand 2 in NECAB1 is predicted to be unable to bind  $\text{Ca}^{2+}$  due to the amino acid composition of the binding loop.

For instance, NECAB3 (also called XB51) has been found to interact with the amino-terminal domain of the adaptor protein X11L which is a major player in the regulation of amyloid  $\beta$ -protein precursor by suppressing its intracellular metabolism and amyloid  $\beta$  (Ab) generation, a hallmark of Alzheimer's disease [Lee 2000].

NECAB2 was found to interact specifically with the Adenosine heptaspanning membrane receptor A2A (A2AR) and the membrane receptor with high relevance in brain function, namely, the metabotropic glutamate type 5 receptor (mGlu5b). Importantly, high concentration of  $\text{Ca}^{2+}$  markedly inhibited the binding of NECAB2 to A2AR in vitro, suggesting that NECAB2 target partner, and consequently the modulation of its function, might be determined by physiological concentration of the intracellular ion. Due to the ability of NECAB2 to interact both with mGlu5b receptor and A2AR, it has also been speculated that it might play a role in the mGlu5b/A2AR oligomer-mediated physiological functions [Canela 2007 and 2009].

Finally, preliminary studies showed the involvement of NECAB1, the less characterised members of the family, in neuronal differentiation and plasticity due to its over-expression in the noradrenergic to cholinergic sympathetic neurons induced development [Apostolova 2007] or in human neural progenitor cells treated with the myocyte enhancer factor 2 (MEF2), which is a transcription factor family highly

expressed in the brain and constitutes a key determinant of neuronal survival, differentiation, and synaptic plasticity [Chan 2015].

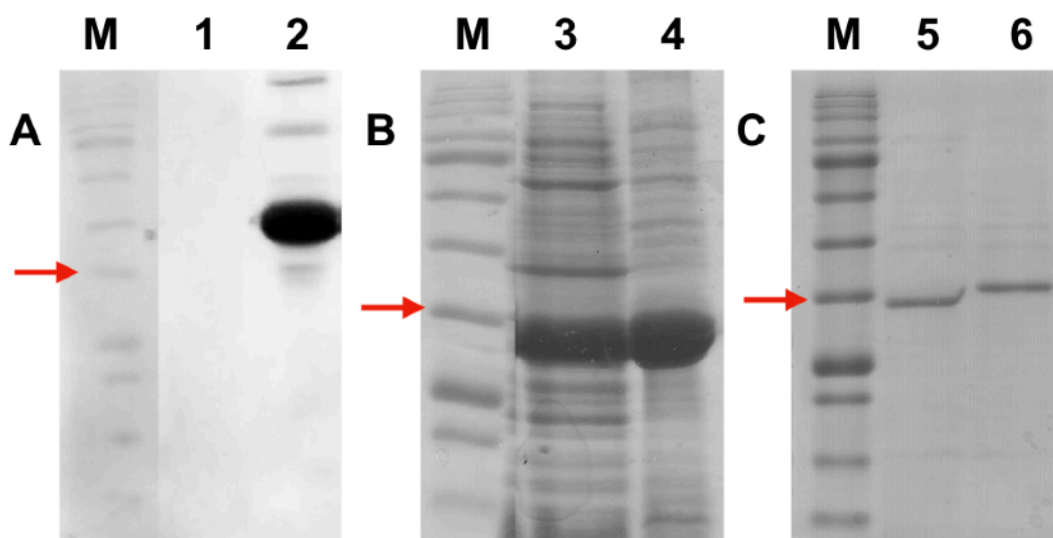
In this Chapter, I report initial attempts in the recombinant production of NECAB proteins and a preliminary analysis of their biochemical features, as NECAB proteins might represent a still uncharacterised and highly relevant new class of CBPs and NCSs.

## **7.2 RESULTS AND DISCUSSION**

### **7.2.1 Production of recombinant NECABs**

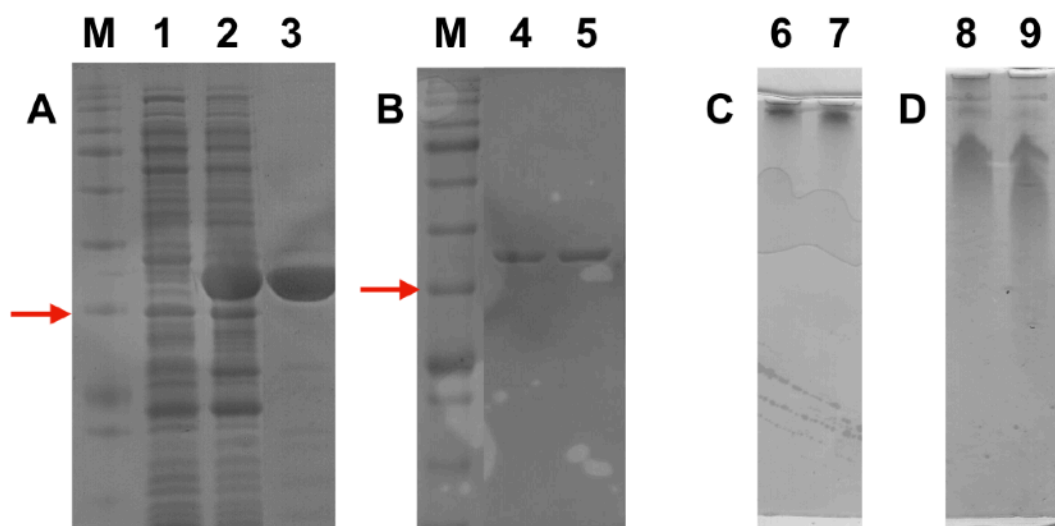
Initial attempts to produce the recombinant NECABs into *E.coli* have been performed by transforming different *E. coli* strains (i.e. BL21, BL21 Codon Plus, BL21 p-Lys-S, C41, C43, C43 p-Lys-S, Rosetta and Origami) with expression vectors coding for NECAB proteins as reported in paragraph 3.1. Unfortunately, NECAB3 was not expressed in any of the tested cell strains, as it was confirmed by detection of the 6His-Tag by Western Blot (data not shown), suggesting a quick protein degradation or inefficiency in protein translation. NECAB2, on the other hand, was expressed in BL21, BL21 Codon Plus and Origami cells; however, no soluble protein was found after cell lysis (Figure 34 A). Thus, we exploited bioinformatic tools (i.e. GeneSilico MetaDisorder Service) to evaluate the presence of disordered regions in the protein primary sequence. Due to the presence of an unstructured segment in the initial N-terminal region, we expressed a truncated version (without the first 56 residues) of NECAB2 (NECAB2DN56), that was found to be present in the soluble fraction of the lysed cells (Figure 34 B). However, purification by Nickel affinity chromatography, revealed that NECAB2DN56 was not able to properly bind the column as it was present also in the flow thorough fraction (data not shown), thus resulting in a low final yield (less than 2 mg/L of culture), unsuitable for a proper biochemical analysis.

Interestingly, despite its high molecular weight, compared to canonical Ca<sup>2+</sup> sensors, and the presence of other domains rather than the EF-hands, NECAB2DN56 showed increased electrophoretic mobility in SDS-PAGE analysis when saturated with Ca<sup>2+</sup> ions, a typical property of Ca<sup>2+</sup> sensors (Figure 34 C).



**Figure 34:** A) Immunoblotting against the His-tag on full length NECAB2 expressed in BL21 Codon Plus cells. Lane 1 represents the soluble fraction, lane 2 the insoluble pellet and lane M the molecular weight markers. The red line correspond to 35 kDa. B) SDS-PAGE of the soluble (lane 3) and insoluble (lane 4) fractions of NECAB2DN56 expression in BL21 Codon Plus cells. C) SDS-PAGE mobility shift analysis of NECAB2DN56 in the presence of 5 mM  $\text{CaCl}_2$  (lane 5) or EGTA (lane 6). Molecular weights of NECAB2 and NECAB2DN56 are 43.2 and 38.1 kDa respectively.

Production of recombinant NECAB1 was successfully achieved in BL21 (DE3) Codon Plus cells after induction with 0.4 mM IPTG at 24°C for 20 hours. The protein was purified by Ni-affinity chromatography, leading to a final yield of 40 mg/L (Chapters 3.1 and 3.2) (Figure 35 A). Initially, as for NECAB2N56, NECAB1 was tested for its ability to respond to  $\text{Ca}^{2+}$  by SDS and NATIVE-PAGE (Figure 35 B, C, D) in order to evaluate the typical  $\text{Ca}^{2+}$  sensor property of improved electrophoretic mobility in the presence of the cation; however, as shown in Figure 35 (lanes 4 and 5) the addition of  $\text{CaCl}_2$  to the protein did not result in an alteration of the electrophoretic mobility. Interesting, we observed that when loaded on a 10% native gel NECAB1 was not capable of entering the acrylamide gel (Figure 35 C) and required a very low percentage of Bis-tris-acrylamide (8%, Figure 35 D) to penetrate the gel therefore suggesting the possible formation of oligomers or aggregates. Thus, our initial analysis was aimed at evaluating if recombinant NECAB1 is correctly folded in vitro and if protein production is associated with protein aggregation or oligomerization.

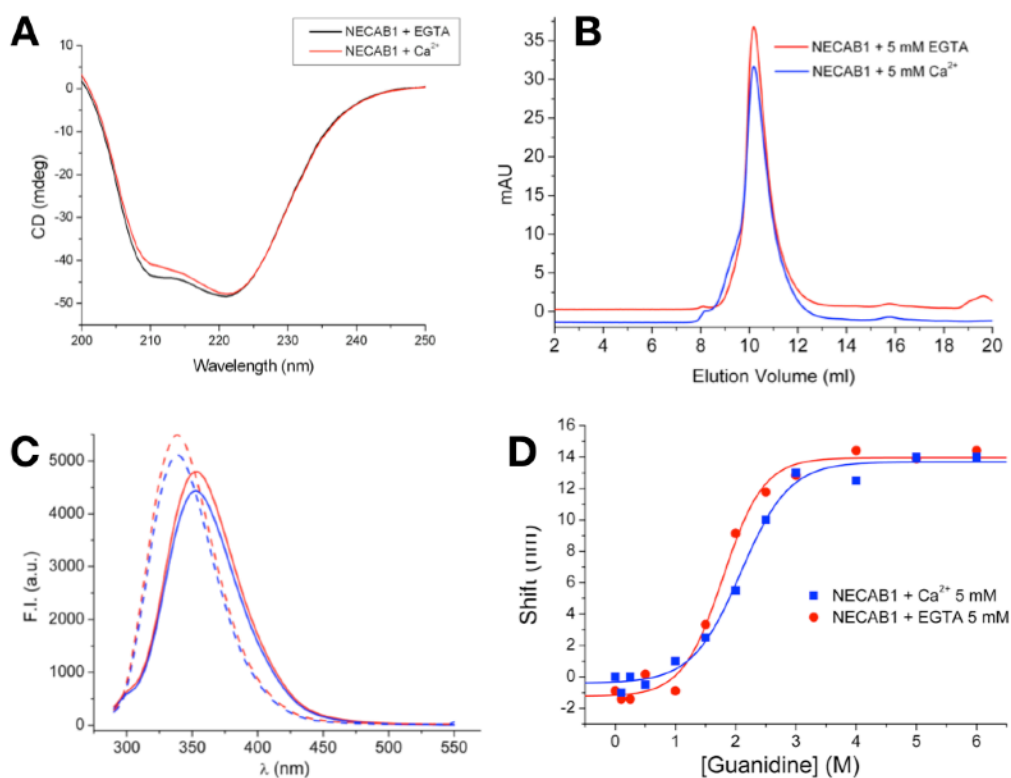


**Figure 35:** A) SDS-PAGE analysis of NECAB1 at various steps of purification. 1) Before induction, 2) after induction and 3) after affinity column. Lane M indicates the marker and the band corresponding to 35 kDa is indicated with a red arrow. MW of NECAB1 is 41.1 kDa. B) SDS-PAGE mobility shift analysis of NECAB1 in the presence (lane 4) or in the absence (lane 5) of Ca<sup>2+</sup> ions. C-D) NATIVE-PAGE mobility shift analysis of NECAB1 in a 10% and 8% gel respectively, in the presence (lanes 6 and 8) or in the absence (lane 7 and 9) of Ca<sup>2+</sup> ions.

Initially, the presence of high molecular weight species was verified by Size Exclusion Chromatography. NECAB1 possesses a Stock Radius of  $74.5 \pm 0.1 \text{ \AA}$  (Figure 36 B), which corresponds to a globular protein with molecular weight around 500 kDa. This high molecular weight species again raises questions regarding the presence of high order oligomer or aggregates in NECAB1 preparation.

However, by far-UV CD spectroscopy, we observed that NECAB1 shows the typical fingerprint of a well folded protein with a high  $\alpha$ -helical content (Figure 36 A) characterised by the two minima at 222 nm and 208 nm. The  $\theta_{222}/\theta_{208}$  ratio was higher than 1, confirming the presence of coiled coil regions as predicted by bioinformatic tools (See paragraph 7.1.2 and [Zhang 1995 and Gagne 1994]). Interestingly, the presence of Ca<sup>2+</sup> led to a reduction only in the 208 nm signal and thus a further increase in the coiled coil content. The ability of Ca<sup>2+</sup> to induce structural rearrangements in the protein conformation suggested that at least one of the two predicted EF-hand in NECAB1 primary sequence could bind to the cation. Thus, we exploited the EDXRF based method described in Chapter 6 to evaluate the Ca<sup>2+</sup>/protein stoichiometry in NECAB1 samples. Since only one EF-hand was

predicted to be functional by PROSITE rule annotations, the protein was tested in a range of concentrations from 600 to 800  $\mu\text{M}$ , as the optimal protein concentration for EDXRF measurement is dependant on both the tested element and the number of binding sites, as specified in the previous Chapter. Experimental stoichiometry was found to be  $1.31 \pm 0.04$  (Average  $\pm$  S.E.M. of six independent samples) and thus confirmed the presence of one functional  $\text{Ca}^{2+}$  binding site which is likely the first N-terminal domain as suggested by prediction tools.



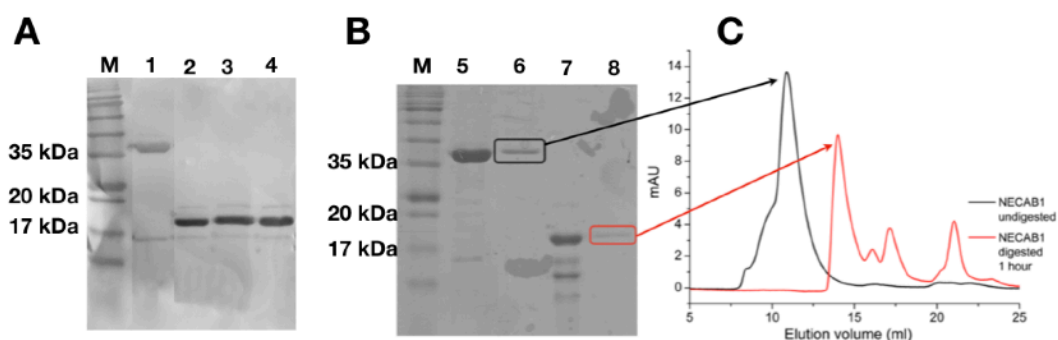
**Figure 36: Biochemical characterisation of NECAB1.** **A)** Far-UV CD spectroscopy of NECAB1 in the presence of 5 mM  $\text{CaCl}_2$  (red line) or 5 mM EGTA (black line). **B)** SEC analysis of NECAB1 on a Superdex 200 column in the presence (blue line) or in the absence (red line) of  $\text{Ca}^{2+}$ . **C)** Tryptophan intrinsic emission fluorescence spectra of NECAB1 in the presence (blue lines) or absence (red lines) of  $\text{Ca}^{2+}$  and in the absence (dashed lines) or in the presence of 6 M Guanidine chloride (solid lines). **D)** Fitting of the titration curve of guanidine-induced unfolding and assuming a 2-state transition model.

To further support the presence of high order oligomers we evaluated the guanidine induced denaturation of NECAB1. In particular, since NECAB1 possesses 6 Trp residues, we evaluated the guanidine-induced unfolding of NECAB1 by following the intrinsic fluorescence of Trp residues at increasing denaturant concentrations.

The maximum emission of NECAB1 shifted from 339 to 353 nm upon the addition of 6 M Guanidine chloride and thus suggests that the Trp residues become exposed to the aqueous solvent (Figure 36 C). Protein denaturation follows a sigmoidal like behaviour in which a single denaturation event can be deduced (Figure 36 D). The presence of  $\text{Ca}^{2+}$  ions, induced a slight increase in the protein stability as the guanidine induced transition midpoint from the folded to the unfolded state of NECAB1 ( $[\text{GdnHCl}]_{1/2}$ ) shifted from  $2.0 \pm 0.1$  M in the presence of 5 mM  $\text{CaCl}_2$  to  $1.8 \pm 0.1$  M in the presence of 5 mM EGTA.

### 7.2.2 Does the C-terminal region of NECAB1 induce the formation of high order oligomers?

In order to understand if a particular domain was responsible for the protein oligomerization, we incubated NECAB1 with trypsin for several time intervals (5 min, 30 min and 120 min) and analysed the obtained fragments by immunoblotting against the His-tag located at the C-terminal of the protein.



**Figure 37:** **A)** Immunoblotting against the C-terminal his-tag of NECAB1. Lane 1 is the undigested protein, lanes 2, 3 and 4 resulted after digestion with trypsin (1:500 w/w) for 5, 30 and 120 min respectively. **B)** SDS-PAGE analysis of NECAB1 before (lane 5) and after (lane 6) SEC analysis and of the 1 hour digested fragment before (lane 7) and after (lane 8) SEC analysis. **C)** Elution profiles of undigested NECAB1 (black line) and 1 hour digested fragments (red line) analysed of a Superdex 200 column.

Figure 37 A shows that even after two hours of incubation with trypsin, it is possible to highlight the presence of an  $\sim 18$  KDa trypsin resistant fragment which was revealed by Western blotting and thus represent the C-terminal of NECAB1, likely including the ABM domain and the coiled coil regions. Interesting, when analysed by size exclusion chromatography, the 1 hour digested NECAB1 fragment was eluted as



a globular species of  $\sim 87$  kDa ( $R_s \sim 36$  Å) (Figure 37 B and C) suggesting that the C-terminal region forms an oligomeric compact core resistant to trypsin digestion. On the other hand the N-terminal region of NECAB1 is quickly digested by the protease, indicating a more unstructured conformation of the initial EF-hands region which likely increase the hydrodynamic radius of the undigested protein over its actual oligomeric state. Thus it might be speculated that the C-terminal region of NECAB1 is the main responsible for protein oligomerization since digested protein still runs as a oligomeric species in SEC analysis. Another interpretation is that the N-terminal region might be capable of recruiting further protein molecules, increasing the effective oligomeric states and exposing a central region to trypsin digestion which also remove the extra monomers interacting at the N-terminal region. Further studies might be necessary to deeply understand the process.

### **7.3 CONCLUSIONS**

In this Chapter we observed preliminary and fascinating aspects of the yet uncharacterised family of NECABs proteins. Though limited, the initial information in this Chapter represents a starting point in the study of the NECAB proteins. Attempts to produce recombinant NECAB2 and NECAB3 in *E. coli* systems, revealed difficulties associated with the correct protein folding and production. In particular, NECAB3 was not found in both the soluble and insoluble fractions of lysed overproducing cells, indicating likely the need of higher organism to obtain a correct protein production. NECAB2, on the other hand, was produced as insoluble protein when expressed as full-length protein, indicating that the disordered N-Terminal region lead to protein aggregation into inclusion bodies. The removal of the first 56 residues improved the solubility of the protein and allowed us to verify the capability of NECAB2 to respond to  $Ca^{2+}$  ions. However, due to the high amount of pure and stable protein necessary for a complete biochemical characterisation, the optimisation of the purification procedure is required and will be the next goal of the laboratory. Among the three isoforms, NECAB2 and NECAB3 are indeed the most studied and likely involved in pathological mechanisms of neurodegeneration. Less is known about NECAB1 but the shared domain composition suggests its role in similar mechanisms. This isoform was successfully produced and initial characterisation confirmed that the binding of one

Ca<sup>2+</sup> ion to the protein induces null or limited variations in the protein conformation, differently from the typical Ca<sup>2+</sup> sensor proteins. The consequences of Ca<sup>2+</sup> binding on NECAB1 remain unclear rising hypothesis about the involvement of the EF-hand domains in different biochemical processes, such as the activation of the catalytic activity of the ABM domain, whose function in mammal cells is not understood yet. Furthermore, NECAB1 showed an evident propensity to form high oligomeric species, it remains unclear whether NECAB1 oligomerization in vitro represents the functional conformation of NECAB1, though it is noteworthy that this propensity might be involved in disease development propensity as observed in many other neuronal proteins [Koo, 1999 and Trojanowski, 1998], thus a deeper insight in the oligomerization process of NECBA1 (e.g. dependence on protein concentration, buffer pH and composition) will be performed in the laboratory.

# *Chapter 8*

*Final discussion*

The importance of  $\text{Ca}^{2+}$  in living cells has been known from a very long time. As summarised by Carafoli [Carafoli 2002], the beginning of the  $\text{Ca}^{2+}$  signaling saga can be placed in London in 1883 when Sidney Ringer [Ringer 1883] was studying the contraction of isolated rat hearts. Nowadays its work would have probably considered inaccurate or sloppy, since he initially resuspended the samples hearts in a saline medium prepared with tap water to observe muscle contraction. When the hard water of London was replaced by distilled water, the beating of the heart became weaker and finally stopped after half an hour, and Ringer understood the need of  $\text{Ca}^{2+}$  salts to maintain contraction. Thus, Ringer had serendipitously realised that  $\text{Ca}^{2+}$ , which before that was considered exclusively a structural element of bones and teeth, had a completely new function as one of the most conserved second messenger currently known. Since that day in London, extensive efforts have been carried out by the scientific community to disentangle the complex mechanisms involved in the regulation and the decoding of the  $\text{Ca}^{2+}$  signals showing the key role of CBPs in such network.

This PhD thesis gives important insight into the study of CBPs, and thus helps in the understanding of  $\text{Ca}^{2+}$  pathways in cells. Plants in particular are known to be capable of exploiting  $\text{Ca}^{2+}$  signalling in a more sophisticated way than animals, and thus we provided a deep characterisation of two *A. thaliana*  $\text{Ca}^{2+}$  sensors in an effort to understand the differences in their sensing properties and the regulation of downstream targets. Thus, in Chapters 4 and 5 we analysed the metal and target binding properties of CML7 and CML19 by applying biochemical and structural methods. Both the proteins were confirmed to act as  $\text{Ca}^{2+}$  sensors in terms of ion induced rearrangements, and showed differences related to the  $\text{Ca}^{2+}$  and target binding properties strongly indicating the involvement in different physiological pathways. Although very efficient and informative, some of the methods used in Chapters 4 and 5 have important limitations in the determination of metal/protein stoichiometry, especially in the study of CBPs, which are generally characterised by the presence of multiple binding sites, clustered affinities and cooperativity events; therefore, in Chapter 6 we optimised a new method that is very relevant for the characterisation of metalloproteins, giving an easy, quick and cheap alternative in the evaluation of metal/protein stoichiometry. Finally, in Chapter 7 our knowledge of CBPs were devoted to the understanding of a new and fascinating class of human CBPs. This peculiar family represents an intriguing connection bridge of structural

domains between several organisms and thus might be involved in still unknown physiological mechanisms. In our studies we just begun to understand  $\text{Ca}^{2+}$  binding properties and evaluated structural aspects that might be correlated to aggregation propensity of NECABs, a well-known phenomenon in the development neurodegeneration.

Overall, this Doctorate work has united many knowledges, form the most brilliant scientists I have ever met, to deeply explore the molecular basis of  $\text{Ca}^{2+}$  response and add just one more drop of understanding in the ocean of the  $\text{Ca}^{2+}$  signalling mystery.

# References

- Acu I.D., Liu T., Suino-Powell K., Mooney S.M., D'Assoro A.B., Rowland N., Muotri A.R., Correa R.G., Niu Y., Kumar R., Salisbury J.L., *Cancer Res*; 70(8) (2010)
- Adams MA, Jia Z. *J. Biol. Chem.* 280, 8358-63, (2005).
- Apostolova G Dorn R., Ka S., Hallböök F, Lundeberg J, Liser K., Hakim V., Brodski C., Michaelidis T.M., Dechanta G., *Mol. Cell. Neurosci.* 35, 397–408 (2007)
- Araki M., Masutani C., Takemura M., et al., *J Biol Chem*; 276:18665–72. (2001)
- Arnold K., Bordoli L., Kopp J., Schwede T., *Bioinformatics* 22, 195–201, (2006)
- Astegno A., Giorgetti, A. Allegrini A., Cellini B., Dominici P., *BioMed Res. Int.* 2013 701536. (2013)
- Astegno A., La Verde V., Marino V., Dell'Orco D., Dominici P., *Biochim. Biophys. Acta* 1864 (3), 297–307. (2016)
- Astegno A., Bonza M.C., Vallone R., La Verde V., D'Onofrio M., Luoni L., Molesini B., Dominici P., *J. Biol. Chem.* 292, 15049–15061, (2017)
- Azimzadeh J., Nacry P., Christodoulidou A., Drevensek S., Camilleri C., Amiour N., Parcy F., Pastuglia M., Bouchez D., *The Plant Cell*, Vol. 20: 2146–2159, (2008)
- Azimzadeh J., Hergert P., Delouvé A., Euteneuer U., Formstecher E., Khodjakov A., Bornens M., *J Cell Biol.* Apr 6;185(1):101-14. (2009)
- Batty D., Rapić-Otrin V., Levine A.S., Wood R.D., *J Mol Biol* ;300:275–90. (2000)
- Baum P., Yip C., Goetsch L., Byers B., *Mol. Cell. Biol.* 8, 5386-5397 (1988)
- Bender K.W., Snedden W.A., *Plant Physiology*, Vol. 163, pp. 486–495, (2013)
- Bender K.W., Dobney S., Ogunrinde A., Chiasson D., Mullen R.T., Teresinski H.J., Singh P., Munro K., Smith S.P., Snedden W.A., *Biochem. J.* 457, 127–136, (2014)
- Benkert P., Biasini P., Schwede T., *Bioinformatics* 27, 343–350, (2011)
- Berridge M.J., Lipp P., Bootman M.D., *Nature Reviews, Molecular Cell Biology*, Vol. 1, 11-21, (2000)
- Bonza M.C., Loro G., Behera S., Wong A., Kudla J., Costa A., *Plant Physiol.* 163 (3) 1230–1241 (2013)

- Boonburapong, B.; Buaboocha, T.; BMC Plant Biol., 7, 4. (2007)
- Bradford M.M., Anal. Biochem. 72, 248–254, (1976)
- Burgess W.H., Jemiolo D.K., Kretsinger R.H., Biochimica et Biophysica Acta, 623, 257-270, (1980)
- Burgoyne R.D., Nat Rev Neurosci. March ; 8(3): 182–193. (2007)
- Bushnell, G. W.; Louie, G. V.; Brayer, G. D. J. Mol. Biol., 214, 585–595. (1990)
- Canela L., Luján R., Lluís C., Burgueño J., Mallol J., Canela E. I., Franco R., Ciruela F., Mol. Cell. Neurosci. 36 1–12 (2007)
- Canela L, Fernández-Dueñas V, Albergaria C, Watanabe M, Lluís C, Mallol J, Canela EI, Franco R, Luján R, Ciruela F, J Neurochem. 111(2):555-67 (2009)
- Carafoli E., PNAS vol. 99 no. 3 1115–1122 (2002)
- Chan S.F., Huang X., McKercher S.R., Zaidi R., Okamoto S.-I., Nakanishi N., Lipton S.A., Genomics Data 3, 24–27, (2015)
- Charbonnier J.B., Renaud E., Miron S., Le Du M.H., Blouquit Y., Duchambon P., Christova P., Shosheva A., Rose T., Angulo J.F., Craescu C.T., J. Mol. Biol. 373 (4) 1032–1046, (2007)
- Chiasson D., Ekengren S.K., Martin G.B., Dobney S.L., Snedden W.A., Plant Molecular Biology, Volume 58, Issue 6, pp 887–897 (2005)
- Chim N, Iniguez A, Nguyen TQ, Goulding CW. J. Mol. Biol. 395, 595-608, (2010).
- Chin D, Means A.R., trends in Cell Biology, Vol. 10, 322-328, (2000)
- Clapham D.E., Cell, Vol. 80, 259-268 (1995)
- Comte M., Maulet Y., Cox J.A., The Biochemical journal 209, 269–272, (1983)
- Cracknell, J. A.; McNamara, T. P.; Lowe, E. D.; Blanford, C. F. Dalton Trans. 40, 6668–6675. (2011)
- Dantas T.J., Wang Y., Lalor P., Dockery P., Morrison C.G., J. Cell Biol. Vol. 193 No. 2 307–318 (2011)
- de Castro E., Sigrist C.J., Gattiker A., Bulliard V., Langendijk-Genevaux P.S., Gasteiger E., Bairoch A., Hulo N., Nucleic Acids Res. 34 W362–5, (2006)
- Del Vecchio A.J., Harper J.D.I., Vaughn K.C., Baron A.T., Salisbury J.L. Overall R.L., Protoplasma 196: 224 (1997)
- Delk N.A., Johnson K.A., Chowdhury N.I., Braam J., Plant Physiol. 139, 240–253. (2005)
- Dkhar E.N., Dkhar P.S., Anal J.M.H., Journal of Chemistry 975810 (2014)

- Dobney S., Chiasson D., Lam P., Smith S.P., Snedden W.A., *J. Biol. Chem.* 284 31647–31657. (2009)
- Dong Z., Saikumar P., Weinberg J.M., Venkatachalam M.A., *Annu. Rev. Pathol. Mech. Dis.* 1:405–34 (2006)
- Durao P., Chen Z., Fernandes A. T., Hildebrandt P., Murgida D. H., Todorovic S., Pereira M. M., Melo E. P., Martins L.O., *J. Biol. Inorg. Chem.*, 13, 183-193, (2008)
- Edel K.H., Marchadier E., Brownlee C., Kudla J., Hetherington A.M., *Current Biology* 27, R667–R679 (2017)
- Fitch M.E., Cross I.V., Ford J.M., *Carcinogenesis*; 24:843–50. (2003)
- Friedberg F., *Mol Biol Rep*, 33(4) 243-52. (2006)
- Fruncillo S., **Trande M.**, Blanford C.F., Astegno A., Wong L.S., *Anal Chem.* 17;91(18):11502-11506 (2019)
- Gagne S.M., Tsuda S., Li M.X., Chandra M., Smillie L.B., L.B. Sykes L.B., *Protein Sci.* 3, 1961–1974, (1994)
- Gajhede, M.; Schuller, D. J.; Henriksen, A.; Smith, A. T.; Poulos, T. L. *Nat. Struct. Biol.*, 4, 1032–1038. (1997)
- Garcia J. S., de Magalhaes C. S., Arruda M. A. Z., *Talanta*, 69, 1–15 (2006)
- Garman, E. F.; Grime, G. W. *Prog. Biophys. Mol. Biol.*, 89, 173–205. (2005)
- Gifford J.L., Walsh M.P., Vogel H.J., *Biochem J.*, 405, 199-221, (2007)
- Gifford J.L., Jamshidiha M., Mo J., Ishida H., Vogel H.J., *Plant Cell* 25, 4512–4524, (2013)
- Grecu D., Blouquit Y., Assairi L., *FEBS Open Bio* 4 33–42, (2014)
- Gut H., Dominici P., Pilati S., Astegno A., Petoukhov M.V., Svergun D.I., Grütter M.G., Capitani G., *J. Mol. Biol.* 392, 334–351, (2009)
- Habermann E., Crowell K., Janicki P., *Arch Toxicol* 54: 61-70, (1983)
- Henrick K., Hirshberg M. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 68, 744-50, (2012)
- Hoeflich K.P., Ikura M., *Cell* 108 (6), 739–742. (2002)
- Hu H., Chazin W.J., *J. Mol. Biol.* 330 (3) 473–484. (2003)
- Hu H., Sheehan J.H., Chazin W.J., *The journal of biological chemistry*, Vol. 279, No. 49, 50895–50903, (2004)
- Huang B., Mengersen A., Lee V.D., *Journal of cell biology*, 107 (1): 133 (1988)
- Ishida H., Vogel H.J., *Protein Pept. Lett.* 13 (5) 455–465. (2006)



- Islam M.D., Book: Calcium signalling, Springer, Dordrecht Heidelberg New York London (2012)
- Ikura M., Clore G.M., Gronenborn A.M., Zhu G., Klee C.B., Bax A., Science 256, 632–638, (1992)
- Kalpage, H. A.; Bazylanska, V.; Recanati, M. A.; Fite, A.; Liu, J.; Wan, J. M.; Mantena, N.; Malek, M. H.; Podgorski, I.; Heath, E. I.; Vaishnav, A.; Edwards, B. F.; Grossman, L. I.; Sanderson, T. H.; Lee, I.; Huttemann, M. FASEB J, 33, 1540–1553. (2019)
- Kelly S.M., Jess T.J., Price N.C., Biochimica et Biophysica Acta 1751, 119–139, (2005)
- Kendrew S.G., Hopwood D.A., Marsh E.N.G., Journal of Bacteriology, Vol. 179, No. 13, p. 4305–4310 (1997)
- Kilmartin, J. V., J. Cell Biol. 162, 1211-1221. (2003)
- Koo E.D., Lansbury P.T., Kelly J.W., Proc. Natl. Acad. Sci. USA Vol. 96, pp. 9989–9990, (1999)
- Kuhlman B., Yang H.Y., Boice J.A., Fairman R., Raleigh D.P., J. Mol. Biol. 270, 640–647, (1997)
- Kumar N., Astegno A., Chen J., Giorgetti A., Dominici P., Int. J. Mol. Sci., 17, 640. (2016)
- La Verde V., Dominici P., Astegno A., bio-protocol, Vol 7, Iss 08, Apr 20, (2017)
- La Verde V., Dominici P., Astegno A., Int. J. Mol. Sci. 19, 1331 (2018)
- La Verde V. (bis), **Trande M.**, D’Onofrio M. Dominici P., Astegno A., International Journal of Biological Macromolecules 108, 1289–1299, (2018)
- Lau S.Y., Taneja A.K., Hodges R.S., J Biol Chem. 10;259(21):13253-61, (1984)
- Leba L.J., Cheval C., Ortiz-Martín I., Ranty B., Beuzón C.R., Galaud J.P., Aldon D., Plant J. Sep;71(6):976-89. (2012)
- Lee D.S., Tomita S., Kirino Y., Suzuki T., The Journal of Biological Chemistry, Vol. 275, No. 30, pp. 23134–23138, (2000)
- Lee WC, Reniere ML, Skaar EP, Murphy ME. J. Biol. Chem. 283, 30957-63, (2008).
- Lemieux MJ, Ference C, Cherney MM, Wang M, Garen C, James MN. J. Struct. Funct. Genomics 6, 245-57, (2005).
- Liang L., Flury S., Kalck V., Hohn B., Molinier J., Plant Molecular Biology 61:345–356, (2006)

- Liao, J.; Deng, J.; Qin, Z.; Tang, J.; Shu, M.; Ding, C.; Liu, J.; Hu, C.; Yuan, M.; Huang, Y.; et al. *Front. Plant Sci.*, 8, 482. (2017)
- Lindahl T., *Nature*;362:709–15. (1993)
- Linse S., Helmersson A., Forsen S., *J. Biol. Chem.* 266, 8050–8054, (1991)
- Linse S., Forsen S., *Adv. Second Messenger Phosphoprotein Res.* 30, 89–151, (1995)
- Lupas, A., Van Dyke, M., and Stock, J., *Science* 252:1162-1164. (1991)
- Ma W., Smigel A., Tsai Y.C., Braam J., Berkowitz G.A., *Plant Physiol.* 148 818e828 (2008)
- Maathuis F.J.; *Curr. Opin. Plant Biol.* 12 (3) 250–258. (2009)
- Madeddu L., Klotz C., Le Caer J.P., Beisson J., *Eur. J. Biochem.* 238, 121-128 (1996)
- Magnan F., Ranty B., Charpentreau M., Sotta B., Galaud J.P., Aldon D., *Plant J.* 56, 575–589. (2008)
- Margui E.; Zawisza B.; Sitko R.; *TrAC, Trends Anal. Chem.* 53, 73–83. (2014)
- Markowicz, A. *Pramana*, 76, 321–329 (2011)
- Marsden B. J., Hodges R. S., Sykes B.D., *Biochemistry* 27, 4198–4206 (1988)
- Marteijs J.A., Lans H., Vermeulen W., Hoeymakers J.H.J., *Nat Rev Mol Cell Biol* 15(7):465–481. (2014)
- Martin R. B., *Book: Bioinorganic Chemistry of Magnesium. In Metal Ions in Biological Systems* (Sigel, H., ed.), pp. 1–13, Marcel Dekker Inc., New York (1990)
- Martinez-Sanz J., Yang A., Blouquit Y., Duchambon P., Assairi L., Craescu C.T., *FEBS Journal* 273 4504–4515 (2006)
- McCormack E., Braam J., *New Phytologist*, 159: 585–598, (2003)
- McCormack E., Tsai Y.C., Braam J., *Trends Plant Sci*, 10(8):383-9, (2005)
- Meador W.E., Means A.R., Quiocho F.A., *Science* 257, 1251–1255, (1992)
- Mertens, M.; Rittmeyer, C.; Kolbesen, B. O. *Spectrochim. Acta, Part B*, 56, 2157–2164 (2001)
- Molinier J., Ramos C., Fritsch O., Hohn B., *Plant Cell*;16(6):1633-43. (2004)
- Mukhi, N.; Dhindwal, S.; Uppal, S.; Kumar, P.; Kaur, J.; Kundu, S. *Biochim. Biophys. Acta, Proteins Proteomics*, 1834, 1944–1956 (2013)
- Munir, S.; Khan, M.R.G.; Song, J.; Munir, S.; Zhang, Y.; Ye, Z.; Wang, T.; *Plant Physiol. Biochem.*, 102, 167–179. (2016)
- Ng J.M., Vermeulen W., van der Horst G.T., et al. *Genes Dev*; 17:1630–45. (2003)

- Nie, S.; Zhang, M.; Zhang, L.; BMC Genom., 18, 842. (2017)
- Ogunrinde A., Munro K., Davidson A., Ubaid M., Snedden W.A., Front. Plant Sci. 8 2175, (2017)
- Ohki S., Ikura M., Zhang M., Biochemistry 36, 4309–4316, (1997)+
- Ouyang H., Vogel H.J., BioMetals, 11, 213–222 (1998)
- Park H.C., Park C.Y., Koo S.C., Cheong M.S., Kim K.E., Kim M.C., et al., Plant Cell Rep. 29, 1297–1304. (2010)
- Perochon A., Aldon D., Galaud J.-P., Ranty B., Biochimie 93, 2048-2053, (2011)
- Polisensky D. H., Braam J., Plant Physiol. 111, 1271–1279. doi: 10.1104/pp.111.4.1271 (1996)
- Popescu A., Miron S., Blouquit Y., Duchambon P., Christova P., Craescu C.T., J. Biol. Chem. 278 (41), 40252–40261, (2003)
- Radu L., Durussel I., Assairi L., Blouquit Y., Miron S., Cox J.A, Craescu C.T., Biochemistry 49 (20) 4383–4394. (2010)
- Raj, S. B.; Ramaswamy, S.; Plapp, B. V. Biochemistry, 53, 5791–5803. (2014)
- Rao S.T., Wu S., Satyshur K.A., Sundaralingam M., Ling K.-Y., Kung C, Protein Sci. 2, 436–447, (1993)
- Reid R. E., Hodges R. S., J. Theor. Biol. 84, 401–444 1 (1980)
- Ringer S., J. Physiol. 4, 29 – 43 (1883)
- Rio-Echevarria I. M., Ponti J., Bogni A., Gilliland D., Altissimo M., Pascolo L., Ceccone G., Gianoncelli A., X-Ray Spectrom. 48, 94–101, (2019)
- Salisbury I.L., Baron A., Surek B. et al. Cell BioI, 99:962-70 (1984)
- Salisbury J.L., Suino K.M., Busby R., Springett M., Curr Biol; 12: 1287–92. (2002)
- Scaloni A., Miraglia N., Orrù S., Amodeo P., Motta A., Marino G., Pucci P., J. Mol. Biol. 277, 945–958, (1998)
- Sciara G., Kendrew S.G., Miele A.E., Marsh N.G., Federici L., Malatesta F., Schimperna G., Savino C., Vallone B., The EMBO Journal Vol. 22 No. 2 pp. 205-215, (2003)
- Scholz, S.S. et al., Mol. Plant 7, 1712–1726 (2014)
- Seeholzer S.H., Cohn M., Putkey J.A., Means A.R., Crespi H.L., Proc. Natl. Acad. Sci. 83, 3634–3638, (1986)
- Sheehan J.H., Bunick C.G., Hu H., Fagan P.A., Meyn S.M., Chazin W.J., J Biol Chem. Feb 3;281(5):2876-81. (2006)
- Shrivastava A., Vipin B. G., Chron. Young Sci., 2, 21-25, (2011)

- Smith J.S., Scholtz J.M., *Biochemistry* 37, 33–40, (1998)
- Sorensen B.R., Eppel J.T., Shea M.A., *Biochemistry* 40 (4),896–903, (2001)
- Spyrakis F., BidonChanal A., Barril X., Luque F.J., *Curr. Top. Med. Chem.* 11 192–210, (2011)
- Sugasawa K., Ng J.M., Masutani C., Iwai S., Van der Spek P.J., Eker A.P., Hanaoka F., Bootsma D., Hoeijmakers J.H., *Mol. Cell.* 2:223–232. (1998)
- Sugita S., Ho A., Südhof T.C. *Neuroscience*;112(1):51-63 (2002)
- Szoboszlai, N.; Polgari, Z.; Mihucz, V. G.; Zaray, G. *Anal. Chim. Acta* 633, 1–18 (2009)
- Teardo E., Carraretto L., Wagner S., Formentin E., Behera S., De Bortoli S., Larosa V., Fuchs P., Lo Schiavo F., Raffaello A., Rizzuto R., Costa A., Schwarzländer M., Szabò I., *Plant Physiol.* 173, 1355–1370, (2017)
- Thompson J.R., Ryan Z.C., Salisbury J.L., Kumar R., *J. Biol. Chem.* 281 (27) 18746–18752, (2006)
- Tidow H., Nissen P., *FEBS J.* 280, 5551–5565, (2013)
- **Trande M.**, Pedretti M., Bonza M.C., Di Matteo A., D’Onofrio M., Dominici P., Astegno A., *Journal of Inorganic Biochemistry* 199, 110796, (2019)
- Trojanowski, J. Q., Goedert, M., Iwatsubo, T. & Lee, V. M. Y. *Cell Death Diff.* 5, 832– 837. (1998)
- Tsai Y.C., Delk N.A, Chowdhury N.I., Braam J., *Plant Signal. Behav.* 2 446e454. (2007)
- Vadassery J. et al., *Plant Physiol.* 159, 1159–1175 (2012)
- Vallone R.; La Verde V., D’Onofrio M., Giorgetti A., Dominici P., Astegno A.,*Protein Sci.* 25, 1461–1471. (2016)
- van der Spek P.J., Eker A., Rademakers S., et al. *Nucleic Acids Res*; 24:2551–9. (1996)
- Vanderbeld B., Snedden W.A., *Plant Mol. Biol.* 64, 683–697. (2007)
- Vekemans B., Janssens K., Vincze L., Adams F., Vanespen P., *X-Ray Spectrom.*, 23, 278-285, (1994)
- Wada T, Shirouzu M, Terada T, Kamewari Y, Park SY, Tame JR, Kuramitsu S, Yokoyama S.; *Proteins* 55, 778-80, (2004).
- Waldron, K. J.; Robinson, N. J. *Nat. Rev. Microbiol.*, 7, 25–35, (2009)
- Wang S.S., Diao W.Z., Yang X., Qiao Z., Wang M., Acharya B.R., Zhang W., *Plant Cell Environ.* Volume38, Issue11, 2372-2386 2015 (2015)

- Weber C., Lee V.D., Chazin W.J., Huang B. *Biol Chem.* 269(22): p. 15795-802. (1994)
- Wiech H., Geier B.M., Paschke T., Spang A., Grein K., Steinkotter J., Melkonian M., Schiebel E., *J. Biol. Chem.* 271 (37), 22453–22461, (1996)
- Weljie A.M., Vogel H.J., *Protein Eng.* 13, 59–66, (2000)
- White, M. D.; Payne, K. A. P.; Fisher, K.; Marshall, S. A.; Parker, D.; Rattray, N. J. W.; Trivedi, D. K.; Goodacre, R.; Rigby, S. E. J.; Scrutton, N. S.; Hay, S.; Leys, D. *Nature* 522, 502. (2015)
- Wilson M.A., Brunger A.T., *J. Mol. Biol.* 301, 1237–1256, (2000)
- Wittershagen, A.; Rostam-Khani, P.; Zickermann, V.; Zickermann, I.; Gemeinhardt, S.; Ludwig, B.; Kolbesen, B. O. *Fresenius' J. Anal. Chem.*, 361, 326–328 (1998)
- Wolf F.I., Torsello A., Fasanella S., Cittadini A.; *Molecular Aspects of Medicine* 24 11–26, (2003)
- Wolfrum U., Giessl A., Pulvermuller A., *Adv Exp Med Biol*, 514, 155-78 (2002)
- Won S.K., Lee Y.J., Lee H.Y., Heo Y.K, Cho M., Cho H.T., *Plant Physiol.* 150 (3) 1459–1473, (2009)
- Wu R, Skaar EP, Zhang R, Joachimiak G, Gornicki P, Schneewind O, Joachimiak A. *J. Biol. Chem.* 280, 2840-6, (2005)
- Yamaguchi T., Aharon G.S., Sottosanto J.B., and Blumwald E., *Proc. Natl. Acad. Sci. U.S.A.* 102, 16107–16112. (2005)
- Yap K.L., Kim J., Truong K., Sherman M., Yuan T., Ikura M., *J Struct Func Genom*, 1: 8. (2000)
- Yiming Y., Hsiau-Wei L., Yang W., Shealy S., Yang J.J., *J. AM. CHEM. SOC.*, 127, 3743-3750 (2005)
- Yoo S.D., Cho Y.-H., Sheen J., *Nat. Protoc.* 2 (7), 1565–1572, (2007)
- Zhang M., Tanaka T., Ikura M., *Nat. Struct. Biol.* 2, 758–767, (1995)
- Zhang, K.; Yue, D.; Wei, W.; Hu, Y.; Feng, J.; Zou, Z.; *Front. Plant Sci.*, 7, 1820; (2016)
- Zhou N.E., Kay C.M., Hodges R.S., *J Biol Chem.* 267: 2664-2670, (1992)
- Zhou Y, Yang W, Kirberger M., Lee H.W., Ayalasomayajula G., Yang J.J., *PROTEINS: Structure, Function, and Bioinformatics* 65:643–655 (2006)
- Zhu X., Dunand C., Snedden W., and Galaud J.-P., *Trends Plant Sci.* 20, 483–489. (2015)



# *Appendices*

## *Papers*



## Cation and peptide binding properties of CML7, a calmodulin-like protein from *Arabidopsis thaliana*

Matteo Trande<sup>a</sup>, Marco Pedretti<sup>a</sup>, Maria Cristina Bonza<sup>b</sup>, Adele Di Matteo<sup>c</sup>, Mariapina D'Onofrio<sup>a</sup>, Paola Dominici<sup>a</sup>, Alessandra Astegno<sup>a,\*</sup>

<sup>a</sup> Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy

<sup>b</sup> Department of Biosciences, University of Milano, Via Celoria 26, 20133 Milano, Italy

<sup>c</sup> Institute of Molecular Biology and Pathology, CNR, Piazzale Aldo Moro 5, Roma 00185, Italy

### ARTICLE INFO

#### Keywords:

Calmodulin-like proteins  
Calcium sensor  
EF-hand  
*Arabidopsis thaliana*  
Melittin

### ABSTRACT

Plants contain a large family of so-called calmodulin-like proteins (CMLs) which differ from canonical calmodulin in that they show greater variability in sequence, length, and number of EF-hand domains. The presence of this extended CML family has raised questions regarding the role of these proteins: are they functionally redundant or do they play specific functions in physiological plant processes? To answer these questions, comprehensive biochemical and structural information on CML proteins is fundamental. Among the 50 CMLs from *Arabidopsis thaliana*, herein we described the ability of CML7 to bind metal ions focusing on the Ca<sup>2+</sup> and Mg<sup>2+</sup> sensing properties, as well as on metal-induced conformational changes. Circular dichroism and nuclear magnetic resonance (NMR) studies indicated that both Ca<sup>2+</sup> and Mg<sup>2+</sup> stabilize CML7, as reflected in conformational rearrangements in secondary and tertiary structure and in increases in thermal stability of the protein. However, the conformational changes that binding induces differ between the two metal ions, and only Ca<sup>2+</sup> binding controls a structural transition that leads to hydrophobic exposure, as suggested by 8-anilino-1-naphthalenesulfonic acid fluorescence. Isothermal titration calorimetry data coupled with NMR experiments revealed the presence of two high affinity Ca<sup>2+</sup>-binding sites in the C-lobe of CML7 and two weaker sites in the N-lobe. The paired nature of these CML7 EF-hands enables them to bind Ca<sup>2+</sup> with positive cooperativity within each globular domain. Our results clearly place CML7 in the category of Ca<sup>2+</sup> sensors. Along with this, the protein can bind to a model target peptide (melittin) in a Ca<sup>2+</sup>-dependent manner.

### 1. Introduction

In plants, calcium (Ca<sup>2+</sup>) ions play key roles in the regulation of many physiological processes. A significant diversity and abundance of proteins that bind Ca<sup>2+</sup> using the EF-hand, an evolutionarily preserved helix-loop-helix motif, have evolved in plants, including calmodulin (CaM) and calmodulin-like proteins (CMLs). CaMs and CMLs are largely considered to act as Ca<sup>2+</sup> sensors that can mediate cellular responses through a well-known mechanism in which the Ca<sup>2+</sup>-bound form of these proteins can interact with specific downstream targets to modulate cellular activity [1].

After plant genome sequencing projects were completed, the presence of a multitude of genes predicted to encode either identical or highly homologous isoforms of CaM and various CMLs has emerged as a typical feature in plants. CaM/CML gene families are represented by seven CaM and 50 CML genes in *Arabidopsis thaliana* [2] and five CaM

and 32 CML genes in rice [3]. Additionally, lists of CaM and CML genes have been identified in many vegetable and fruit crops, such as tomato, [4], strawberry [5], grapevine [6], and cabbage [7]. Interestingly, while CaMs are well-conserved in all eukaryotes, the CML family is unique to the plant kingdom. Similarly to CaM, CMLs are characterized by the presence of EF-hand motifs and no other known functional domains [2].

The presence of a large CML family in plants has raised fundamental and still open questions regarding the evolutionary selection in preserving such apparently redundant family in the plant genome. To this end, the CML field has experienced an explosion of investigations in the last years, mainly focused on: i) elucidation of the metal binding ability and the structural features of CMLs [8–14]; ii) identification of specific targets of CMLs as well as the processes in which they are involved [10,15–22]; iii) determination of the expression patterns and regulation at the protein and messenger RNA levels during different processes, e.g.

\* Corresponding author.

E-mail address: [alessandra.astegno@univr.it](mailto:alessandra.astegno@univr.it) (A. Astegno).

<https://doi.org/10.1016/j.jinorgbio.2019.110796>

Received 9 April 2019; Received in revised form 26 July 2019; Accepted 31 July 2019

Available online 01 August 2019

0162-0134/ © 2019 Elsevier Inc. All rights reserved.



plant development or response to external stimuli [4–7,23]. Acquired knowledge has provided evidence that CMLs are not a simple CaM safety backup system; rather, they are likely to play specific functions in coordinating plant responses to endogenous and environmental stimuli. However, although significant insights into the function and regulation of various CMLs have been provided, and a holistic view of how the CML genes function at different levels is emerging, how this family can contribute to cellular signaling and still maintain properties that are dissimilar from CaM and other  $\text{Ca}^{2+}$  sensors is still elusive.

Accumulating evidence suggests that the biochemical properties of individual CMLs, such as the number of  $\text{Ca}^{2+}$  binding sites, cooperativity, metal affinity, capacity to distinguish between  $\text{Ca}^{2+}$  and other physiological ions (i.e.  $\text{Mg}^{2+}$ ), and conformational rearrangement upon binding to metal ions are crucial parameters that contribute to the specificity of various  $\text{Ca}^{2+}$  mediated responses. Sequence divergence among CMLs may lead to different binding properties, creating diverse subsets of CMLs capable of perceiving stimulus-specific  $\text{Ca}^{2+}$  signals and regulate specific targets, thus coupling the perception of different signals to precise physiological responses. In this light, the comparison of the  $\text{Ca}^{2+}$  binding and structural properties of CMLs which are related to conserved eukaryotic CaM with those more divergent from CaM may help to distinguish between specificity and redundancy in the plant CaM/CML families.

According to a recent phylogenetic analysis performed taking into the account the occurrence of CaM and CML proteins in the genomes of the green lineage [24], CMLs recognized as being closely related to CaMs (subgroup Ia) were distributed to subgroups Ib and Ic, while CMLs that diverge more from conserved CaMs could be categorized in seven subgroups (from II to VIII). Herein, we investigated the isoform CML7 from *A. thaliana* (At1g05990). Interestingly, even if CML7 has been classified as a member of subgroup VII, it possesses significant sequence similarity with CaM2 from *A. thaliana* (~ 45% of amino acid sequence identity). CML7 is a 16.9 kDa protein composed of 150 amino acid residues and characterized by four EF-hands that are predicted to be functional  $\text{Ca}^{2+}$  binding sites by PROSITE ProRule annotation [25], based on the presence of canonical  $\text{Ca}^{2+}$ -coordinating groups [26]. These observations suggest the potential capacity of CML7 to bind  $\text{Ca}^{2+}$  *in vitro*.

Functional analyses, albeit limited, have revealed that CML7 is likely involved in root development processes, in which the protein appears to act as a negative regulator, inhibiting hair tip growth [27]. However, its properties as a  $\text{Ca}^{2+}$  sensor, its downstream targets, and the mechanism of action by which CML7 can negatively transduce  $\text{Ca}^{2+}$  signals into a physiological response have not been elucidated yet.

Accordingly, we characterized CML7 from *A. thaliana* in terms of  $\text{Ca}^{2+}$  binding affinity, structural properties, and conformational perturbation upon metal binding using spectroscopic and calorimetric techniques, and we determined its subcellular localization using confocal microscopy. We examined the behavior of CML7 also in the presence of  $\text{Mg}^{2+}$ , a physiologically relevant cation that, due to its chemical properties similar to  $\text{Ca}^{2+}$  and its high concentration in plant cells (*in vivo*, cytosolic  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  concentrations are about 100–200 nM and 0.5–1 mM, respectively) could represent a possible  $\text{Ca}^{2+}$  competitor towards CML7 binding [28]. Moreover, we studied the interaction of CML7 with the model CaM-binding peptide melittin.

We believe that the structural information obtained in our study, in addition to the demonstration that CML7 behaves as a  $\text{Ca}^{2+}$  sensor, will help aid in better understanding the specificity of CML7 in the evolutionary panorama of the multigene family of CMLs present in plants.

## 2. Experimental procedures

### 2.1. DNA constructs

The cDNA sequence for CML7 (At1g05990) was obtained from the Arabidopsis Information Resource (TAIR) into a pUNI51 vector. The

coding sequence was polymerase chain reaction (PCR)-amplified using the following primers: forward, 5'-CATATGGATCCGACAGAGCTAA AAC, and reverse, 5'-GGATCCCTACAAAGAATTTAAACCACCG. The forward primer carried a *NdeI* restriction site (bold), while the reverse primer contained a stop codon and a *BamHI* restriction site (bold). For protein expression, the cDNA sequence was further cloned using standard laboratory protocols into pET21a vector using *NdeI/BamHI* restriction sites and the plasmid construct was verified by DNA sequencing.

For CML7 subcellular localization studies, the complete coding sequence (CDS) of the red fluorescent protein (RFP) was amplified by PCR using Phusion® High-Fidelity DNA Polymerase (New England Biolabs, USA) according to the manufacturer's protocol. The following primers, forward 5'-GCGGCCGCGATGGCTCTCCGAGGACG-3' and reverse 5'-GCGGCCGCTACCCGGGTGCTCCAGTACTGT-3' were used to add a *NotI* site (bold) to both ends of the PCR product. Following *NotI* digestion, RFP was inserted into the polylinker of the plant expression vector pGREEN0029::ter [29] under the control of the cauliflower mosaic virus (CaMV) 35S promoter (pGREEN0029::RFP). The complete CDS of CML7 was PCR amplified using forward 5'-ACGCTCGAGATGG ATCCGACAGAGCTAAAACG-3' and reverse 5'-CGGAATTTCCAAAGAAT TAAACCAC-3' primers to add *XhoI* and *EcoRI* restriction sites (bold) at 5' and 3' respectively during amplification. The reverse primer also eliminates the stop codon at the 3' of CML7 CDS. Following a *XhoI/EcoRI* digestion, CML7 CDS was inserted into pGREEN0029::RFP for the in frame fusion at the C-terminus with RFP. Sequencing confirmed the accuracy of the construct.

### 2.2. Protein production

Recombinant CML7 was expressed in *E. coli* strain BL21(DE3)pLysS (Stratagene) and purified by heat treatment and  $\text{Ca}^{2+}$ -dependent phenyl-Sepharose chromatography. Expression was induced by adding 0.4 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) to 1 L of bacterial culture [pET21a-CML7] grown in LB medium supplemented with 100  $\mu\text{g}/\text{mL}$  ampicillin and 34  $\mu\text{g}/\text{mL}$  chloramphenicol to exponential phase and incubating at 24 °C for 16 h.

Cells were isolated by centrifugation (5000g for 15 min), and the pellet was resuspended in 50 mM Tris-HCl, 150 mM KCl, 0.5 mM dithiothreitol (DTT) pH 7.5 buffer containing 0.1 mg/mL lysozyme, stirred at room temperature for 30 min, and then sonicated on ice. After removing cell debris by centrifugation, the supernatant was subject to heat treatment (6 min at 85 °C followed by 40 min on ice). The precipitate was removed by centrifugation and the heat-treated lysate containing 5 mM  $\text{CaCl}_2$  was loaded onto a phenyl-Sepharose column, previously equilibrated with buffer containing 50 mM Tris-HCl, 0.5 mM  $\text{CaCl}_2$ , 0.5 mM DTT pH 7.5. The column was then washed with the same buffer containing 150 mM KCl. CML7 was eluted in 50 mM Tris-HCl, 150 mM KCl pH 7.5, 2 mM ethylene glycol-bis(2-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid (EGTA) buffer. The purity of CML7 was estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis and the protein concentration was determined by the Bradford assay [30]. Final yield was about 20 mg/L culture.

$^{15}\text{N}$  and  $^{13}\text{C}$  labeled CML7 samples for NMR experiments were obtained using M9 minimal medium supplemented with 1 g/L  $^{15}\text{NH}_4\text{Cl}$  and 4 g/L  $^{13}\text{C}$ -glucose as the sole sources of nitrogen and carbon. In M9 medium the final yield was about 10 mg/L culture.

CaM1 from *A. thaliana* was produced as described [31].

### 2.3. Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker Avance III spectrometer (Bruker, Karlsruhe, Germany) operating at 600.13 MHz proton Larmor frequency, and equipped with a cryogenic probe.  $^1\text{H}$ - $^{15}\text{N}$  Heteronuclear Single Quantum Coherence (HSQC) experiments were recorded at 298 K on a  $^{15}\text{N}$  labeled sample, in 50 mM

Tris-HCl, 50 mM KCl, 0.5 mM DTT, 7% D<sub>2</sub>O pH 7.5 at protein concentration of 270 μM in the presence of EGTA (5 mM), or MgCl<sub>2</sub> (5 mM EGTA + 10 mM MgCl<sub>2</sub>) or CaCl<sub>2</sub> (5 mM). The data matrix consisted of 2048 (F2, <sup>1</sup>H) × 256 (F1, <sup>15</sup>N) complex data points, spectral windows of 8417.509 Hz (<sup>1</sup>H) × 2189.445 Hz (<sup>15</sup>N), 8 transients, and 1.5 s relaxation delay. Standard triple resonance experiments, HNCACB, CBCA(CO)NH, and <sup>1</sup>H-<sup>15</sup>N Nuclear Overhauser Effect Spectroscopy (NOESY) HSQC were recorded at 298 K on <sup>15</sup>N,<sup>13</sup>C-labeled samples at protein concentration of 500 μM to achieve partial sequence-specific backbone atoms assignment of CML7 bound to Ca<sup>2+</sup>. The HNCACB, and CBCA(CO)NH spectra were acquired with a data matrix consisting of 2 K (F3, <sup>1</sup>H) × 90 (F2, <sup>15</sup>N) × 120 (F1, <sup>13</sup>C) complex points, spectral windows of 8417.51 Hz (<sup>1</sup>H) × 2189.42 Hz (<sup>15</sup>N) × 11,318.16 Hz (<sup>13</sup>C), 8 transients, and 1.5 s relaxation delay. The <sup>1</sup>H-<sup>15</sup>N NOESY HSQC spectrum was acquired with a data matrix consisting of 2 K (F3, <sup>1</sup>H) × 40 (F2, <sup>15</sup>N) × 128 (F1, <sup>1</sup>H) complex points, spectral windows of 8417.51 Hz (<sup>1</sup>H) × 2189.42 Hz (<sup>15</sup>N) × 8417.51 Hz (<sup>1</sup>H), 16 transients, and 120 ms mixing time. Data processing was performed with Topspin3.2 (Bruker, Karlsruhe, Germany).

#### 2.4. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed on a TA nanocalorimeter (TA Instruments). In each experiment, 1.5 μL injections of 3 mM CaCl<sub>2</sub> were made to a 200 μL sample cell containing 200 μM of apo- or Mg<sup>2+</sup>-bound CML7 in 50 mM Tris-HCl, 150 mM KCl, 0.5 mM DTT pH 7.5 buffer. Mg<sup>2+</sup>-loaded CML7 was obtained by incubation of apo-protein with 10 mM MgCl<sub>2</sub> for 15 min at 25 °C. Experiments were performed at 25 °C with injections at 300-s intervals and 250 rpm stirring speed. The buffer was decalcified using a Chelex-100 ion-exchange resin (Sigma) as described [8,32] and all solutions were degassed prior to each experiment. Control experiments (injecting ligand solutions into the buffer) were conducted to obtain a baseline for each experiment and determine the heat of dilution/mixing. These values were subtracted from the experimental runs in the presence of protein. Analysis of the data was performed using Origin-based software. The reported values represent the mean ± SEM of at least three independent titrations using at least two different CML7 preparations.

#### 2.5. Circular dichroism spectroscopy

All circular dichroism (CD) spectra were recorded on a Jasco J-1500 spectropolarimeter equipped with a thermostatically controlled Peltier type compartment, using the protocols described in [14,15,33]. Briefly, near UV (320–250 nm) and far UV (250–200 nm) spectra of ~ 100 μM and ~12 μM CML7, respectively, were collected at 25 °C after consecutive additions of 1 mM EGTA, 2 mM MgCl<sub>2</sub> and 2 mM CaCl<sub>2</sub>. Quartz cuvettes were used both for near UV (1 cm) and far UV (0.1 cm). CD spectra of protein/peptide complex were obtained using a protein:peptide ratio of 1:1 in 50 mM Tris-HCl, 150 mM KCl, 0.5 mM DTT pH 7.5 in the presence of 2 mM EGTA or 2 mM CaCl<sub>2</sub>. Spectra from five scans were averaged and were corrected for background signal by subtracting out the spectrum of the buffer. Denaturation profiles were recorded in the same conditions as for far UV spectra by following CD signal at 222 nm between 4 and 96 °C (scan rate 90 °C/h).

#### 2.6. Fluorescence spectroscopy

8-Anilino-1-naphthalene-sulfonate (ANS) fluorescence spectra were measured using a Jasco FP8200 spectrofluorometer to probe the changes in hydrophobicity of CML7 following Mg<sup>2+</sup> and Ca<sup>2+</sup> binding, as described in [14,33]. Protein at 1 μM was incubated with 15 μM ANS and fluorescence was measured after 1 h incubation at room temperature with EGTA (2 mM), or MgCl<sub>2</sub> (0.5 mM EGTA + 5 mM MgCl<sub>2</sub>) or CaCl<sub>2</sub> (2 mM). ANS fluorescence spectra were recorded at room temperature with slit widths set to 5 nm. Excitation wavelength was

380 nm and emission spectra were collected in the range 400–650 nm.

For the CML7/melittin and AtCaM1/melittin interaction the single Trp residue of melittin was excited at 295 nm, and fluorescence emission spectra were recorded with a single scan over the range 305–450 nm. Melittin (2 μM) was titrated with increasing concentrations of AtCaM1 or CML7 in 50 mM Tris-HCl, 150 mM KCl, 0.5 mM DTT pH 7.5 in the presence of 2 mM CaCl<sub>2</sub>. Titration experiments were performed by monitoring the blue-shift in wavelength (λ) of the peptide emission peak. The dissociation constant (K<sub>d</sub>) was determined according to a tight binding hypothesis [34–36]:

$$fb = \frac{\lambda - \lambda_0}{\lambda_{\max} - \lambda_0} = \frac{e_0 + x + K_d - \sqrt{[(e_0 + x + K_d)^2 - 4e_0x]}}{2e_0} \quad (1)$$

where fb is the fraction bound, x is the protein concentration, e<sub>0</sub> is the peptide concentration, λ is the value of wavelength of the peptide emission peak as a function of protein concentration, λ<sub>0</sub> is the value at zero protein concentration and λ<sub>max</sub> is the value at saturating protein concentration.

#### 2.7. Limited proteolysis

CML7 at 0.5 mg/mL concentration was digested with trypsin (1:500 w/w) in 50 mM Tris-HCl, 150 mM KCl, 0.5 mM DTT pH 7.5 in the presence of 5 mM CaCl<sub>2</sub> or 5 mM EGTA, respectively, at 25 °C. At various times (0, 1, 5, 10, 20, 40, 60, 120, and 180 min), 20 μL aliquots were taken for 15% SDS-PAGE analysis. The reaction was stopped by sample boiling and addition of reducing Laemmli buffer. Bands in denaturing gel were visualized by Coomassie Blue staining.

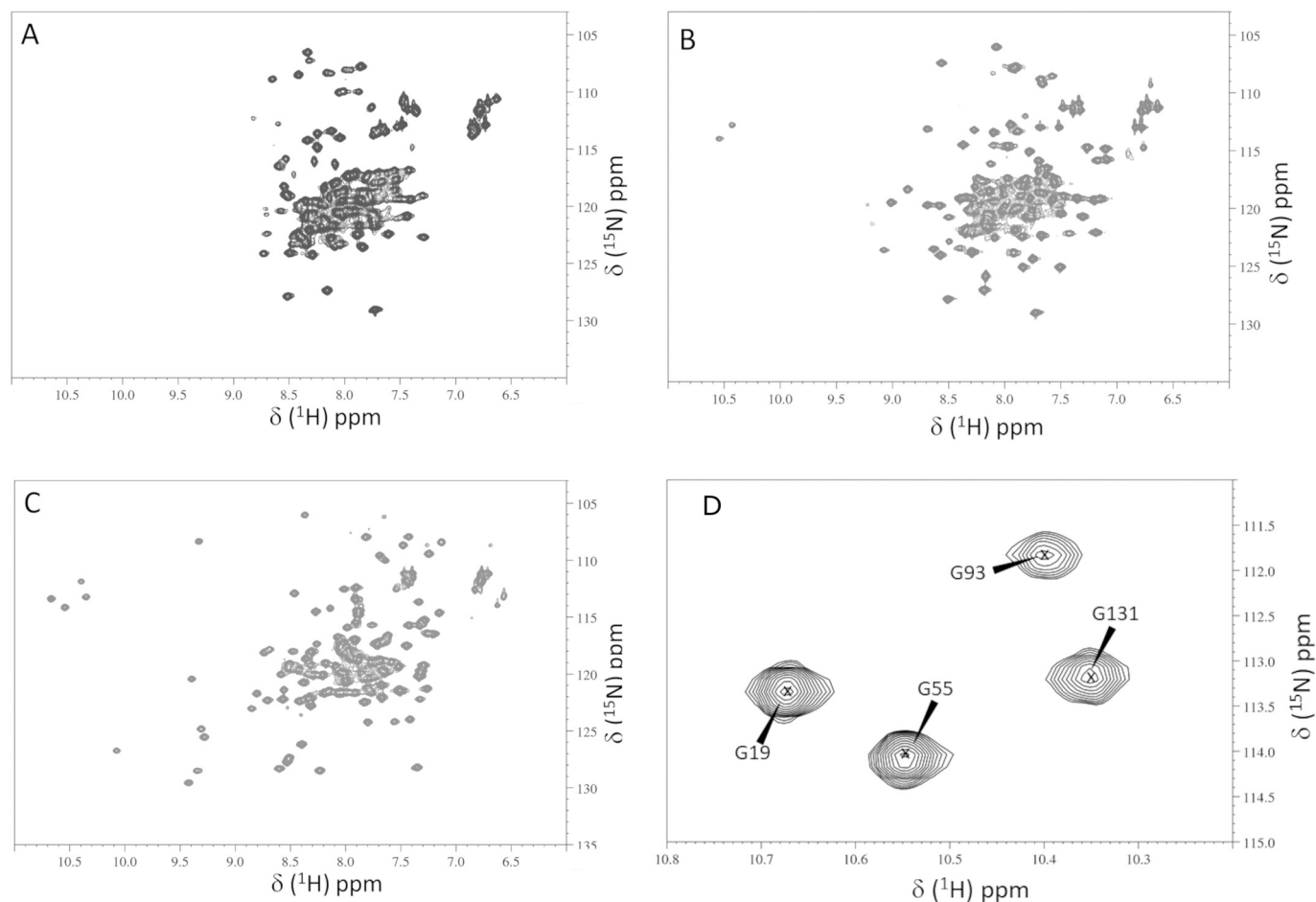
#### 2.8. Ca<sup>2+</sup>-dependent electrophoresis mobility shift assay

CML7 recombinant protein was analyzed by SDS-PAGE (15%) or native PAGE (12%) after incubation (30 min at room temperature) with 5 mM CaCl<sub>2</sub> or 5 mM EGTA in the presence and absence of DTT, as previously described [8,14,32].

Peptide-protein stoichiometry was analyzed by incubating AtCaM1 or CML7 (20 μM) for 20 min at room temperature with increasing concentrations of melittin in 100 mM Tris-HCl pH 7.5 in the presence of 5 mM CaCl<sub>2</sub> or 5 mM EGTA, as described [33]. Samples were analyzed under native conditions on a 12% continuous gel containing 5 mM CaCl<sub>2</sub> or 5 mM EGTA.

#### 2.9. A. thaliana protoplast isolation and transformation

Mesophyll protoplasts were isolated and transiently transformed according to [37]. Leaves from four week old *A. thaliana* Col-0 plants were collected and cut in 0.5–1 mm strips with fresh razor blades and put into multiwell cell culture in the presence of the following enzymatic plasmolysis mix: 20 mM KCl, 0.4 M D-Mannitol, 20 mM 2-(N-Morpholino)ethanesulfonic acid, 4-Morpholineethanesulfonic acid (MES), 1.25% R10 Cellulase, 0.3% R10 Macerozyme. Enzymes were pre-heated at 55 °C for 10 min and 10 mM CaCl<sub>2</sub> and 0.1% bovine serum albumin (BSA) were added to the mix after the incubation. The enzymatic mix was filtered and added to the cut leaves. A 10 min vacuum favored enzymes entering the leaves through the cuticle. Digestion was performed in the dark for about 3 h without shaking. Protoplasts were separated from the cell debris by filtering the mixture through a nylon sieve of 50 μm pore size. Protoplasts were spun at 100g for 6 min in presence of 50 mM CaCl<sub>2</sub>, washed once in cold W5 washing solution (154 mM NaCl, 125 mM CaCl<sub>2</sub>, 5 mM KCl, 2 mM MES) and resuspended in the same solution. After incubation on ice for 30 min, protoplasts were resuspended in 600 μL MMg (0.4 M D-Mannitol, 15 mM MgCl<sub>2</sub>, 4 mM MES pH 5.7) and centrifuged at 100g for 1 min. For the co-transformation with the free enhanced green fluorescent protein (EGFP), the DNA constructs (5–10 μg of pGREEN0029::CML7-RFP and



**Fig. 1.** Two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of apo- (A),  $\text{Mg}^{2+}$ -bound (B),  $\text{Ca}^{2+}$ -bound (C)  $^{15}\text{N}$ -CML7 and zoom of the Gly<sub>6</sub> downfield peaks of  $\text{Ca}^{2+}$ -bound CML7 (D). The assignment of the Gly of the EF-hands is indicated by the position number of the residue. The spectra were recorded at 600 MHz at 298 K. All samples were at a protein concentration of 270  $\mu\text{M}$  in 50 mM Tris-HCl, 50 mM KCl, 0.5 mM DTT, 7%  $\text{D}_2\text{O}$  pH 7.5.

pGREAT::eGFP; [38]) were mixed together before adding them to 100  $\mu\text{L}$  MMg isolated protoplasts and then an equal amount of polyethylene glycol (PEG) solution (40% v/v PEG4000, 0.2 M D-Mannitol, 1 mM  $\text{CaCl}_2$ ) was added, mixed gently and incubated at room temperature for 20 min. The protoplasts suspension was then diluted by adding a 4-fold volume of W5 to remove PEG. The supernatant was carefully removed after 3 min centrifugation at 100g and the transformed protoplasts resuspended in 1 mL W5 solution and maintained at 23  $^\circ\text{C}$  in the dark. Protoplasts were microscopically analyzed 16–24 h after transformation.

### 2.10. Confocal laser scanning microscopy

Confocal microscopy analysis was performed using a Nikon Eclipse Ti2 inverted microscope, equipped with a Nikon A1R+ laser scanning device (Nikon, <http://www.nikon.com/products/microscope-solutions/lineup/confocal/a1/index.htm>). EGFP and chlorophyll were excited with the 488 nm laser and the emission was collected at 500–550 nm and 663–738 nm, respectively. RFP was excited with the 561 nm laser and the emission was collected at 570–620 nm. The pinhole was set to 1.2 airy unit. Images were acquired by a CFI Plan Apo Lambda 60X Oil (N.A. 1.4) objective and were analyzed using Fiji software (<https://fiji.sc>).

### 2.11. Homology modelling and sequence analysis

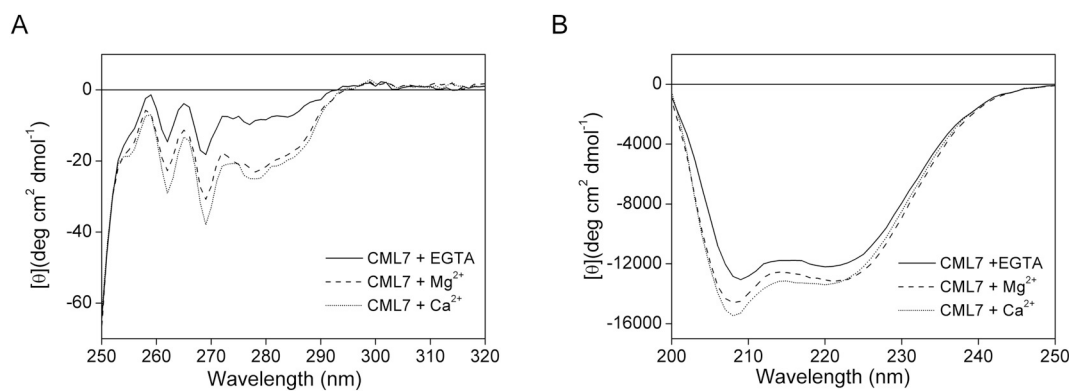
The homology model of the CML7 from *A. thaliana* in the  $\text{Ca}^{2+}$ -

bound form was generated using SWISS-MODEL Homology Modelling server ([swissmodel.expasy.org](http://swissmodel.expasy.org)) [39,40]. The template chosen for modelling procedure was the CaM from *Paramecium tetraurelia* (Protein Data Bank (PDB) entry: 1clm [41]) that shares with CML7 the higher sequence identity (42.0%) among proteins whose structures have been solved so far. The Global Model Quality Estimation (GMQE) index was 0.77, and the Qualitative Model Energy Analysis (QMEAN) potential scoring function was 1.21, indicating a good reliability for the obtained model. Sequence alignments were performed with NPS@ server [42]. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) (<http://www.cgl.ucsf.edu/chimera>). Molecular surface was coloured with the Coulombic Surface colouring (colouring parameters:  $\mu = 4$ , thresholds  $\pm 12$  kcal mol $^{-1}$  at 298 K) implemented in Chimera (<http://www.cgl.ucsf.edu/chimera>).

## 3. Results

### 3.1. CML7 production

CML7 from *A. thaliana* was successfully expressed as a recombinant protein in *E. coli* and purified to homogeneity by heat treatment and hydrophobic interaction chromatography; CML7 binds to a phenyl-sepharose column in the presence of  $\text{Ca}^{2+}$  and is released by the addition of EGTA (see Section 2.2).  $\text{Ca}^{2+}$ -dependent hydrophobic binding to phenyl-sepharose is typical of classic  $\text{Ca}^{2+}$  sensors, and thus represents



**Fig. 2.** Near and far UV CD spectra of CML7. (A) Near UV CD spectra of 100  $\mu\text{M}$  CML7 and (B) far UV CD spectra of 12  $\mu\text{M}$  CML7 after sequential addition of 1 mM EGTA (solid line), 2 mM  $\text{MgCl}_2$  (dashed line) and 2 mM  $\text{CaCl}_2$  (dotted line).

a first clue indicating that CML7 has  $\text{Ca}^{2+}$  sensor properties. Moreover, CML7 displayed a large electrophoretic mobility shift in the presence of  $\text{Ca}^{2+}$ , in both native and denaturing conditions (Supplementary Fig. S1), which is another characteristic of  $\text{Ca}^{2+}$  sensors. Both apo- and  $\text{Ca}^{2+}$ -bound CML7 show a single band on gel electrophoresis in the presence of DTT (Supplementary Fig. S1, lanes 1 and 2), while additional bands were present in the absence of DTT for both forms, independently of the presence of  $\text{Ca}^{2+}$  (Supplementary Fig. S1, lanes 3 and 4). This behavior is likely due to the presence of oligomers through formation of intermolecular disulphide bridges via Cys oxidation.

### 3.2. Apo-, $\text{Mg}^{2+}$ - and $\text{Ca}^{2+}$ -bound states of CML7 exist in distinct structural conformations

A hallmark of CaM and related  $\text{Ca}^{2+}$  sensors is the ability to experience a reversible conformational change upon  $\text{Ca}^{2+}$  binding that leads to the exposure of hydrophobic regions through which these proteins bind and regulate their downstream targets. Two-dimensional NMR spectra ( $^1\text{H}$ - $^{15}\text{N}$  HSQC) of  $^{15}\text{N}$ -labeled CML7 were collected to monitor conformational changes in CML7 resulting from  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  binding to the protein (Fig. 1). In each spectrum, the peaks are representative of protons in the main and side chain amides, and their dispersion is a good probe of protein conformation. As shown in Fig. 1, the apo-,  $\text{Mg}^{2+}$ - and  $\text{Ca}^{2+}$ -bound states of CML7 exist in distinct structural conformations. Easily observed in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra, upon addition of either  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  ions to the apo-protein a better dispersion of NMR signals was observed, and new peaks appeared, including downfield shifted peaks between 10 and 11 ppm in the  $^1\text{H}$  dimension and 111–115 ppm in the  $^{15}\text{N}$  dimension, which are indicative of metal binding in the EF-hand motifs. In particular, the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra indicate that the  $\text{Mg}^{2+}$ -bound conformation of CML7 is separate from both its apo- and  $\text{Ca}^{2+}$ -bound states (Fig. 1). The broadening of the existing resonances line and the fewer number of peaks observed in the spectrum of  $\text{Mg}^{2+}$ -CML7, compared to that of the  $\text{Ca}^{2+}$ -CML7, suggest that  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  bind to CML7 with different affinities and that the  $\text{Mg}^{2+}$ -bound form of CML7 may still possess a degree of flexibility. Despite this partial conformational instability, the dispersion of resonances observed in the two-dimensional spectrum of  $\text{Mg}^{2+}$ -CML7 is distinct from the narrow range of  $^1\text{H}$  chemical shifts observed in the apo-CML7 spectrum and favors a more folded conformation of the protein when bound to  $\text{Mg}^{2+}$ . The addition of  $\text{Ca}^{2+}$  to CML7 promoted a further change in chemical shifts that describes the transition between a closed apo conformation and an open  $\text{Ca}^{2+}$ -bound state, analogous to that observed with both mammalian and plant CaMs [8–10,13–15,43,44].

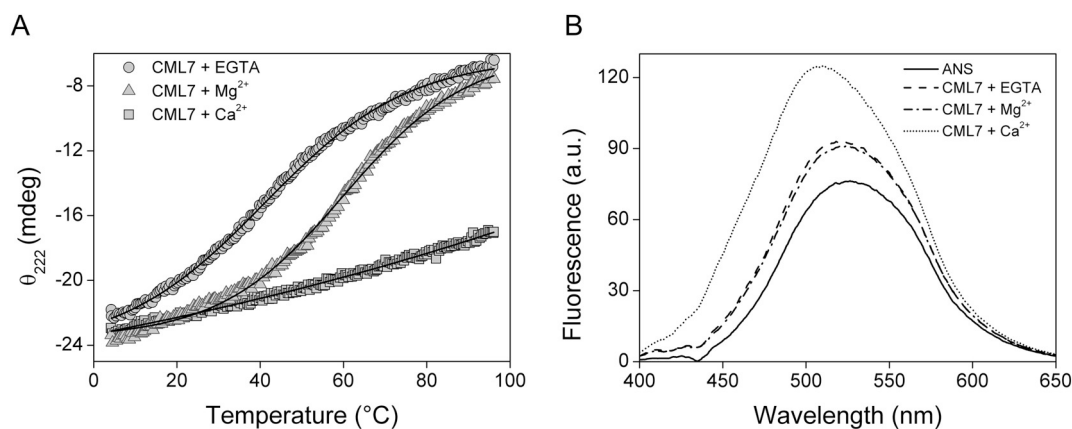
Many EF-hand proteins, including CML7, are characterized by a hydrogen bond network in metal-bound loops that causes the appearance of downfield-shifted peaks in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra, which

belong to the backbone amides of glycine residues present at position 6 (Gly<sub>6</sub>) of the EF-hands when bound to metal [14,15,43,45,46]. The extent of this downfield shift enables these resonances to act as indicators for the  $\text{Ca}^{2+}$ -bound state of a given EF-hand, providing indication on the stoichiometry of ion binding. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $\text{Ca}^{2+}$ -bound CML7 showed four downfield shifted resonances (Fig. 1). Based on standard triple resonance NMR experiments acquired on  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled CML7 in its  $\text{Ca}^{2+}$  bound state, we assigned the four resonances at positions 10.68, 10.55, 10.40 and 10.35 ppm ( $^1\text{H}$  dimension) to the Gly<sub>6</sub> of the EF-hands, Gly19 (EF-1), Gly55 (EF-2), Gly93 (EF-3), and Gly131 (EF-4), respectively (Fig. 1D). Thus, the presence of four low-field resonances corresponding to Gly<sub>6</sub> of the four EF-hands strongly suggests that CML7 can bind  $\text{Ca}^{2+}$  ions with a stoichiometry of 1:4 protein to metal ratio.

The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectral features of CML7 were consistent with our CD experiments. The near UV CD fingerprint of CML7, which reflects protein tertiary structure through local arrangements of aromatic residues, closely resembles that of CaM. Bands at 269 and 262 nm arising from the side chains are very well resolved with a broad minimum for Tyr at around 280 nm. No peaks corresponding to Trp residue are present as CML7 lacks Trp. Binding of  $\text{Mg}^{2+}$  to apo-CML7 produced significant changes in the near UV CD spectrum, with a decrease in the negative ellipticity both in the Phe and Tyr bands (Fig. 2A, dashed line). Addition of  $\text{Ca}^{2+}$  refined the spectrum, slightly increasing the signal intensity, consistently with further changes in the tertiary structure upon  $\text{Ca}^{2+}$  binding (Fig. 2A, dotted line). These results were supported by CD spectra in the far UV region which indicated CML7 to possess significant  $\alpha$ -helical structure and characteristic negative ellipticity bands at 222 nm and 208 nm in both the apo- and metal-bound forms (Fig. 2B). Upon  $\text{Ca}^{2+}$  binding, CML7 showed a change of the spectrum characterized by an increase in intensity and variations in shape. These are visible as a change in the  $\theta_{222}/\theta_{208}$  ratio (0.94 for apo-CML7 and 0.86 for  $\text{Ca}^{2+}$ -bound CML7), which could be ascribed mainly to helix reorientation within the tertiary structure, and not merely to a change in  $\alpha$ -helical content [47,48].

CD data clearly indicated that both  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  ions induce conformational changes, affecting the secondary and tertiary structure of CML7. The metal binding also resulted in an increase of the thermal stability of the protein. Fig. 3A shows the thermal unfolding profiles of apo- and metal-bound CML7 obtained by following the CD signal at 222 nm as a function of temperature. In comparison to the sample containing CML7 and EGTA, the melting temperature ( $T_m$ ) of CML7 with  $\text{Mg}^{2+}$  increased from  $40.9 \pm 0.8$  °C to  $60.0 \pm 0.4$  °C (Table 1). In the presence of  $\text{Ca}^{2+}$ , no estimate of  $T_m$  was possible as the denaturation curve increases monotonically with increasing temperature up to  $\sim 95$  °C, and even at 95 °C the protein maintains a partially folded state. Along with this, the  $\text{Ca}^{2+}$ -bound form of CML7 was more stable against proteolysis by trypsin compared to the apo-form (Supplementary Fig.





**Fig. 3.** CML7 thermal stability and ANS fluorescence. (A) Thermal denaturation profiles of 12  $\mu\text{M}$  CML7 in the presence of 1 mM EGTA (circle), 2 mM  $\text{MgCl}_2$  (triangle) and after addition of 2 mM  $\text{CaCl}_2$  (square). (B) ANS fluorescence spectra of CML7 upon incubation with 2 mM EGTA (dashed line), 0.5 mM EGTA plus 5 mM  $\text{MgCl}_2$  (dot dashed line) and 2 mM  $\text{CaCl}_2$  (dotted line). The fluorescence spectrum of ANS alone is also shown (solid line).

**Table 1**

Analysis of ANS and CD thermal denaturation data for CML7.

	Blue shift (nm) <sup>a</sup>	$F_{\text{max}}/F_{\text{ref}}$ <sup>b</sup>	$T_m$ ( $^{\circ}\text{C}$ ) <sup>c</sup>
CML7 + EGTA	$7 \pm 1$	$1.2 \pm 0.1$	$40.9 \pm 0.8$
CML7 + $\text{MgCl}_2$	$6 \pm 2$	$1.2 \pm 0.1$	$60.0 \pm 0.4$
CML7 + $\text{CaCl}_2$	$21 \pm 2$	$1.6 \pm 0.1$	$> 95$

<sup>a</sup> Shift of the wavelength of maximum emission of CML7 + ANS mixture following excitation at 380 nm as compared to that of ANS alone.

<sup>b</sup> Ratio between maximal fluorescence emission of CML7 + ANS mixture and that of ANS alone.

<sup>c</sup> Melting temperature obtained by data fitting in thermal denaturation profiles.

S2).

This data suggests that the binding of  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$  strongly stabilizes CML7 by leading to a rearrangement in both the secondary and the tertiary structures. However, for each metal the stabilized final conformation is distinct. Using the hydrophobic fluorescent probe ANS, we demonstrated that the binding of  $\text{Ca}^{2+}$  induces a domain opening in CML7 leading to the exposure of hydrophobic regions, whereas  $\text{Mg}^{2+}$  does not. As shown in Fig. 3B and Table 1, upon addition of  $\text{Ca}^{2+}$  to CML7 a blue-shift in the emission peak of ANS fluorescence (from 525 nm to 504 nm) was observed along with a  $\sim 2$ -fold increase in the maximal emission intensity, indicating that the hydrophobic regions of the recombinant CML7 were exposed to solvent, and therefore accessible to the ANS probe. On the contrary, the addition of  $\text{Mg}^{2+}$  to CML7 did not cause a significant change in the ANS fluorescence spectrum compared to that observed in the apo form. However, few variations in the emission spectra were observed for either apo-CML7 or  $\text{Mg}^{2+}$ -bound CML7 compared to the spectrum of ANS alone, suggesting that some hydrophobic exposure is present even in the absence of  $\text{Ca}^{2+}$  (Fig. 3B and Table 1). Thus, we showed that, in accordance with a putative function as a  $\text{Ca}^{2+}$  sensor, CML7 undergoes a large structural rearrangement leading to hydrophobic exposure only upon binding  $\text{Ca}^{2+}$ .

### 3.3. The C-lobe contains the higher affinity $\text{Ca}^{2+}$ binding sites

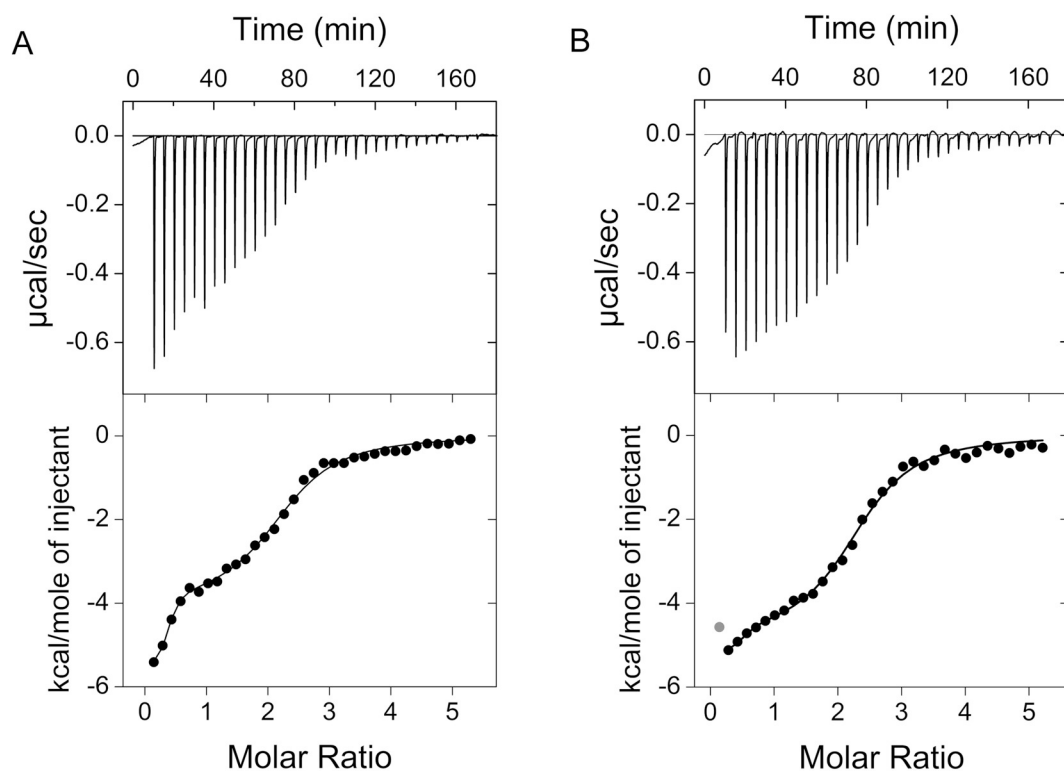
We assessed  $\text{Ca}^{2+}$  binding to CML7 by ITC, starting from conditions of decalcified protein and buffer (Fig. 4). Data were fit using a sequential binding model with the stoichiometry of  $\text{Ca}^{2+}$  binding fixed to four (as deduced by NMR experiments) and following a procedure used previously for mammalian CaM and similar EF-hand containing  $\text{Ca}^{2+}$ -binding proteins [14,15,33,43,49–52]. The values obtained for the four macroscopic binding constants showed that the  $\text{Ca}^{2+}$  binding sites in

CML7 are grouped into two pairs of different strength (Table 2): a first set of sites that comprises two high-affinity sites ( $K_{d1} = 0.9 \pm 0.2 \mu\text{M}$ ,  $K_{d2} = 0.5 \pm 0.1 \mu\text{M}$ , respectively) and a second set that likely represents the binding of two  $\text{Ca}^{2+}$  ions with lower affinity ( $K_{d3} = 14.7 \pm 0.8 \mu\text{M}$ ,  $K_{d4} = 29.4 \pm 4.7 \mu\text{M}$ ).  $\text{Ca}^{2+}$  binding to CML7 followed an exothermic reaction, displaying a negative enthalpy change, which can be ascribed almost entirely to  $\text{Ca}^{2+}$ -induced conformational changes in the N- and C-terminal domains of CML7, since the  $\Delta H$  associated with  $\text{Ca}^{2+}$  binding/desolvation can be approximated to be zero [43,49].

To learn more about the pathways of  $\text{Ca}^{2+}$  binding to CML7,  $\text{Ca}^{2+}$  was titrated into a solution of apo-CML7 and two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded at each step. As shown in Fig. 5A, the Gly<sub>6</sub> peaks for the C-terminal lobe (Gly93 and Gly131) appeared simultaneously at  $\text{Ca}^{2+}$ : protein ratios lower than those of the N-terminal Gly<sub>6</sub> (Gly19 and Gly55) which develop together only at higher  $\text{Ca}^{2+}$ : protein ratios. These observations indicate that the high affinity sites are located in the C-lobe of CML7 and allowed us to assign the dissociation constants determined from ITC  $K_{d1}$  and  $K_{d2}$  to the C-terminal domain (higher affinity  $\text{Ca}^{2+}$ -binding sites), and  $K_{d3}$  and  $K_{d4}$  to the N-terminal EF-hands (lower affinity  $\text{Ca}^{2+}$ -binding sites).

The paired nature of these plant CML7 EF-hands and the fact that two of the sites have higher  $\text{Ca}^{2+}$  affinity than the other two enabled us to estimate the cooperativity within each EF-hand pair. The cooperativity in a two-site system (EF-hand pair) can be quantified in terms of free energy, with the term  $-\Delta\Delta G$  reflecting the increase in affinity of the second EF-hand as a result of  $\text{Ca}^{2+}$  binding to the first [53,54]. If the affinity for the second ion is greater than the first, this term is positive, indicating positive cooperativity. As shown in Supplementary Table S1, each pair of CML7 EF-hands displayed positive cooperativity. Thus, we can conclude that the two sites of the C-terminal domain in CML7 bind  $\text{Ca}^{2+}$  most strongly in a cooperative manner.

We next investigated the influence of  $\text{Mg}^{2+}$  on the  $\text{Ca}^{2+}$ -binding thermodynamic parameters as well as association constants of CML7 by ITC.  $\text{Ca}^{2+}$  titrations were performed upon incubation of CML7 with 10 mM  $\text{MgCl}_2$  (Fig. 4B, Table 2, Supplementary Table S1). Although the four EF-hand  $\text{Ca}^{2+}$ -binding sites still bind  $\text{Ca}^{2+}$  with positive cooperativity, the presence of 10 mM  $\text{Mg}^{2+}$  slightly decreased the  $\text{Ca}^{2+}$  association constants of sites 1 and 2 ( $\sim$ up to 6-fold), while it did not significantly modify the  $\text{Ca}^{2+}$  association constants of the lower affinity sites (sites 3 and 4) of CML7. In the presence of  $\text{Mg}^{2+}$ , the  $\text{Ca}^{2+}$ -binding preference of the EF-hands of CML7 remains unchanged, and the C-terminal sites are occupied before the N-terminal sites, as observed in NMR titration experiments (data not shown). Importantly,  $\text{Ca}^{2+}$  addition to  $\text{Mg}^{2+}$ -bound CML7 leads to the disappearance of the peaks associated with  $\text{Mg}^{2+}$  bound form in  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum and the



**Fig. 4.** ITC analysis of the  $\text{Ca}^{2+}$  binding to apo- and  $\text{Mg}^{2+}$ -bound CML7. Representative raw trace (top) of the calorimetric titration of  $\text{CaCl}_2$  into 200  $\mu\text{M}$  apo-CML7 (A) and 200  $\mu\text{M}$  CML7 pre-incubated with 10 mM  $\text{MgCl}_2$  (B), and integrated binding isotherm (bottom). Curve fitting was performed using a four-site sequential binding model. The ligand dilution blank experiments ( $\text{Ca}^{2+}$  ions titrated into buffer only) were subtracted from the binding isotherm obtained in the presence of protein.

**Table 2**

Thermodynamic parameters for  $\text{Ca}^{2+}$  binding to CML7. The mean values from triplicate experiments and SEM using at least two different CML7 preparations are presented.

	No $\text{Mg}^{2+}$			+ 10 mM $\text{Mg}^{2+}$		
	$K_a$ ( $\text{M}^{-1}$ )	$K_d$ ( $\mu\text{M}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )	$K_a$ ( $\text{M}^{-1}$ )	$K_d$ ( $\mu\text{M}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )
Site 1	$1.1\text{E}6 \pm 2.9\text{E}5$	$0.9 \pm 0.2$	$-6.6\text{E}3 \pm 0.8\text{E}3$	$2.6\text{E}5 \pm 6.3\text{E}4$	$3.8 \pm 0.9$	$-5.9\text{E}3 \pm 0.9\text{E}3$
Site 2	$2.2\text{E}6 \pm 4.1\text{E}5$	$0.5 \pm 0.1$	$-1.6\text{E}3 \pm 0.4\text{E}3$	$3.2\text{E}5 \pm 6.2\text{E}4$	$3.1 \pm 0.6$	$-1.9\text{E}3 \pm 0.3\text{E}3$
Site 3	$6.8\text{E}4 \pm 3.9\text{E}3$	$14.7 \pm 0.8$	$-3.1\text{E}3 \pm 0.7\text{E}3$	$4.9\text{E}4 \pm 7.3\text{E}3$	$20.4 \pm 3.0$	$-2.6\text{E}3 \pm 0.6\text{E}3$
Site 4	$3.4\text{E}4 \pm 5.4\text{E}3$	$29.4 \pm 4.7$	$-1.5\text{E}3 \pm 0.4\text{E}3$	$2.6\text{E}4 \pm 8.4\text{E}3$	$38.5 \pm 6.9$	$-0.8\text{E}3 \pm 0.1\text{E}3$

Sites 1 and 2 can be assigned to the C-terminal lobe and sites 3 and 4 to the N-terminal lobe. The domain-specific assignment is based on NMR analysis (see text for explanation).

concomitant appearance of the peaks characteristic of the  $\text{Ca}^{2+}$ -bound CML7; the resulting spectrum is nearly superimposable with that of  $\text{Ca}^{2+}$ -CML7 (Figs. 1C, 5). This provides evidence of  $\text{Mg}^{2+}$  displacement by  $\text{Ca}^{2+}$ , and that CML7 attains the same conformation after binding  $\text{Ca}^{2+}$ , in both the presence and absence of  $\text{Mg}^{2+}$ .

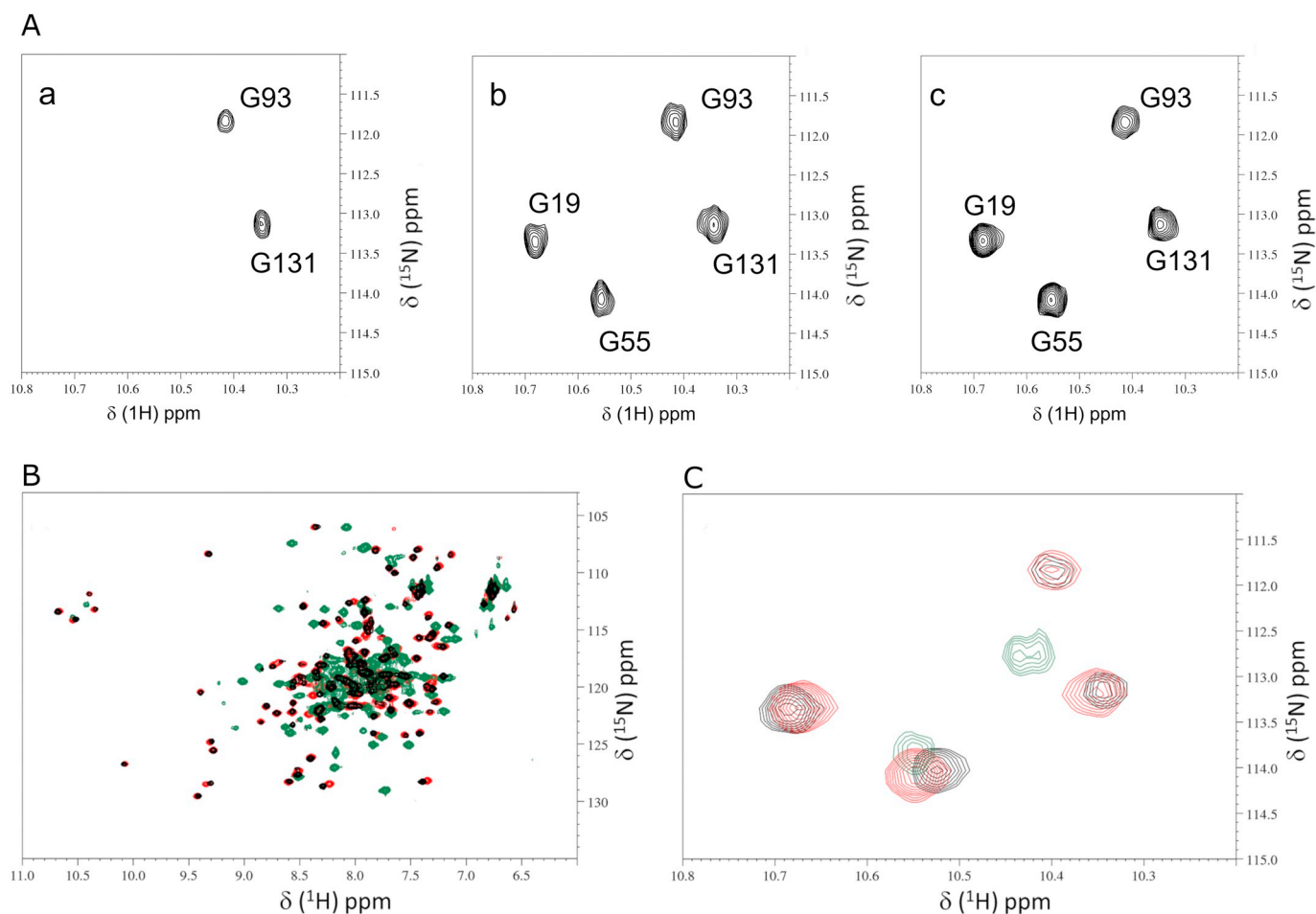
### 3.4. Melittin binds to CML7

Melittin (GIGAVLKVLTGTPALISWIKRKRQ) is an amphipathic model peptide that is widely used to study the interaction of downstream targets with proteins of the CaM family. It is well-known that melittin interacts with  $\text{Ca}^{2+}$ -saturated bovine CaM to form a 1:1 complex with nanomolar affinity [55], assuming an  $\alpha$ -helix conformation upon binding [56]. In an effort to demonstrate that CML7 behaves as a member of CaM family towards the interaction of targets and since the CML7 physiological target in plants has not been identified, we examined CML7 for interaction with melittin. A canonical plant CaM (AtCaM1 from *A. thaliana*) was used as a positive control in our

experiments.

First, we monitored the interaction of CML7 or AtCaM1 with melittin in the presence of increasing amounts of the peptide by native PAGE (Fig. 6A). Nondenaturing gel band shift electrophoresis directly confirmed that melittin forms a 1:1 complex with AtCaM1 in the presence of  $\text{Ca}^{2+}$ , but not in its absence (i.e. in the presence of 5 mM EGTA (Fig. 6A)). The peptide also forms a complex with CML7 in the presence of  $\text{Ca}^{2+}$ . Upon addition of melittin, the single band reflecting pure CML7 disappeared and a new protein band of lower mobility appeared, indicating the formation of a new complex. At 1:1 protein:peptide ratio, almost all the free CML7 disappeared, suggesting the 1:1 formation of a peptide/CML7 complex. At higher concentrations of melittin, a smearing of the band representing the complex was observed, likely indicating that additional binding occurs (up to 4 mol of melittin/mol of CML7).

Formation of CML7/melittin and AtCaM1/melittin complexes was also studied by fluorescence spectroscopy. Both CML7 and AtCaM1 do not have any Trp residues, while melittin possesses one Trp (Trp19).



**Fig. 5.** NMR analysis of the  $\text{Ca}^{2+}$  binding to apo- and  $\text{Mg}^{2+}$ -bound CML7. (A) Downfield region of selected  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of CML7 recorded as a function of increasing  $\text{Ca}^{2+}$  concentration. The molar ratio of  $\text{Ca}^{2+}$ : protein was 2 (a), 4.5 (b), 10 (c). (B) Superimposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra recorded on  $^{15}\text{N}$ -CML7 after addition of 5 mM  $\text{CaCl}_2$  (red), 20 mM  $\text{MgCl}_2$  (green), or 20 mM  $\text{MgCl}_2$  + 20 mM  $\text{CaCl}_2$  (black). (C) Zoom of the Gly downfield peaks. The spectra were recorded at 600 MHz and 298 K. All samples were at protein concentration of 270  $\mu\text{M}$  in 50 mM Tris, 50 mM KCl, 0.5 mM DTT, 7%  $\text{D}_2\text{O}$  pH 7.5.

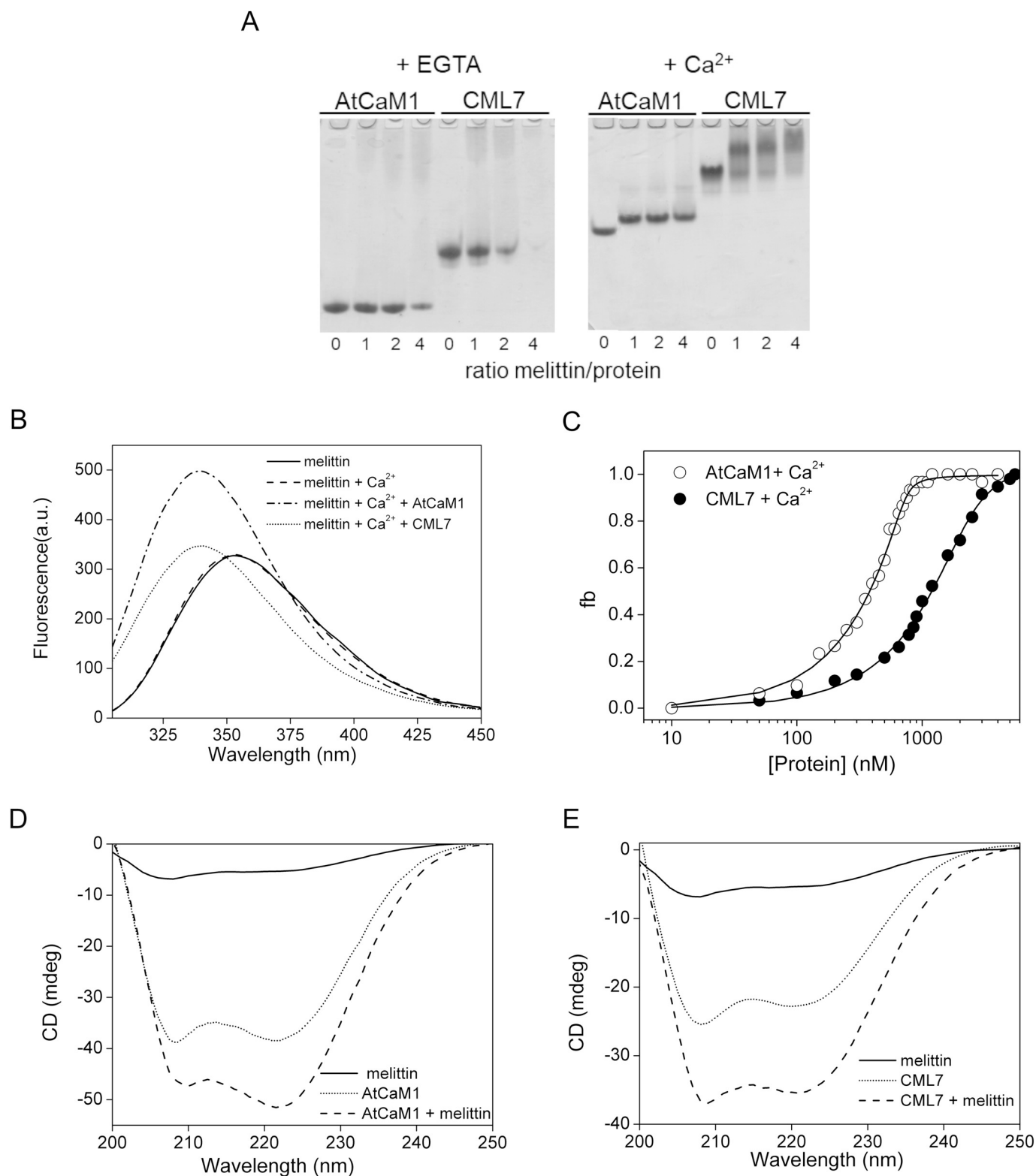
This intrinsic fluorophore is a useful probe to monitor binding of the peptide to the protein since, upon complex formation it usually experiences changes in its spectral properties (maximum emission wavelength and/or fluorescence intensity). As expected for a solvent exposed Trp residue, for melittin alone, the maximum fluorescence emission wavelength is 354 nm. Upon the addition of  $\text{CaCl}_2$  to the melittin solution, no changes in the fluorescence spectrum were observed, while the fluorescence was blue shifted to 339 nm and 338 nm upon binding to AtCaM1 and CML7, respectively (in the presence of saturating concentrations of  $\text{Ca}^{2+}$ ) (Fig. 6B). This large blue shift is indicative of an augmented hydrophobicity of the Trp environment, suggesting that interaction between proteins and melittin occurs. The weaker intensity of the Trp fluorescence emission when melittin is bound to CML7 compared to AtCaM1 may be due to increased quenching either by the surrounding amino acid side chains or from a greater exposure to solvent. The microenvironment of the Trp residue of melittin could be perturbed in a different way in the two complexes. However, it might also suggest that the Trp is less rigidly held in the binding pocket, therefore we cannot exclude the presence of conformational changes which could potentially alter the fluorescence properties of melittin in a different manner in the two complexes.

To estimate the apparent affinity between proteins and melittin, titration experiments in Trp fluorescence were performed (Fig. 6C). Our data showed that, in the presence of  $\text{Ca}^{2+}$ , AtCaM1 binds melittin with high affinity ( $K_d$  of  $8 \pm 2$  nM) and with 1:1 stoichiometry, in line with previously reported data for bovine CaM/melittin interaction [55].

Binding of melittin to  $\text{Ca}^{2+}$ -bound CML7 also occurs with high affinity ( $K_d$  of  $0.3 \pm 0.1$   $\mu\text{M}$ ), even if lower compared to AtCaM1.

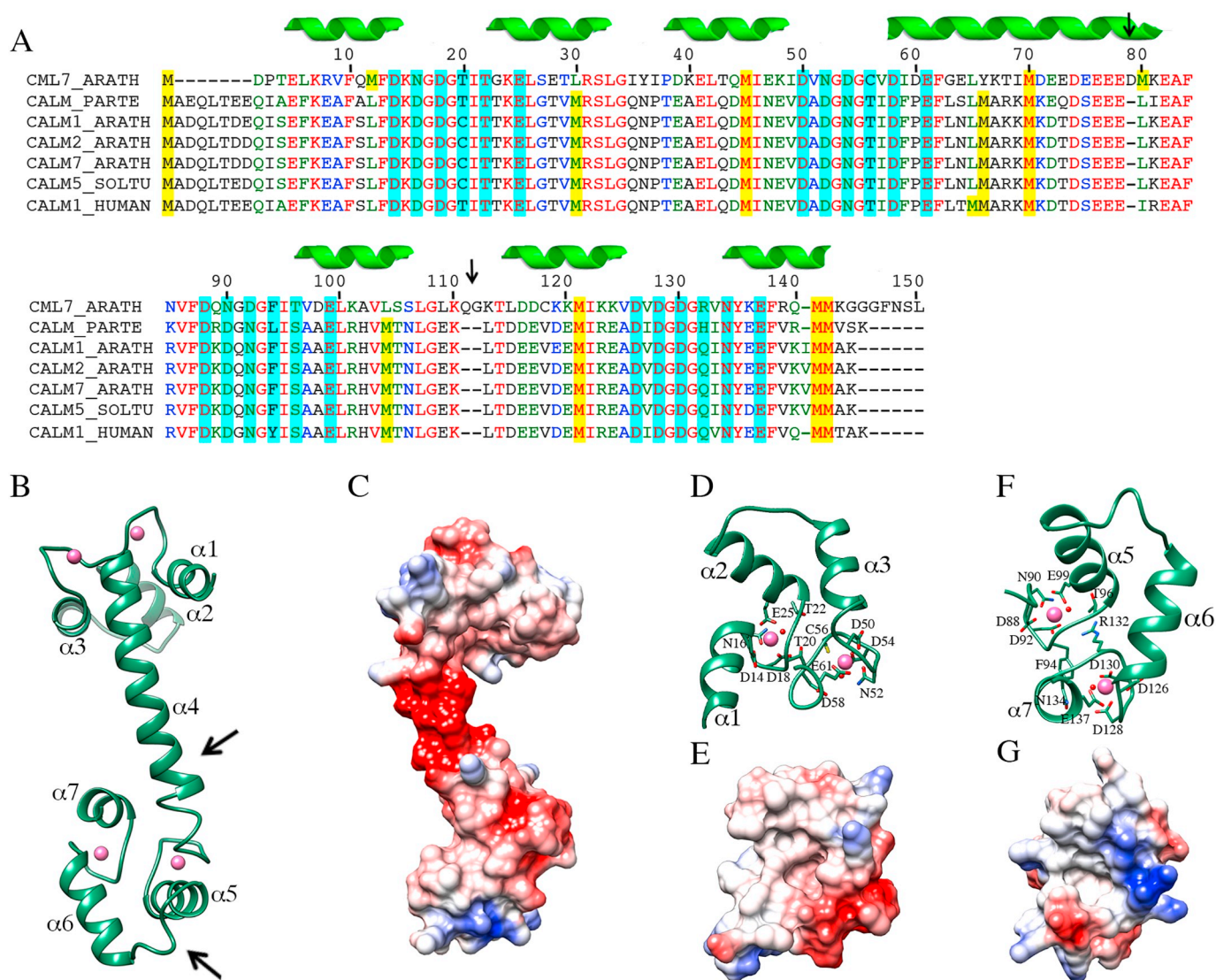
Furthermore, we acquired CD spectra to examine whether binding of melittin to AtCaM1 or CML7 is associated with changes in secondary structure. As shown in Fig. 6D and 6E, CD signals of both AtCaM1 and CML7 were increased following the addition of melittin. In aqueous solution, free melittin has a small  $\alpha$ -helical content. The observation of  $\alpha$ -helix formation is not uncommon in monomeric peptides that are 20–25 residues in length [57,58]. The far UV CD spectrum of AtCaM1 and CML7 showed two minima peaks at 208 and 222 nm, which is typical of a protein with a significant  $\alpha$ -helical content, as described above [33,59]. Addition of equimolar melittin to AtCaM1 or CML7 yielded a spectrum that was more negative than the sum of each taken separately, suggesting that the  $\alpha$ -helical content of the  $\text{Ca}^{2+}$ -protein/peptide complex increases (Fig. 6D, E). The higher value of CD signal observed is attributable to induced helicity in melittin, since  $\text{Ca}^{2+}$ -CaM itself does not gain secondary structure following binding to target peptide [60,61]. Moreover, in the case of AtCaM1, the shape of the spectrum representing the complex was also different from the individual components. In the presence of melittin and  $\text{Ca}^{2+}$ , the  $\theta_{222}/\theta_{208}$  ratio for AtCaM1 approaches the value of  $\sim 1.08$  observed for coiled-coil motifs. Thus, it is possible that a tight interaction between  $\alpha$ -helices, which in certain regions may coil together, occurs when AtCaM1 binds melittin.

Importantly, in the presence of EGTA, CD spectra were characterized by only a modest change upon addition of melittin to CML7



**Fig. 6.** Interaction of melittin with AtCaM1 and CML7. (A) Native PAGE in the presence of  $\text{CaCl}_2$  or EGTA. (B) Complex formation monitored by Trp fluorescence. Emission spectra, recorded in the presence of  $\text{CaCl}_2$ , of melittin (dashed line), melittin/AtCaM1 (dash dotted line) and melittin/CML7 complex (dotted line). The spectrum of melittin before addition of  $\text{CaCl}_2$  (solid line) is also shown. (C) Fluorescence titration of AtCaM1 (open circles) or CML7 (filled circles) with melittin in the presence of  $\text{CaCl}_2$ . The  $K_d$  value is determined by fitting the fraction bound (fb) to Eq. (1). (D–E) Far UV CD spectra changes in secondary structure for AtCaM1/melittin (D) and CML7/melittin complex (E) in the presence of  $\text{CaCl}_2$ . Melittin alone (solid line), AtCaM1 or CML7 alone (dotted line), and protein/peptide complex at 1:1 molar ratio (dashed line).





**Fig. 7.** Sequence alignment and molecular modelling of CML7. (A) Sequence alignment of CML7\_ARATH (CML7 from *A. thaliana*; Uniprot: Q9LNE7), CALM\_PARTE (CaM from *P. tetraurelia*; Uniprot: P07463), CALM1\_ARATH (CaM1 from *A. thaliana*; Uniprot: PODH95), CALM2\_ARATH (CaM2 from *A. thaliana*; Uniprot: PODH97), CALM7\_ARATH (CaM7 from *A. thaliana*; Uniprot: P59220), CALM5\_SOLTU (CaM5 from *S. tuberosum*; Uniprot: Q7DMN9), CALM1\_HUMAN (CaM1 from *H. sapiens*; Uniprot: PODP23); residues involved in  $\text{Ca}^{2+}$  coordination are highlighted in light blue, Met in yellow; black arrows represent amino acid insertion in the CML7 sequence. Numbering and secondary structure elements assignments are highlighted for CML7. (B) Cartoon representation of the CML7 molecular modelling;  $\text{Ca}^{2+}$  is represented as a pink sphere; arrows represent region of amino acid insertion in the CML7 sequence. (C) Coulombing surface coloring of the CML7 molecular surface. (D–F) Close view of the N-lobe and the C-lobe respectively. Residues involved in  $\text{Ca}^{2+}$  coordination are represented in stick, water molecules (red spheres) were transferred from the template (E–G) Coulombing surface coloring of the CML7 N-lobe and the C-lobe.

(Supplementary Fig. S3), suggesting that the interaction takes place mainly in the presence of  $\text{Ca}^{2+}$ .

Our results showed that CML7 has a high affinity interaction with melittin, a model CaM interacting peptide, in a  $\text{Ca}^{2+}$ -dependent manner. Therefore, we can envisage that CML7 may be involved in  $\text{Ca}^{2+}$  mediated protein-protein interactions even if its physiological partner/s is still unidentified.

### 3.5. CML7 sequence analysis and homology model

A homology model of the CML7 from *A. thaliana* in the  $\text{Ca}^{2+}$  bound form was generated using the structure of the CaM from *Paramecium tetraurelia* (PDB entry: 1clm [41]) as template. The model assumes the dumbbell-like structure typical of CaM proteins formed by two structurally similar globular domains (the N-term lobe and the C-term lobe), containing a pair of  $\text{Ca}^{2+}$ -binding motifs connected by a central  $\alpha$ -helix (Fig. 7B, C). It was recently reported that while the conformation of

animal and *P. tetraurelia* CaM proteins are very similar, CaM6 from potato [62] and CaM7 from *A. thaliana* [63] show a different orientation of the central helix. The central linker connecting the two globular domains has been described as “magical” because it is highly flexible, allowing CaM to interact with a wide range of targets in a remarkable variety of conformations [64–66]. For this reason, while the CML7 model can be considered highly accurate in the  $\text{Ca}^{2+}$  binding domains, the region corresponding to the central helix may be less accurate. Moreover, CML7 has an amino acid insertion (Asp79) in the C-terminal end of the central helix (Fig. 7A, B) that may lead the linker to adopt different orientations compared to typical CaMs.

The  $\text{Ca}^{2+}$  binding site is composed of 12 residues in which positions 1, 3, 5, 7, 9 and 12 are occupied by residues that, together with a water molecule, contribute to  $\text{Ca}^{2+}$  coordination. Among them, residues in positions 1, 3, 5, and 12 participate in  $\text{Ca}^{2+}$  binding through side chain oxygen, residue in position 7 with main chain oxygen while residue in position 9 coordinates a water molecule (Fig. 7A, D, F) [26]. Analysis of

the primary structure of CML7 showed that all the four EF-hands in CML7 resemble those found in CaM proteins, especially in the positions devoted to  $\text{Ca}^{2+}$  binding (Fig. 7A), with minor differences. In the EF-hand 1, position 3, which is usually occupied by an Asp residue, is replaced by an Asn in CML7 (Fig. 7A). This conservative substitution is also present in other CMLs from *A. thaliana* belonging to the VII subgroup in the classification performed by Zhu [24], with the only exception of CML6 (Supplementary Fig. S4). EF-hand 2 of CML7 shows an Asp to Asn substitution in position 3, while position 5 (in which an Asn is usually present) is occupied by an Asp residue. Interestingly, EF-hand 2 is less conserved in the CMLs belonging to subgroup VII and appears to be not functional in CML42, CML43, and CML44 (Supplementary Fig. S4) [9,10,67]. Similarly to EF-hand 2 of CML7, also in EF-hand 3 Asp to Asn substitutions in positions 3 and 5 are present. The residues responsible for  $\text{Ca}^{2+}$  coordination are fully conserved in EF-hand 4. Interestingly, a two-residue insertion is present in the loop connecting EF-hand 3 and EF-hand 4 and it seems to be a general feature in the CML proteins belonging to subgroup VII (Supplementary Fig. S4).

In analogy to what reported for CaMs, it is likely that the binding of  $\text{Ca}^{2+}$  to each globular domain in CML7 changes the inter-helical angles in the EF-hand motifs, inducing exposure of hydrophobic sites that favors the interactions with downstream partners. As shown in Figs. 7E and G, the Coulombic surface of both N-terminal and C-terminal lobes of the model of CML7 in the  $\text{Ca}^{2+}$ -bound form revealed the presence of hydrophobic patches in which Met residues likely play a key role in the interaction with targets [68]. CML7 primary sequence contains eight Met residues; two of them are located in the N-lobe (Met12 and Met45), two in the central region (Met70 and Met80), and three in the C-lobe (Met121, Met141 and Met142). Only five of them are conserved in other CaMs (Fig. 7A). In addition to the exposure of hydrophobic residues upon  $\text{Ca}^{2+}$  binding, the ability of CaM to recognize different interacting molecules is assured by the presence of the helix [68,69]. In CML7, this helix appears to contain a negatively charged region that is longer than that present in CaM proteins (Fig. 7A, C). This region is quite variable among CMLs belonging to subgroup VII (Supplementary Fig. S4), and this diversity may contribute to different modes of target recognition and/or binding topology that distinguish CMLs from one another.

### 3.6. CML7 is localized in the cytoplasm

Protein localization is a necessary step if we aim to understand the function of CML7 in a physiological context and examine functional redundancy between CaM and CML isoforms. To determine the subcellular localization of CML7 in plant cells, we fused the RFP at the C-terminus end of CML7 and placed it under the control of the CaMV35S promoter. *A. thaliana* protoplasts cotransformed with the CML7-RFP fusion and the free EGFP showed green fluorescence with typical cytoplasmic localization (Fig. 8A). The red fluorescence from the CML7-RFP fusion was also restricted to cytoplasm (Fig. 8B), perfectly

overlapping with the green fluorescence (Fig. 8E).

## 4. Discussion

$\text{Ca}^{2+}$  has a central role in adaptive developmental and stress response programs in plants, regulating many important physiological processes. A number of abiotic and biotic stimuli elicit specific  $\text{Ca}^{2+}$  signals that usually occur via changes in cytosolic free  $\text{Ca}^{2+}$  concentration, and the shape of these  $\text{Ca}^{2+}$  signals is believed to specify the nature of the response [21,22]. The mechanisms by which changes in intracellular  $\text{Ca}^{2+}$  concentration can lead to diverse responses is a topic of widespread interest among plant biologists that has recently focused on the roles of CMLs as specialized  $\text{Ca}^{2+}$  sensors [20–22,70,71]. Importantly, the plant CML protein family is revealing its diversification among  $\text{Ca}^{2+}$ -binding proteins with respect to CaM [12,20–22,70–73].

There are 50 isoforms of CML in *A. thaliana*, and characterization of the different isoforms is essential to understand their functional significance. To this end, we have recently described the ion and target binding properties of several *A. thaliana* CMLs including CML14, CML19, and CML36 [8,14,15]. This investigation reports on the  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  sensing properties of CML7, as well as on metal induced conformational changes and binding to the model target peptide melittin.

Our results provide significant information regarding modulation of the structure of CML7 in the presence of saturating concentrations of  $\text{Ca}^{2+}$ . In the absence of  $\text{Ca}^{2+}$  ions, the CD spectrum shows extensive helical content, while ANS fluorescence suggests that the protein has a closed conformation and no or only limited exposure of hydrophobic surfaces. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum is characterized by broad peaks, indicating conformational heterogeneity in the absence of  $\text{Ca}^{2+}$ . Under saturating  $\text{Ca}^{2+}$  conditions, transition between the heterogeneous closed conformation and the open conformation took place.  $\text{Ca}^{2+}$ -dependent exposure of hydrophobic regions as measured by ANS fluorescence,  $\text{Ca}^{2+}$ -dependent interaction of CML7 with hydrophobic phenyl-Sepharose, and increased mobility in electrophoresis in the presence of  $\text{Ca}^{2+}$ , collectively corroborated the significant  $\text{Ca}^{2+}$ -induced conformational change in CML7.

Moreover, molecular modelling of CML7 was performed in the  $\text{Ca}^{2+}$ -bound form. Our model suggests that the overall structure in CML7 is well conserved and assumes the dumbbell-like structure typical of CaM proteins, in which the N-lobe and C-lobe are devoted to  $\text{Ca}^{2+}$  binding through conserved or conservatively mutated residues. CML7 possesses four EF-hand motifs and the ability of all of them to bind  $\text{Ca}^{2+}$  was confirmed by the appearance in the NMR spectrum of  $\text{Ca}^{2+}$ -bound CML7 of four downfield-shifted resonances, which belong to the conserved Gly residues at position 6 of  $\text{Ca}^{2+}$ -occupied EF-hands. This stoichiometry was further supported by ITC analysis and by both amino acid sequence predictions and inspection of the CML7 model generated *in silico*. According to the ITC and NMR results, CML7 was shown to possess two high affinity  $\text{Ca}^{2+}$  sites in the C-lobe and two weaker sites

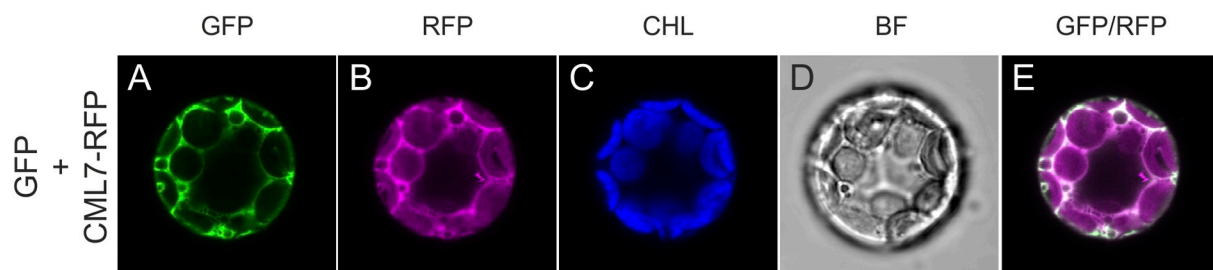


Fig. 8. Subcellular localization of CML7-RFP in *A. thaliana* mesophyll protoplasts. The CML7-RFP fusion protein was transiently coexpressed in protoplasts with free EGFP. Fluorescence was analyzed by confocal microscopy. (A) EGFP fluorescence (green) from a single focal plane of a representative protoplast. (B) RFP fluorescence (magenta) from the same protoplast as shown in (A). (C) Chlorophyll autofluorescence of chloroplasts. (D) Bright field image. (E) Overlay image of A and B.

in the N-lobe. This high/moderate  $\text{Ca}^{2+}$  affinity appears to be an emerging pattern in CMLs [12–14], and a wide range of  $\text{Ca}^{2+}$  affinity (nM- $\mu\text{M}$  range) has been determined for this class of  $\text{Ca}^{2+}$  sensors [9,10,13,15]. The growing idea is that CMLs might possess different  $\text{Ca}^{2+}$  binding affinities to specifically sense a broad range of cytosolic free  $\text{Ca}^{2+}$  variations, being thus key players in  $\text{Ca}^{2+}$  signaling transduction. Along with this, the apparent affinity for  $\text{Ca}^{2+}$  for most CMLs differs considerably from the universal  $\text{Ca}^{2+}$  sensor CaM [9,10,13,15,33,43]. Importantly, positive cooperativity of  $\text{Ca}^{2+}$  binding is observed within each globular domain of CML7, in line with the paired nature of these plant CML7 EF-hands, and as frequently observed in EF-hand containing proteins [26,43,54,74].

We also studied the effect of  $\text{Mg}^{2+}$  on the binding of  $\text{Ca}^{2+}$  and the conformation of CML7. The interaction between  $\text{Mg}^{2+}$  and CML7 is crucial for understanding how CML7 specifically responds to micromolar variations in free cytosolic  $\text{Ca}^{2+}$  concentrations in the presence of  $\sim 1000$ -fold excess of the chemically similar  $\text{Mg}^{2+}$  ions.  $\text{Ca}^{2+}$  affinities of CML7 decreased by  $\sim 4$ – $6$ -fold for C-terminal sites and  $\sim 1.3$ -fold for N-terminal sites when assessed in the presence of  $\text{Mg}^{2+}$ . Importantly, the order of  $\text{Ca}^{2+}$  binding to CML7 remained the same as that in the absence of  $\text{Mg}^{2+}$ , and similar positive cooperativity was also observed. The modest effect of  $\text{Mg}^{2+}$  on  $\text{Ca}^{2+}$  affinity of CML7, which is similar to the  $\text{Mg}^{2+}$ -based change in  $\text{Ca}^{2+}$  affinity seen for mammalian CaM [49], likely allows the EF-hands of CML7 to be classified as  $\text{Ca}^{2+}$ -specific sites [26]. The effect of  $\text{Mg}^{2+}$  on  $\text{Ca}^{2+}$  binding is still controversial and competition or allosteric effects of  $\text{Mg}^{2+}$  with respect to  $\text{Ca}^{2+}$  binding have been suggested [26,49,75,76]. Moreover, the presence of auxiliary  $\text{Mg}^{2+}$  binding sites distinct from the EF-hand  $\text{Ca}^{2+}$  binding sites in CaM has been proposed [44]. NMR spectra of  $\text{Mg}^{2+}$ -CML7 showed two downfield shifted  $\text{Gly}_6$  resonances, pointing to the  $\text{Mg}^{2+}$  binding sites being located within the EF-loop structure. However, only the comparison of the crystal structure of  $\text{Mg}^{2+}$ -bound CML7 with the  $\text{Ca}^{2+}$ -bound CML7 could provide more insight into this aspect.

Owing to the excess  $\text{Mg}^{2+}$  concentration in the resting cell, many  $\text{Ca}^{2+}$ -specific EF-hands have enough affinity for this cation to be partially or even fully saturated with  $\text{Mg}^{2+}$  [26,77]. The crucial factor for efficient  $\text{Ca}^{2+}$  signaling is the difference in the resulting metal-bound protein conformation. Binding of  $\text{Mg}^{2+}$  to apo-CML7 induced changes in the secondary structure similar to those induced by  $\text{Ca}^{2+}$ . Moreover, both  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  stabilize CML7, as reflected in a large increase in the unfolding temperature. However, the  $\text{Mg}^{2+}$ -bound conformation of CML7 differs from the  $\text{Ca}^{2+}$ -bound form of the protein, as observed in NMR spectra, and CML7 does not experience a marked increase in surface hydrophobicity in the presence of  $\text{Mg}^{2+}$ , as demonstrated by the ANS data. This suggests that the  $\text{Mg}^{2+}$ -bound protein and the  $\text{Mg}^{2+}$ -bound loops likely remain in a “closed” conformation, analogously to that seen with  $\text{Mg}^{2+}$ -bound mammalian CaM [46], and therefore  $\text{Mg}^{2+}$  is not predicted to replace  $\text{Ca}^{2+}$  in hydrophobic-dependent target interaction. Consistent with this, the bidimensional NMR spectrum of  $\text{Ca}^{2+}$ -saturated CML7 in the presence of  $\text{Mg}^{2+}$  is essentially identical to that of the  $\text{Mg}^{2+}$ -free,  $\text{Ca}^{2+}$ -bound protein, indicating that  $\text{Ca}^{2+}$ -saturated CML7 has the same three-dimensional conformation both in the absence and in the presence of excess  $\text{Mg}^{2+}$ .

Overall, the results obtained indicate that CML7 belongs to the group of classical  $\text{Ca}^{2+}$  sensors instead of  $\text{Ca}^{2+}$ -buffering proteins, which experience only minor structural variations and bind very tightly to  $\text{Ca}^{2+}$  [26]. Working as a  $\text{Ca}^{2+}$  sensor, CML7 is therefore expected to interact with and/or regulate specific target proteins. Inspection of the coulombic surface of CML7 in our *in silico* model indicates that the inner surfaces of the N- and C-term lobes contain Met-rich hydrophobic patches, while acidic clusters are present in the linker region. The presence of these hydrophobic and acidic regions is likely instrumental for interaction with target proteins and suggests the necessity for preserving the contact sites with the effector molecule with respect to CaM. CaM-binding sites are usually positively charged, amphiphilic  $\alpha$ -helical

peptides with bulky hydrophobic anchor residues (e.g. Trp) that are often located at one or both ends of the helix [78]. Thus, CaM binding to its targets usually involves both hydrophobic and electrostatic interactions [60,61]. Melittin, a small peptide of 26 residues, is widely used to study protein–peptide interactions in  $\text{Ca}^{2+}$ -binding proteins, including CaM, and can provide insight into target recognition mode of CaM [55,79,80]. Earlier studies of the  $\text{Ca}^{2+}$ -CaM/melittin complex have shown that the binding of melittin to bovine CaM occurs with high affinity ( $K_d$  in nanomolar range) and in a parallel orientation, meaning that the N- and C-terminal regions of the peptide are associated with the N- and C-term domains of the protein, respectively [56,79]. Herein, by non-denaturing-PAGE, CD spectroscopy, and steady state Trp fluorescence we showed that plant CaMs (i.e., CaM1 from *A. thaliana*) are also able to bind melittin peptide with a 1:1 stoichiometry and very high affinity, in agreement with the previously reported bovine  $\text{Ca}^{2+}$ -CaM/melittin interaction [55]. Importantly, we further demonstrated that CML7 binds the natural peptide melittin in the presence of  $\text{Ca}^{2+}$ . According to our results, the melittin/CML7 interaction has a 1:1 stoichiometry with a  $K_d$  of  $0.3 \pm 0.1 \mu\text{M}$ . Binding of melittin to  $\text{Ca}^{2+}$ -AtCaM1 and  $\text{Ca}^{2+}$ -CML7 occurred in an  $\alpha$ -helical conformation, and the interaction involved the Trp residue of the peptide which becomes solvent shielded.

The binding of melittin definitely places CML7 in the group of  $\text{Ca}^{2+}$  sensors, but more insightful information will require identification and examination of the formation of complexes between CML7 and peptides that arise from its physiological targets. Moreover, although important clues can be obtained by the analysis of CML complexes with short target peptides, these complexes do not always perfectly reflect the biophysical and structural properties of the full-length protein. Therefore, in the absence of physiologically significant functional information, the structural and biochemical data should be interpreted with caution.

An involvement of CML7 in root development has been proposed [27], even if the downstream targets of CML7 have not yet been identified. Despite the large number of molecular interactions that take place with members of the CaM superfamily, specific differences in the structure of physiological targets are likely to ensure specificity of interactions. Indeed, there may be large differences in target binding among CMLs, and ultimately each CML/target interaction must be verified separately. However, it is worth noting that the structural properties and the mechanism of action of CML7 and AtCaM1, as predicted from our results, seem to be very similar. Thus, it is unclear how they can have non-overlapping physiological functions. Clearly, for  $\text{Ca}^{2+}$  sensors it is crucial to consider the  $\text{Ca}^{2+}$  binding properties in connection with the target. Moreover, factors other than  $\text{Ca}^{2+}$  binding, such as relative abundance of  $\text{Ca}^{2+}$  sensor and target, tissue expression, and cellular localization as well as phosphorylation may have significant implications for target regulation mechanism. Subcellular localization is an important attribute of protein function, since, usually,  $\text{Ca}^{2+}$  sensors and their interacting targets colocalize. CML7 was empirically demonstrated to be cytosolic, therefore providing further clues into understanding of the  $\text{Ca}^{2+}$ /CML7 signal transduction pathway. However, only the identification of the physiological targets of CML7 will allow us to obtain deeper insights into the cellular events that this protein participates in and to determine its selectivity and mechanism of action.

## Acknowledgments

We thank Prof. A. Costa (University of Milan) for his precious help with confocal analysis, which was carried out at the NOLIMITS imaging facility at the University of Milan. We also thank the Centro Piattaforme Tecnologiche of the University of Verona for providing access to the NMR spectrometer, nanoITC calorimeter, and CD spectropolarimeter. This work has been partially supported by the Italian MIUR-PRIN 2017 grant No. 2017ZBBYNC (to AA, MCB, ADM).

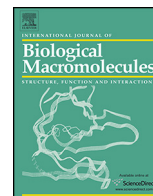


## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jinorgbio.2019.110796>.

## References

- [1] E.A. Permyakov, R.H. Kretsinger, *Calcium Binding Proteins*, John Wiley & Sons, Inc., 2010, pp. 567–573.
- [2] E. McCormack, J. Braam, *New Phytologist* 159 (2003) 585–598, <https://doi.org/10.1046/j.1469-8137.2003.00845.x>.
- [3] B. Boonburapong, T. Buaboocha, *BMC Plant Biology* 7 (2007) 4 BioMed Central, London.
- [4] S. Munir, M.R.G. Khan, J. Song, S. Munir, Y. Zhang, Z. Ye, T. Wang, *Plant Physiol. Biochem.* 102 (2016) 167–179.
- [5] K. Zhang, D. Yue, W. Wei, Y. Hu, J. Feng, Z. Zou, *Frontiers in Plant Science*, vol. 7, Frontiers Media S.A. (2016) 1820.
- [6] E. Vandelle, A. Vannozzi, D. Wong, D. Danzi, A.M. Digby, S. Dal Santo, A. Astegno, *Plant Physiol. Biochem.* 129 (2018) 221–237.
- [7] N. Guo, G. Wang, M. Zong, S. Han, F. Liu, *Gene* 677 (2018) 232–244.
- [8] R. Vallone, V. La Verde, M. D'Onofrio, A. Giorgetti, P. Dominici, A. Astegno, *Protein Sci.* 25 (2016) 1461–1471.
- [9] K.W. Bender, S. Dobney, A. Ogunrinde, D. Chiasson, R.T. Mullen, H.J. Teresinski, P. Singh, K. Munro, S.P. Smith, W.A. Snedden, *Biochem. J.* 457 (2014) 127–136.
- [10] S. Dobney, D. Chiasson, P. Lam, S.P. Smith, W.A. Snedden, *J. Biol. Chem.* 284 (2009) 31647–31657.
- [11] S.S. Scholz, J. Vadassery, M. Heyer, M. Reichelt, K.W. Bender, W.A. Snedden, W. Boland, A. Mithofer, *Mol. Plant* 7 (2014) 1712–1726.
- [12] V. La Verde, P. Dominici, A. Astegno, *Int. J. Mol. Sci.* 19 (2018) 1331.
- [13] A. Ogunrinde, K. Munro, A. Davidson, M. Ubaid, W.A. Snedden, *Front. Plant Sci.* 8 (2017) 2175.
- [14] V. La Verde, M. Trande, M. D'Onofrio, P. Dominici, A. Astegno, *Int. J. Biol. Macromol.* 108 (2018) 1289–1299.
- [15] A. Astegno, M.C. Bonza, R. Vallone, V. La Verde, M. D'Onofrio, L. Luoni, B. Molesini, P. Dominici, *J. Biol. Chem.* 292 (2017) 15049–15061.
- [16] K.M. Cho, H.T. Nguyen, S.Y. Kim, J.S. Shin, D.H. Cho, S.B. Hong, S.H. Ok, *New Phytol.* 209 (2) (2016) 664–678.
- [17] T. Yamaguchi, G.S. Aharon, J.B. Sottosanto, E. Blumwald, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 16107–16112.
- [18] S.C. Popescu, G.V. Popescu, S. Bachan, Z. Zhang, M. Seay, M. Gerstein, M. Snyder, S.P. Dinesh-Kumar, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 4730–4735.
- [19] A. Perochon, S. Dieterle, C. Pouzet, D. Aldon, J.P. Galaud, B. Ranty, *Biochem. Biophys. Res. Commun.* 398 (2010) 747–751.
- [20] D. Aldon, M. Mbengue, C. Mazars, J.-P. Galaud, *Int. J. Mol. Sci.* 19 (2018) 665.
- [21] A. Perochon, D. Aldon, J.P. Galaud, B. Ranty, *Biochimie* 93 (2011) 2048–2053.
- [22] B. Ranty, D. Aldon, V. Cotellet, J.-P. Galaud, P. Thuleau, C. Mazars, *Frontiers in plant science*, vol. 7, Frontiers Media S.A. (2016) 327.
- [23] S. Nie, M. Zhang, L. Zhang, *BMC Genomics* 18 (2017) 842.
- [24] X. Zhu, C. Dunand, W. Snedden, J.-P. Galaud, *Trends Plant Sci.* 20 (2015) 483–489.
- [25] E. de Castro, C.J. Sigrist, A. Gattiker, V. Bulliard, P.S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch, N. Hulo, *Nucleic Acids Res.* 34 (2006) W362–W365.
- [26] J.L. Gifford, M.P. Walsh, H.J. Vogel, *Biochem. J.* 405 (2007) 199–221.
- [27] S.K. Won, Y.J. Lee, H.Y. Lee, Y.K. Heo, M. Cho, H.T. Cho, *Plant Physiol.* 150 (3) (2009) 1459–1473, <https://doi.org/10.1104/pp.109.140905>.
- [28] B.M. Waters, *New Phytol.* 190 (2011) 510–513.
- [29] M.C. Bonza, G. Loro, S. Behera, A. Wong, J. Kudla, A. Costa, *Plant Physiol.* 163 (3) (2013) 1230–1241, <https://doi.org/10.1104/pp.113.226050>.
- [30] M.M. Bradford, *Anal. Biochem.* 72 (1976) 248–254.
- [31] H. Gut, P. Dominici, S. Pilati, A. Astegno, M.V. Petoukhov, D.I. Svergun, M.G. Grütter, G. Capitani, *J. Mol. Biol.* 392 (2009) 334–351.
- [32] V. Marino, A. Astegno, M. Pedroni, F. Piccinelli, D. Dell'Orco, *Nanoscale* 6 (2014) 412–423.
- [33] A. Astegno, V. La Verde, V. Marino, D. Dell'Orco, P. Dominici, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1864 (2016) 297–307.
- [34] A. Allegrini, A. Astegno, V. La Verde, P. Dominici, *J. Biochem.* 161 (2017) 349–360.
- [35] A. Astegno, A. Giorgetti, A. Allegrini, B. Cellini, P. Dominici, *Biomed. Res. Int.* 2013 (2013) 701536.
- [36] A. Astegno, E. Maresi, M. Bertoldi, V. La Verde, A. Paiardini, P. Dominici, *Biochem. J.* 474 (2017) 939–955.
- [37] S.-D. Yoo, Y.-H. Cho, J. Sheen, *Nat. Protoc.* 2 (7) (2007) 1565–1572 PubMed PMID: 17585298.
- [38] E. Teardo, L. Carraretto, S. Wagner, E. Formentin, S. Behera, S. De Bortoli, V. Larosa, P. Fuchs, F. Lo Schiavo, A. Raffaello, R. Rizzuto, A. Costa, M. Schwarzländer, I. Szabò, *Plant Physiol.* 173 (2017) 1355–1370.
- [39] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, *Bioinformatics* 22 (2006) 195–201.
- [40] P. Benkert, M. Biasini, T. Schwede, *Bioinformatics* 27 (2011) 343–350.
- [41] S.T. Rao, S. Wu, K.A. Satyshur, M. Sundaralingam, K.-Y. Ling, C. Kung, *Protein Sci.* 2 (1993) 436–447.
- [42] C. Combet, C. Blanchet, C. Geourjon, G. Deléage, *Trends Biochem. Sci.* 25 (2000) 147–150.
- [43] J.L. Gifford, M. Jamshidiha, J. Mo, H. Ishida, H.J. Vogel, *Plant Cell* 25 (2013) 4512–4524.
- [44] S. Ohki, M. Ikura, M. Zhang, *Biochemistry* 36 (1997) 4309–4316.
- [45] M. Ikura, O. Minowa, K. Hikichi, *Biochemistry* 24 (1985) 4264–4269.
- [46] A. Malmendal, G. Carlstrom, C. Hambraeus, T. Drakenberg, S. Forsen, M. Akke, *Biochemistry* 37 (8) (1998) 2586–2595.
- [47] M. Zhang, T. Tanaka, M. Ikura, *Nat. Struct. Biol.* 2 (1995) 758–767.
- [48] S.M. Gagne, S. Tsuda, M.X. Li, M. Chandra, L.B. Smillie, B.D. Sykes, *Protein Sci.* 3 (1994) 1961–1974.
- [49] R. Gilli, D. Lafitte, C. Lopez, M. Kilhoffer, A. Makarov, C. Briand, J. Haiech, *Biochemistry* 37 (1998) 5450–5456.
- [50] R. Dagher, S. Peng, S. Gioria, M. Fève, M. Zeniou, M. Zimmermann, C. Pigault, J. Haiech, M.-C. Kilhoffer, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1813 (2011) 1059–1067.
- [51] S. Mukherjee, P.M. Mohan, K.V. Chary, *Biochemistry* 46 (2007) 3835–3845.
- [52] G. Wu, Z. Gao, A. Dong, S. Yu, *Int. J. Biol. Macromol.* 50 (2012) 1011–1017.
- [53] S. Linse, P. Brodin, C. Johansson, E. Thulin, T. Grundstrom, S. Forsen, *Nature* 335 (1988) 651–652.
- [54] S. Linse, A. Helmersson, S. Forsen, *J. Biol. Chem.* 266 (1991) 8050–8054.
- [55] M. Comte, Y. Maulet, J.A. Cox, *The Biochemical journal* 209 (1983) 269–272.
- [56] S.H. Seeholzer, M. Cohn, J.A. Putkey, A.R. Means, H.L. Crespi, *Proc. Natl. Acad. Sci.* 83 (1986) 3634–3638.
- [57] B. Kuhlman, H.Y. Yang, J.A. Boice, R. Fairman, D.P. Raleigh, *J. Mol. Biol.* 270 (1997) 640–647.
- [58] J.S. Smith, J.M. Scholtz, *Biochemistry* 37 (1998) 33–40.
- [59] A. Astegno, E. Maresi, V. Marino, P. Dominici, M. Pedroni, F. Piccinelli, D. Dell'Orco, *Nanoscale* 6 (2014) 15037–15047.
- [60] M. Ikura, G.M. Clore, A.M. Gronenborn, G. Zhu, C.B. Klee, A. Bax, *Science* 256 (1992) 632–638.
- [61] W.E. Meador, A.R. Means, F.A. Quioco, *Science* 257 (1992) 1251–1255.
- [62] C.H. Yun, J. Bai, D.Y. Sun, D.F. Cui, W.R. Chang, D.C. Liang, *Acta Crystallogr D Biol Crystallogr* 60 (2004) 1214–1219.
- [63] S. Kumar, M. Mazumder, N. Gupta, S. Chattopadhyay, S. Gourinath, *FEBS Lett.* 590 (2016) 3029–3039.
- [64] M. Ikura, J.B. Ames, *Proc. Natl. Acad. Sci. U. S. A.* 103 (2006) 1159–1164.
- [65] A. Villarreal, M. Tagliatalata, G. Bernardo-Seisdedos, A. Alaimo, J. Aguirre, A. Alberdi, C. Gomis-Perez, M.V. Soldovieri, P. Ambrosino, C. Malo, P. Areso, *J. Mol. Biol.* 426 (15) (2014) 2717–2735.
- [66] M.A. Wilson, A.T. Brunger, *J. Mol. Biol.* 301 (2000) 1237–1256.
- [67] X. Zhu, C. Dunand, W. Snedden, J.P. Galaud, *Trends Plant Sci.* 20 (2015) 483–489.
- [68] H. Tidow, P. Nissen, *FEBS J.* 280 (2013) 5551–5565.
- [69] F. Spyrikis, A. Bidon-Chanal, X. Barril, F.J. Luque, *Curr. Top. Med. Chem.* 11 (2011) 192–210.
- [70] K.W. Bender, W.A. Snedden, *Plant Physiol.* 163 (2013) 486–495.
- [71] H. Zeng, L. Xu, A. Singh, H. Wang, L. Du, B.W. Poovaiah, *Frontiers in Plant Science*, vol. 6, Frontiers Media S.A., 2015, p. 600.
- [72] T.A. DeFalco, K.W. Bender, W.A. Snedden, *Biochem. J.* 425 (2010) 27–40.
- [73] C. Cheval, D. Aldon, J.-P. Galaud, B. Ranty, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1833 (2013) 1766–1771.
- [74] S. Linse, S. Forsen, *Adv. Second Messenger Phosphoprotein Res.* 30 (1995) 89–151.
- [75] A. Malmendal, J. Evenas, S. Forsen, M. Akke, *J. Mol. Biol.* 293 (1999) 883–899.
- [76] A. Malmendal, J. Evenas, E. Thulin, G.P. Gippert, T. Drakenberg, S. Forsen, *J. Biol. Chem.* 273 (1998) 28994–29001.
- [77] Z. Grabarek, *Biochim. Biophys. Acta* 1813 (2011) 913–921.
- [78] K.T. O'Neil, W.F. DeGrado, *Trends Biochem. Sci.* 15 (1990) 59–64.
- [79] A. Scaloni, N. Miraglia, S. Orrù, P. Amodeo, A. Motta, G. Marino, P. Pucci, *J. Mol. Biol.* 277 (1998) 945–958.
- [80] A.M. Weljie, H.J. Vogel, *Protein Eng.* 13 (2000) 59–66.



# Binding of calcium and target peptide to calmodulin-like protein CML19, the centrin 2 of *Arabidopsis thaliana*



Valentina La Verde, Matteo Trande, Mariapina D'Onofrio, Paola Dominici, Alessandra Astegno\*

Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134, Italy

## ARTICLE INFO

### Article history:

Received 14 June 2017

Received in revised form 7 November 2017

Accepted 8 November 2017

Available online 10 November 2017

### Keywords:

Calmodulin-like protein

Centrin

Ca<sup>2+</sup> sensor

## ABSTRACT

Calmodulin-like protein 19 (CML19) is an *Arabidopsis* centrin that modulates nucleotide excision repair (NER) by binding to RAD4 protein, the *Arabidopsis* homolog of human Xeroderma pigmentosum complementation group C protein. Although the necessity of CML19 as a part of the RAD4 plant recognition complex for functional NER is known at a cellular level, little is known at a molecular level. Herein, we used a combination of biophysical and biochemical approaches to investigate the structural and ion and target-peptide binding properties of CML19. We found that CML19 possesses four Ca<sup>2+</sup>-specific binding sites, two of high affinity in the N-terminal domain and two of low affinity in the C-terminal domain. Binding of Ca<sup>2+</sup> to CML19 increases its alpha-helix content, stabilizes the tertiary structure, and triggers a conformational change, resulting in the exposure of a hydrophobic patch instrumental for target protein recognition. Using bioinformatics tools we identified a CML19-binding site at the C-terminus of RAD4, and through *in vitro* binding experiments we analyzed the interaction between a 17-mer peptide representing this site and CML19. We found that the peptide shows a high affinity for CML19 in the presence of Ca<sup>2+</sup> (stoichiometry 1:1) and the interaction primarily involves the C-terminal half of CML19.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

Calcium (Ca<sup>2+</sup>) is a universal second messenger with crucial roles in several plant developmental processes and stress responses [1–3]. Indeed, one of the first events linked to external stimuli perception is intracellular change of free Ca<sup>2+</sup> concentrations. These variations, also called Ca<sup>2+</sup> signatures, vary in their spatio-temporal patterns (distribution in the cell, amplitude, frequency and kinetics) based on the specific stimulus and, thus, are thought to transmit the primary information from the environmental changes. However, the conversion of this information into biological appropriate responses requires a specific decoding process via Ca<sup>2+</sup>-binding proteins.

In plants, three main families of Ca<sup>2+</sup> binding proteins have been characterized: the calmodulin (CaM) and calmodulin-like (CML) proteins, the Ca<sup>2+</sup>-dependent protein kinases (CDPKs) and the calcineurin-B like (CBL) proteins [4–6]. Common to all these proteins is the presence of EF-hand domains which consist of a helix-loop-helix structure that is responsible for the binding of Ca<sup>2+</sup> ions. The coordination of a Ca<sup>2+</sup> selective metal binding loop in

canonical EF-hands has been shown to form a pentagonal bipyramidal arrangement with amino acids at positions 1, 3, 5, 7, 9, and 12 of the loop region [7]. An accepted mechanism for these proteins is that the binding of Ca<sup>2+</sup> ions triggers a conformational change resulting in the modulation of enzymatic activity of the Ca<sup>2+</sup> sensor itself or allowing it to interact with downstream targets [8].

While the role of CaM is well described, CML proteins remain largely uncharacterized at both the structural and functional levels. Examples of CML proteins functions have been covered in recent reviews [1,2,9–11], providing evidence that these proteins are likely not the result of functional redundancy, but have central roles in the coordination of plant responses to different external stimuli. However, due to the absence of empirical data on the role of Ca<sup>2+</sup> binding to CMLs, very few CMLs have been demonstrated to work as Ca<sup>2+</sup> sensors and/or have been thoroughly characterized from biophysical and structural points of view.

In the present work, we characterize CML19 from *Arabidopsis thaliana*, which is also known as centrin 2. Centrins are a class of highly evolutionarily conserved CML proteins found in protists, fungi, plants, and animals [12–15]. In general, centrins are acidic proteins of about 20 kDa that are associated with centrosomes and the contractile fiber system with a role in different cellular functions such as centrosome modulation and nucleotide excision repair (NER) of DNA. As for CaMs, centrins fold into two indepen-

\* Corresponding author at: Department of Biotechnology, University of Verona, Strada Le Grazie 15, Verona (VR), Italy.

E-mail address: [alessandra.astegno@univr.it](mailto:alessandra.astegno@univr.it) (A. Astegno).

dent domains joined by an  $\alpha$ -helical flexible linker region. Each domain contains a pair of EF-hands that can bind two  $\text{Ca}^{2+}$  ions. The presence of additional amino acids at the N-terminal domain and a conserved aromatic residue (Tyr or Phe) at the C-terminal end are features that are unique to centrin among the members of the large CaM superfamily.

CML19 contains four EF-hands and is highly related to human centrin 2 (HsCen2), possessing about 56% amino acid identity and 71% similarity within their EF-hand motifs [16]. HsCen2 plays a role in the NER system, which is responsible for repair of DNA damage induced by UV through interaction of its C-terminal domain with the Xeroderma pigmentosum group C (XPC) protein as the principal NER damage recognition factor [17,18]. Similar to HsCen2, it has been shown that CML19 modulates NER by  $\text{Ca}^{2+}$  dependent complex formation with RAD4 protein, the Arabidopsis homolog of human XPC protein [19]. Although the role of CML19 as a part of the RAD4 plant recognition complex for functional NER is known at a cellular level, little is known at a molecular level.

In the present study, we report an in depth examination of the biophysical and structural properties of CML19 by determining its detailed  $\text{Ca}^{2+}$  binding parameters and  $\text{Ca}^{2+}$ -dependent functional properties. Moreover, we identified the specific CML19-binding site at the C-terminus of its physiological partner RAD4 and characterized the molecular association of a 17-mer peptide representing this site and CML19. Collectively, our data support conclusive evidence that CML19 displays characteristic properties of a  $\text{Ca}^{2+}$  sensor protein and that the interaction of the peptide with the centrin is strictly  $\text{Ca}^{2+}$  dependent and primarily involves the C-terminal half of CML19.

## 2. Materials and methods

### 2.1. Protein production

The complete cDNA of *Arabidopsis thaliana* CML19 (At4g37010) in pUNI51 vector was obtained from Arabidopsis Information Resource (TAIR). The clone contains a premature stop codon, which was removed using the Quick Change Mutagenesis kit (Agilent) and confirmed by DNA sequencing. Cloning and expression procedures were as described [20]. The corrected coding region was amplified using a forward primer containing a *NcoI* site and 6 His-tag (5'-CCATGGGCCATCATCATCATCATTCGGAAGCAGCAC-3') and a reverse primer containing a *BamHI* site and a stop codon (5'-GGATCCTAGCCGTAAGAGGTTCTTTCATCATCTTC-3'). The amplicon was then sub-cloned into pET15b and the recombinant vector was used to transform DH5 $\alpha$  cells.

The construct encoding N-terminally truncated CML19 (residues 20–167) was generated by PCR using the cDNA of CML19 full length as template, a forward primer containing a *NcoI* site (5'-CCATGGGCCACCACCACCACCACCTTGACCAAT-CAGAAG-3'), and the same reverse primer used for the full length CML19. The verified cDNA sequence was then cloned into the corresponding restriction sites of pET15b for cytosolic expression in *E.coli* Rosetta.

The expression of the truncated variant (hereafter referred to as CML19) was performed by growing the cell culture for 18 h at 25 °C in LB medium following addition of 0.4 mM IPTG.

The bacterial pellet obtained from 1-L culture was resuspended in 5 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM imidazole containing 0.1 mg/mL lysozyme, 100  $\mu\text{g}/\text{mL}$  DNAase (Sigma), 10 mM  $\text{MgCl}_2$ , stirred at room temperature for 30 min, and then sonicated on ice. After centrifugation, the supernatant containing the soluble protein was loaded onto a Ni-sepharose column that was pre-equilibrated with 5 mM Tris-HCl at pH 7.5, 150 mM KCl and 10 mM imidazole buffer. A linear gradient from 10 to 500 mM imidazole was then

applied. The fractions containing CML19 protein were pooled and extensively dialyzed against 5 mM Tris-HCl at pH 7.5, 150 mM KCl. The protein concentration was estimated using the Bradford assay (Sigma) [21]. The final yield was about 25 mg/L culture.

$^{15}\text{N}$  and  $^{13}\text{C}$  labeling for NMR spectroscopy was obtained using M9 minimal medium supplemented with 1 g/L  $^{15}\text{NH}_4\text{Cl}$  and 4 g/L  $^{13}\text{C}$ -glucose as the sole sources of nitrogen and carbon. In M9 medium the final yield was about 10 mg/L culture.

The peptide corresponding to the RAD4 portion comprised between residues 762–778 (RAD4p, RWYQLLSILTRERLKN) was synthesized by GenScript USA Inc. (New Jersey, 144 USA). The purity of RAD4p, estimated by HPLC, was >90% and its concentration was determined using its predicted molar extinction coefficient ( $\epsilon_{280} = 6990 \text{ M}^{-1} \text{ cm}^{-1}$ , <http://web.expasy.org/protparam/>).

### 2.2. Size exclusion chromatography

The Stokes radius ( $R_s$ ) of apo-,  $\text{Mg}^{2+}$ -bound and  $\text{Ca}^{2+}$ -bound CML19 was estimated by size exclusion chromatography (SEC) in 5 mM Tris-HCl, 150 mM KCl pH 7.5 using a Superose 12 column (10/300 GL, GE Healthcare, Uppsala, Sweden) as described [22–25]. The reported values represent means  $\pm$  S.E of triplicate experiments.

### 2.3. Electrophoretic mobility shift assays

The  $\text{Ca}^{2+}$ -induced mobility shift of CML19 was assayed on a 15% SDS-PAGE gel after incubating 2  $\mu\text{g}$  of CML19 with 5 mM  $\text{CaCl}_2$  or 5 mM EGTA for 30 min at RT as described [22,23].

Peptide-binding stoichiometry was analyzed by incubation of CML19 (20  $\mu\text{M}$ ) 20 min at RT with increasing concentrations of RAD4p (5–80  $\mu\text{M}$ ) in 100 mM Tris-HCl, 4 M urea, and 5 mM  $\text{CaCl}_2$  or 5 mM EGTA pH 7.5 as described [22]. Samples were assayed under native conditions on a 12.5% continuous gel containing 5 mM  $\text{CaCl}_2$  or 5 mM EGTA and 4 M urea to prevent non-specific interactions.

### 2.4. Nuclear magnetic resonance (NMR) spectroscopy

All NMR experiments were performed on a 600 MHz Bruker Avance III spectrometer (Bruker, Karlsruhe, Germany), equipped with a cryogenic probe. The spectra were recorded in a buffer containing 5 mM Tris-HCl, 50 mM KCl pH 7.5 at protein concentration of 0.5 mM in the presence of 5 mM EGTA, 5 mM  $\text{MgCl}_2$ , or 5 mM  $\text{CaCl}_2$  at 25 °C.

A standard pulse sequence was used to record  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single-quantum coherence (HSQC) spectra, with pulsed field gradients for cancellation of artifacts and for the solvent signal suppression.  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments were acquired with a data matrix consisting of 2K (F2,  $^1\text{H}$ )  $\times$  256 (F1,  $^{15}\text{N}$ ) complex points, spectral windows of 8417.51 Hz ( $^1\text{H}$ )  $\times$  2128.63 Hz ( $^{15}\text{N}$ ), 8 transients, and 1.2 s relaxation delay.

Standard triple resonance experiments, HNCA, HN(CO)CA, HNCACB and CBCA(CO)NH were recorded on  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled samples to achieve partial sequence-specific backbone atoms assignment of CML19 bound to  $\text{Ca}^{2+}$ . The spectra were acquired with a data matrix consisting of 2K (F3,  $^1\text{H}$ )  $\times$  90 (F2,  $^{15}\text{N}$ )  $\times$  120 (F1,  $^{13}\text{C}$ ) complex points, spectral windows of 8417.51 Hz ( $^1\text{H}$ )  $\times$  2128.63 Hz ( $^{15}\text{N}$ )  $\times$  11318.16 Hz ( $^{13}\text{C}$ ), 8 transients, and 1.5 s relaxation delay.

All spectra were processed with Topspin3.2 (Bruker, Karlsruhe, Germany).

### 2.5. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed using a TA Instrument Nano ITC at 25 °C with 250 rpm stirring speed and

50 injections at 300 s intervals. The binding of  $\text{Ca}^{2+}$  to apo- or  $\text{Mg}^{2+}$ -saturated CML19 was determined by titrating 1  $\mu\text{L}$  of 4 mM  $\text{CaCl}_2$  solution into the sample cell (200  $\mu\text{L}$ ) containing 120  $\mu\text{M}$  protein.  $\text{Mg}^{2+}$ -saturated CML19 was obtained by incubation of apo-CML19 with 5 mM  $\text{MgCl}_2$  for 15 min at 25 °C. The titration of  $\text{Mg}^{2+}$  into apo-CML19 was performed by injecting 2  $\mu\text{L}$  of 10 mM  $\text{MgCl}_2$  solution into 120  $\mu\text{M}$  apo-CML19. The buffer used was 5 mM Tris-HCl, 150 mM KCl, pH 7.5 and was decalcified using a Chelex-100 ion-exchange resin (Sigma) as previously described [23,26]. All solutions were thoroughly degassed prior to each experiment. In order to remove contaminating  $\text{Ca}^{2+}$ , the syringes and ITC sample cell were soaked with EDTA and extensively washed with  $\text{Ca}^{2+}$ -free buffer before sample loading.

Control experiments, consisting of injecting ligand solutions into the buffer, were performed to evaluate the heat of dilution which was subtracted from the raw heats in the presence of protein.

Data were analyzed using Origin-based software and fitted using the sequential binding site model. The reported values represent the mean  $\pm$ S.E. of at least three independent titrations using two different CML19 preparations.

### 2.6. Limited proteolysis

CML19 at 28  $\mu\text{M}$  was incubated with trypsin (1:500 w/w) in 5 mM Tris-HCl, 150 mM KCl pH 7.5 in the presence of 5 mM  $\text{CaCl}_2$ , or 5 mM EGTA, respectively, at 25 °C. The reaction was followed on SDS-PAGE by withdrawing a 15  $\mu\text{L}$  aliquot at regular times and stopping the reaction by boiling for 5 min in reducing sample buffer. The intensity of protein band stained with Coomassie blue were analyzed using Chemidoc Image Lab Software (Biorad).

### 2.7. Spectroscopic measurements

Far-UV CD spectra were obtained using a Jasco J-710 spectropolarimeter equipped with a thermostatically controlled compartment. Five scan averaged spectra were recorded between 200 and 250 nm at 25 °C, using a 0.1 cm quartz cell, with a response time of 4 s and a scan rate of 50 nm/min. Samples were dissolved in 5 mM Tris-HCl, 150 mM KCl pH 7.5. The apo-form was obtained by the addition of 5 mM EGTA. The  $\text{Mg}^{2+}$ -CML19 and  $\text{Ca}^{2+}$ -CML19 forms contained 0.5 mM EGTA and 5 mM  $\text{MgCl}_2$  or 5 mM  $\text{CaCl}_2$ , respectively. CD spectra of the CML19-peptide complex were obtained using a protein:peptide ratio of 1:1 in 5 mM Tris-HCl, 150 mM KCl pH 7.5 in the presence of 5 mM EGTA or 5 mM  $\text{CaCl}_2$ .

Fluorescence spectra were recorded on a Jasco FP8200 spectrofluorometer. 1-anilino-8-naphthalene sulfonate (ANS) fluorescence emission was used to evaluate the hydrophobicity of CML19 upon  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  binding and was recorded as described [22].

For the CML19/RAD4p interaction the only Trp residue of RAD4p was selectively excited at 295 nm, and fluorescence emission was recorded from 305 to 450 nm. RAD4p (1  $\mu\text{M}$ ) was titrated with increasing concentrations of CML19 in 50 mM Hepes, 150 mM KCl, pH 7.5 in the presence of 5 mM  $\text{CaCl}_2$  or 5 mM EGTA, respectively.

The  $K_d$  value in the presence of  $\text{Ca}^{2+}$  was calculated using the following equation according to a tight binding hypothesis as described [27–29]:

$$Y = Y_{\max} \frac{e_0 + l_0 + K_d - \sqrt{(e_0 + l_0 + K_d)^2 - 4e_0l_0}}{2e_0} \quad (1)$$

where  $e_0$  and  $l_0$  are the concentrations of RAD4p and CML19 respectively,  $K_d$  is the equilibrium dissociation constant,  $Y$  is the fluorescence change at the CML19 concentration  $l_0$ ,  $Y_{\max}$  is the fluorescence change at saturating CML19 concentrations.

## 3. Results

### 3.1. $\text{Ca}^{2+}$ -induced conformational changes

To avoid experimental difficulties associated with the aggregation propensity of recombinant full length CML19, which was observed also for other centrins [30,31], we focused our studies on a truncated variant of CML19 which is missing the first 19 amino acids (referred to as CML19, see Section 2.1), as they were predicted to be intrinsically disordered by GeneSilico MetaDisorder Service analysis [32].

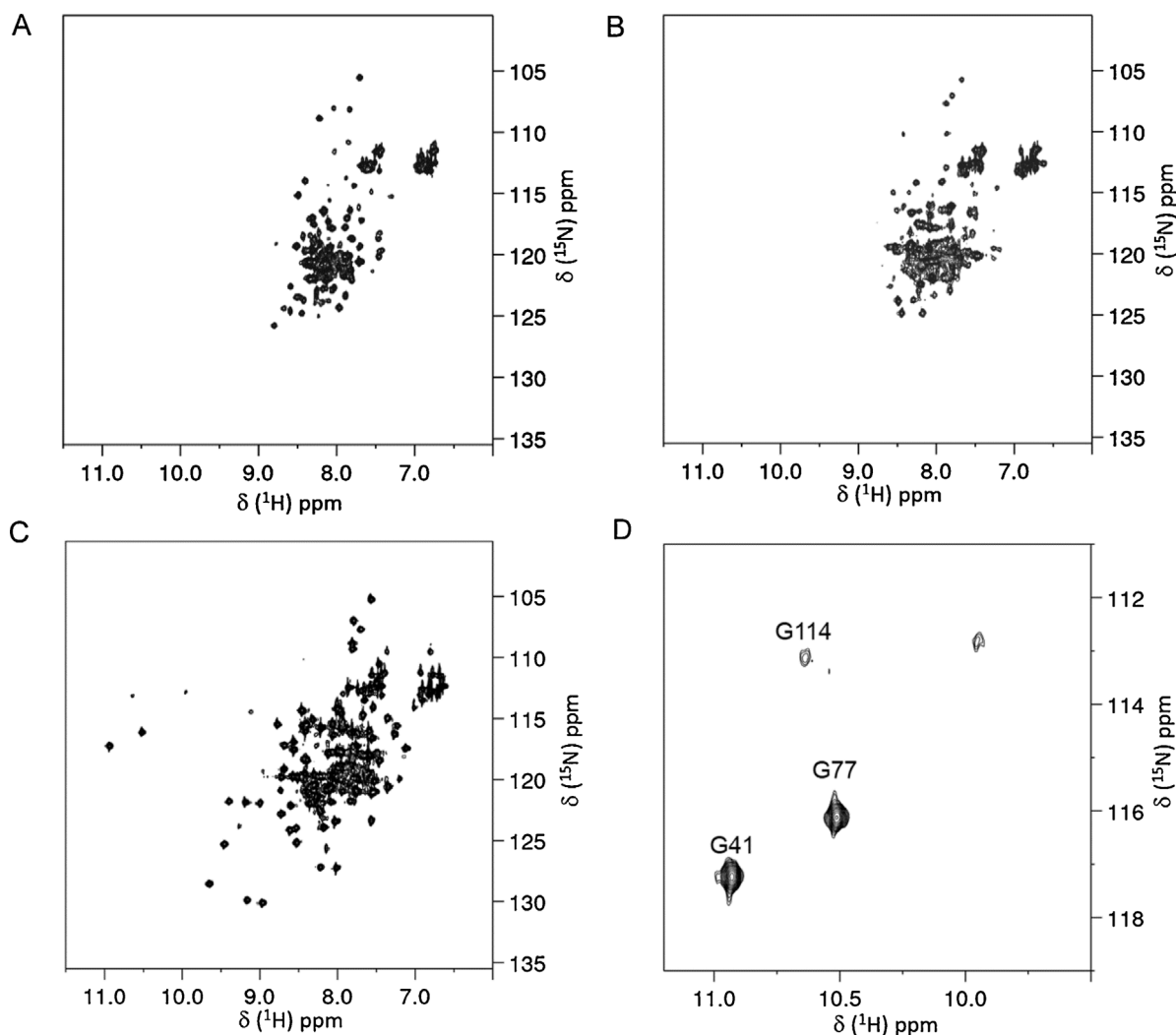
Notably, in plant cells resting cytosolic free  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  concentration is  $\sim$ 100 nM and 0.5–2 mM, respectively [33,34]. Therefore, to analyze the ability of CML19 in responding specifically to  $\text{Ca}^{2+}$  signals in the presence of  $\sim$ 10<sup>3</sup> fold excess of  $\text{Mg}^{2+}$ , we investigated the dynamic and structural properties of CML19 not only in the presence of  $\text{Ca}^{2+}$  but also upon addition of the similar divalent cation  $\text{Mg}^{2+}$ .

Fig. 1 illustrates the comparison of the two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC spectra of CML19 in the presence of saturating amounts of EGTA,  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$ . Each peak in the spectra corresponds to a protein HN group and gives information on the resonance frequencies of the respective <sup>15</sup>N (vertical scale) and <sup>1</sup>H (horizontal scale) nuclei.

The NMR spectrum of the apo-protein (Fig. 1A) consists of cross peaks within a narrow range of chemical shifts (7.2–8.8 ppm) and is characterized by a poor peak dispersion. Moreover, many peaks are broadened beyond detection; this is a clear evidence that the protein, in the absence of  $\text{Ca}^{2+}$ , is characterized by conformational heterogeneity [35]. The addition of  $\text{Mg}^{2+}$  to apo-CML19 (Fig. 1B) did not alter substantially the HSQC spectrum, likely suggesting the absence of specific  $\text{Mg}^{2+}$  binding to the protein. On the contrary, when saturating amounts of  $\text{Ca}^{2+}$  were added to the apo-protein (Fig. 1C), the NMR spectrum changed dramatically, showing an increased number of well-dispersed peaks, meaning that the protein undergoes a conformational rearrangement with a predominant ordered structure. There were a number of distinct chemical shift changes detected in the NMR spectrum of CML19 upon  $\text{Ca}^{2+}$  binding. In particular, four peaks appeared in the low field-shifted <sup>1</sup>H region at about 10 ppm in the presence of  $\text{Ca}^{2+}$  (Fig. 1D); these signals usually arise from highly conserved glycines in position six (Gly<sub>6</sub>) of EF-hand  $\text{Ca}^{2+}$ -binding loops, as a consequence of the hydrogen-bonding network produced following  $\text{Ca}^{2+}$  binding [20,36–41]. In particular, based on standard triple resonance NMR experiments acquired on <sup>15</sup>N,<sup>13</sup>C-labeled samples, we found that the four resonances between 10 and 11 ppm in the <sup>1</sup>H<sub>N</sub> dimension correspond to four glycines and unambiguously assigned the peaks at 10.9, 10.5, and 10.63 ppm to the Gly<sub>6</sub> of the EF-hands, Gly41 (EF-1), Gly77 (EF-2), and Gly114 (EF-3), respectively (Fig. 1D). We were not able to assign the remaining downfield-shifted peak at 9.95 ppm, although it is likely presumed to arise from the corresponding Gly<sub>6</sub> of EF-4 (Gly150). Notably, only partial backbone atoms assignment was obtained. The majority of the resonances of the  $\text{Ca}^{2+}$ -bound CML19 that escaped detection belongs to the C-terminal region. This observation is consistent with the fact that C-terminal half is more flexible, while the N-terminal half shows largely shifted and narrow peaks, indicating a more regular fold.

Analysis of far-UV CD spectra of CML19 in the apo- and holo-states gave complementary information on the secondary structure. The far-UV CD spectra of both apo- and holo-CML19 showed the typical shape and amplitude of a protein with a high amount of helical structure, namely, minima at 208 and 222 nm (Fig. 2A).  $\text{Ca}^{2+}$  binding caused an increase in molar ellipticity as for other centrins [42–44], leading to a switch in the  $\theta_{222}/\theta_{208}$  ratio from 0.79 in the apo-form to 0.87 in the  $\text{Ca}^{2+}$ -bound form (Table 1).





**Fig. 1.** Two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of  $^{15}\text{N}$ -CML19 in its apo-state (A),  $\text{Mg}^{2+}$ - (B),  $\text{Ca}^{2+}$ - (C) bound forms, and the zoom of the glycine downfield peaks of CML19 in complex with  $\text{Ca}^{2+}$  ions (D). The assignment of the glycines of the EF-hands is indicated by the position number of the residue. All the spectra were recorded at 600 MHz and at 298 K. Protein samples were at concentration of 0.5 mM in 5 mM Tris-HCl, 50 mM KCl pH 7.5.

**Table 1**  
 $\Theta_{222}/\Theta_{208}$  ratio of CML19 in the absence and in the presence of divalent cations.

	$\Theta_{208}$ (mdeg)	$\Theta_{222}$ (mdeg)	$\Theta_{222}/\Theta_{208}$
CML19+EGTA	-13.6	-10.7	0.79
CML19+ $\text{Mg}^{2+}$	-13.9	-11.3	0.81
CML19+ $\text{Ca}^{2+}$	-14.3	-12.5	0.87

This may be explained as an increase in  $\alpha$ -helix content due to a stabilization/compactness of some helical elements and/or a reorientation of the  $\alpha$ -helices within the tertiary structure [45]. The analysis of the  $\Theta_{222}/\Theta_{208}$  ratio indicated that the shape of the spectrum corresponding to the  $\text{Ca}^{2+}$ -bound is distinguished from that of the  $\text{Mg}^{2+}$ -bound form. In fact, the spectra exhibited a remarkably similar shape for apo- and  $\text{Mg}^{2+}$ -bound CML19, as quantitatively demonstrated by nearly the same  $\Theta_{222}/\Theta_{208}$  ratio (0.79 versus 0.81, respectively, Table 1).

These CD results are consistent with NMR spectra, as both analyses indicated that  $\text{Mg}^{2+}$  did not induce significant conformational changes, and the apo- and  $\text{Mg}^{2+}$ -bound CML19 species possess highly similar structural properties. It is likely that the minor changes observed on both NMR and CD spectra upon  $\text{Mg}^{2+}$  addition to the protein are the result of non-specific binding of  $\text{Mg}^{2+}$

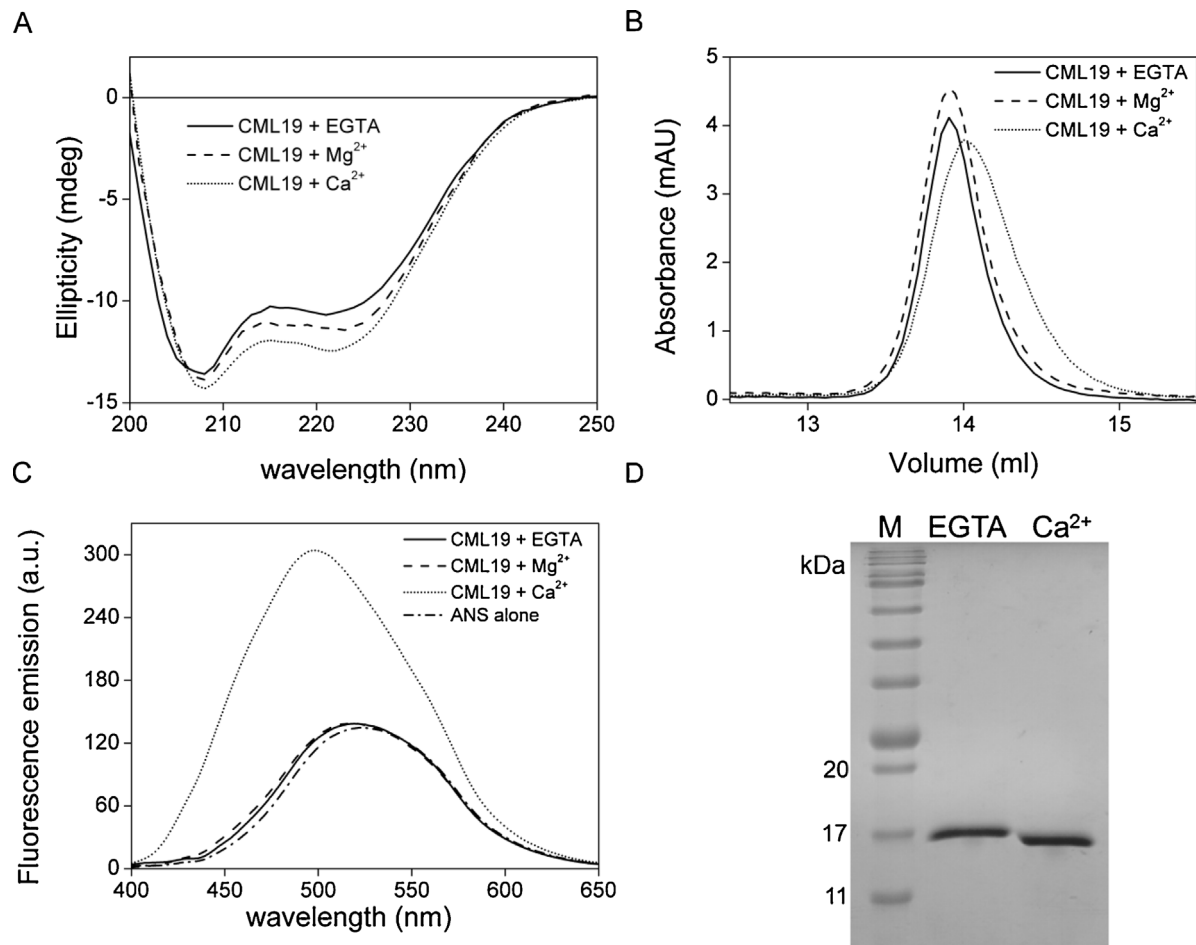
ions (as also indicated by ITC data, see Section 3.2), perhaps to sites that differ from the main  $\text{Ca}^{2+}$  binding sites.

Limited trypsin proteolysis provided additional support for the improved stability of CML19 in the presence of  $\text{Ca}^{2+}$ ; in 5 mM EGTA the protein underwent complete digestion, whereas in the presence of 5 mM  $\text{Ca}^{2+}$  CML19 was stable to proteolysis for more than 60 min (Supplemental Fig. S1).

SEC studies were conducted to determine the Stokes radii ( $R_s$ ) for the peaks corresponding to apo-,  $\text{Ca}^{2+}$ -bound and  $\text{Mg}^{2+}$ -bound CML19, using proteins with known  $R_s$  as standards (Fig. 2B). SEC permits resolving the variations in the hydrated shape of  $\text{Ca}^{2+}$  sensor proteins upon metal binding, thus allowing the possibility to monitor  $\text{Ca}^{2+}$ -induced conformational changes. Recombinant CML19 was found to have a  $R_s$  of  $26.06 \pm 0.02$  Å in the presence of EGTA. The relatively high  $R_s$  was understandable likely because of the elongated shape of CML19, as for many  $\text{Ca}^{2+}$  sensor proteins, e.g. CaMs and centrin from different organisms [22,25,42,46,47]. Addition of  $\text{Ca}^{2+}$  to CML19 caused a decrease in its  $R_s$  ( $25.51 \pm 0.02$  Å), which is likely the result of the conformational change that CML19 undergoes upon  $\text{Ca}^{2+}$  binding. In contrast,  $\text{Mg}^{2+}$  binding did not induce a variation in  $R_s$  ( $26.00 \pm 0.09$  Å) and, therefore, in the hydration shell structure compared to apo-protein.

$\text{Ca}^{2+}$ -induced conformational changes in  $\text{Ca}^{2+}$  sensors are often accompanied by exposure of hydrophobic surfaces, which repre-





**Fig. 2.** Metal-induced conformational changes to CML19. (A) Far-UV CD spectra at 25 °C of 8  $\mu\text{M}$  CML19 in the presence of 5 mM EGTA (solid line), 5 mM  $\text{MgCl}_2$  plus 0.5 mM EGTA (dashed line), and 5 mM  $\text{CaCl}_2$  (dotted line). (B) Superose 12 column 10/300GL elution profiles of CML19 in the presence of 5 mM EGTA (solid line), 5 mM  $\text{MgCl}_2$  plus 0.5 mM EGTA (dashed line), 5 mM  $\text{CaCl}_2$  (dotted line). (C) ANS fluorescence spectra of CML19 in the presence of 5 mM EGTA (solid line), 5 mM  $\text{MgCl}_2$  plus 0.5 mM EGTA (dashed line), and 5 mM  $\text{CaCl}_2$  (dotted line). The spectra of ANS alone (dash dotted line) is also shown. (D) SDS-PAGE mobility shift of CML19 in the presence of 5 mM EGTA or 5 mM  $\text{CaCl}_2$ . For all the experiments the buffer was 5 mM Tris-HCl, 150 mM KCl pH 7.5.

sents a crucial step in signal transduction. A specific conformational probe for such a  $\text{Ca}^{2+}$  effect on surface hydrophobicity of  $\text{Ca}^{2+}$  sensors is the fluorescence of the hydrophobic fluorophore ANS, since its emitted fluorescence is altered after binding to hydrophobic surfaces. As shown in Fig. 2C, the probe alone exhibits a maximum emission peak at  $\sim 520$  nm. When  $\text{Ca}^{2+}$  binds to CML19, a  $\sim 20$  nm blue shift in the ANS maximum emission ( $\lambda_{\text{max}} \sim 500$  nm) occurred with a 2.2-fold increase in fluorescence intensity in comparison to the spectrum of the ANS-CML19 complex in the presence of EGTA. In contrast, binding of  $\text{Mg}^{2+}$  to the apo-protein led to an emission spectrum that was approximately identical to that of ANS-CML19 in the presence of EGTA, therefore indicating a  $\text{Ca}^{2+}$  and not  $\text{Mg}^{2+}$ -dependent exposure of hydrophobic patches. Moreover, recombinant CML19 shifted to a faster-migrating conformation in the  $\text{Ca}^{2+}$ -bound form during SDS-PAGE (Fig. 2D), which is a hallmark of  $\text{Ca}^{2+}$  sensor proteins [22,23,48].

### 3.2. Calcium binding to CML19

To investigate the intrinsic  $\text{Ca}^{2+}$  binding properties of CML19 we used isothermal titration calorimetry (ITC), which permits estimation of the values of apparent dissociation constants ( $K_d$ ) for multiple binding sites and the associated changes in enthalpy ( $\Delta H$ ). A representative ITC isotherm is shown in Fig. 3A and the thermodynamic parameters of the interaction are listed in Table 2. Titration

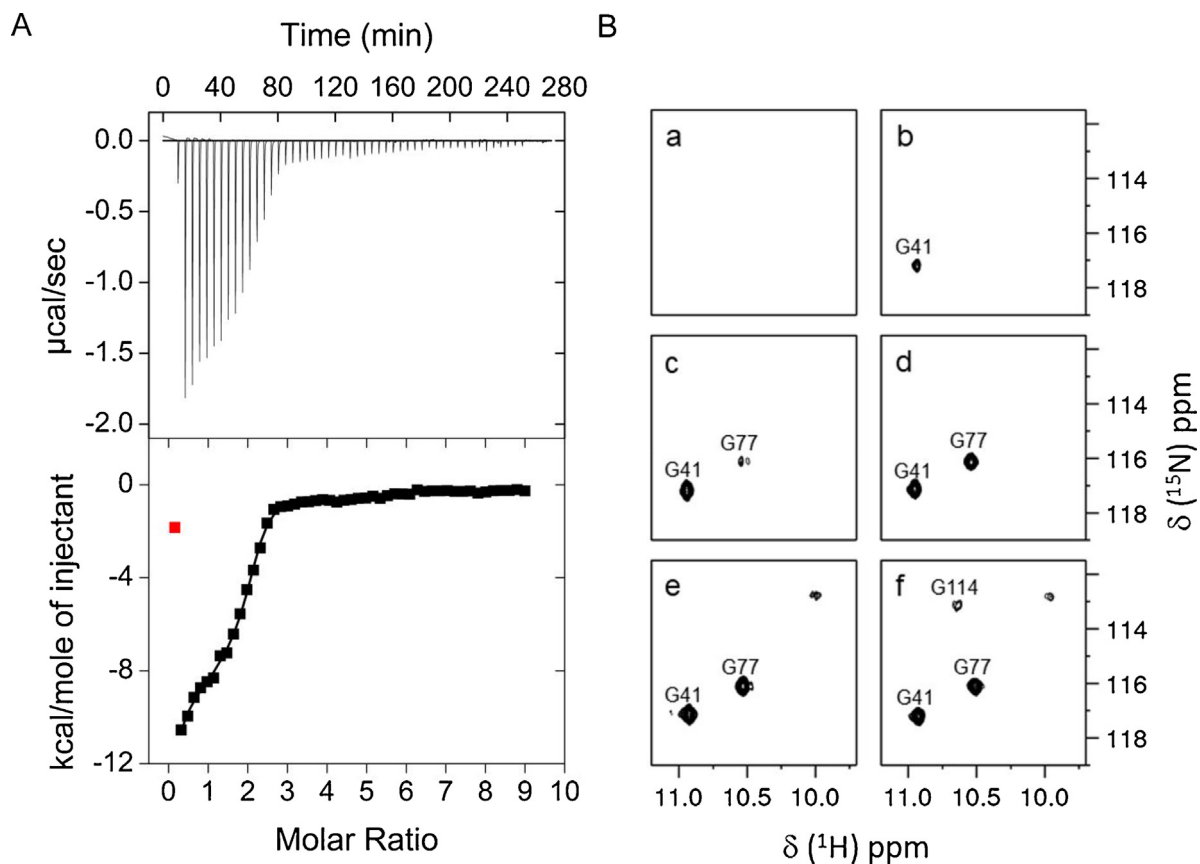
**Table 2**

Thermodynamic parameters describing the binding of  $\text{Ca}^{2+}$  to apo-CML19 obtained from ITC. A four-site sequential binding model was used to fit ITC data which gives the  $\Delta H$  and  $\text{Ca}^{2+}$  association constants ( $K_A$ ) values corresponding to the strength of the first, second, third, and fourth  $\text{Ca}^{2+}$  binding site on CML19. The mean values  $\pm$  S.E. from triplicate experiments are presented.

	$K_A$ ( $\text{M}^{-1}$ )	$\Delta H$ ( $\text{kcal mol}^{-1}$ )
Site 1	$2.9 \text{ E}5 \pm 7.1 \text{ E}4$	$-1.2 \text{ E}4 \pm 1.9 \text{ E}3$
Site 2	$4.1 \text{ E}5 \pm 6.8 \text{ E}4$	$-4.4 \text{ E}3 \pm 0.6 \text{ E}3$
Site 3	$1.4 \text{ E}4 \pm 3.8 \text{ E}3$	$-2.1 \text{ E}3 \pm 0.9 \text{ E}3$
Site 4	$3.7 \text{ E}3 \pm 0.8 \text{ E}3$	$-1.9 \text{ E}3 \pm 0.7 \text{ E}3$

of  $\text{CaCl}_2$  into apo-CML19 results in an exothermic curve demonstrating the binding of four  $\text{Ca}^{2+}$  ions to CML19, two with higher affinity ( $K_{d1} = 3.5 \pm 0.8 \mu\text{M}$ ,  $K_{d2} = 2.4 \pm 0.4 \mu\text{M}$ ) and two with lower affinity ( $K_{d3} = 71 \pm 19 \mu\text{M}$ ,  $K_{d4} = 270 \pm 58 \mu\text{M}$ ) and suggesting a positive cooperativity of  $\text{Ca}^{2+}$  binding to the high-affinity sites. Notably, the results obtained are consistent with the sequence-based prediction of four functional  $\text{Ca}^{2+}$ -binding sites based on PROSITE-ProRule annotation [49].

Since at resting levels the cytosolic free  $\text{Mg}^{2+}$  concentration in plant cells is high (0.5–2 mM), we next determined the  $\text{Ca}^{2+}$  binding isotherm upon incubation of CML19 with 5 mM  $\text{MgCl}_2$ . The  $\text{Ca}^{2+}$  binding curve of the  $\text{Mg}^{2+}$ -saturated CML19 was very similar to that obtained by titration of  $\text{CaCl}_2$  into apo-CML19 and could be fit by the binding of four  $\text{Ca}^{2+}$  ions without significant changes in the



**Fig. 3.** Calcium binding to CML19. (A) ITC analysis of the  $\text{Ca}^{2+}$  binding to apo-CML19. Representative raw trace minus baseline (top) of the calorimetric titration of 4 mM  $\text{CaCl}_2$  into 120  $\mu\text{M}$  apo-CML19 and integrated binding isotherm (bottom). Curve fitting was performed using a sequential binding site model. (B) Downfield region of  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of CML19 recorded as a function of increasing  $\text{Ca}^{2+}$  concentration. The molar ratio of  $\text{Ca}^{2+}$ :protein in each case was 0 (a), 0.1 (b), 0.25 (c), 0.67(d), 8 (e), 16 (f).

derived thermodynamic parameters (data not shown). Since  $\text{Mg}^{2+}$  does not influence the curve, we conclude that all four sites are  $\text{Ca}^{2+}$ -specific and CML19 is a  $\text{Ca}^{2+}$  signaling selective protein that is not able to bind  $\text{Mg}^{2+}$  through the EF-hand domains.

Importantly, addition of  $\text{MgCl}_2$  to apo-CML19 under the same experimental conditions used for addition of  $\text{CaCl}_2$  to apo-CML19 exhibited a profile that was not ascribable to a binding event in terms of either heat change or signal intensity (Supplemental Fig. S2).

The presence of two classes of  $\text{Ca}^{2+}$ -specific binding sites with different affinity was clearly supported by NMR spectra. The existence of distinct structural conformations in apo- and  $\text{Ca}^{2+}$ -bound states of CML19 (Fig. 1) can be exploited to understand the binding pathway of the  $\text{Ca}^{2+}$  ions, in particular by monitoring the downfield shift  $^1\text{H}$  resonances of Gly<sub>6</sub> of the EF-hand binding loop which act as indicators for the  $\text{Ca}^{2+}$ -bound state of a given EF-hand. Therefore,  $\text{Ca}^{2+}$  titration of apo-CML19 was monitored through the acquisition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra. As shown in Fig. 3B, in the  $\text{Ca}^{2+}$  titration of CML19 the NH peaks of the  $\text{Ca}^{2+}$ -bound N-terminal Gly<sub>6</sub> residues of CML19 (Gly41 and Gly77) develop firstly, while the other two glycine peaks (Gly114 and the peak presumed to arise from the  $\text{Ca}^{2+}$ -bound Gly150 of the EF-4) appear at higher  $\text{Ca}^{2+}$  concentrations. Therefore, the Gly<sub>6</sub> residues of the N-terminal lobe begin to reach their  $\text{Ca}^{2+}$ -saturated chemical shift and intensity at  $\text{Ca}^{2+}$ :protein ratios much lower than those of the C-terminal sites. These data clearly confirmed that all four EF-hands of CML19 contain a functional  $\text{Ca}^{2+}$ -binding site, and allowed us to assign the dissociation constants to specific lobes of CML19:  $K_{d1}$  and  $K_{d2}$  to the N-terminal lobe (higher affinity  $\text{Ca}^{2+}$ -binding sites), while  $K_{d3}$

and  $K_{d4}$  to the C-terminal EF-hands (lower affinity  $\text{Ca}^{2+}$ -binding sites). Importantly, this experiment provides information not only on the presence of a Gly<sub>6</sub> downfield-shifted  $^1\text{H}$  peak at its  $\text{Ca}^{2+}$ -bound chemical shifts but on the occupancy (by means of signal intensity) of the  $\text{Ca}^{2+}$ -bound state of an EF-hand. As a consequence, higher than stoichiometric  $\text{Ca}^{2+}$ :protein ratios were necessary to obtain saturation.

### 3.3. Interaction with RAD4 peptide

Although the involvement of CML19 in the RAD4 plant recognition system for functional NER has been described [19], structural and biophysical data for this interaction is almost non-existent. In particular, when we decided to investigate the CML19/RAD4 interaction, it was known that the CML19-binding site was in the C-terminal portion of RAD4 [19], but there was no clear indication about which specific peptide segment of RAD4 was involved in the interaction. The binding sites in the centrin targets possess some common key features: they are relatively small (15–30 residues long), able to adopt, upon centrin binding, an amphipathic  $\alpha$ -helical structure, and they generally possess a hydrophobic W<sup>1</sup>-L<sup>4</sup>-L<sup>8</sup> triad. The HsCen2 binding site of human homolog XPC has been well characterized (interaction  $\text{Ca}^{2+}$ -dependent,  $K_a = 2.2 \times 10^8 \text{ M}^{-1}$ , stoichiometry 1:1) [31,50] and identified as the hydrophobic W<sup>848</sup>-L<sup>851</sup>-L<sup>855</sup> triad. Therefore, we have used bioinformatics methods to analyze the RAD4 sequence and choose an amphipathic  $\alpha$ -helix peptide with a 1,4,8 motif that would be hypothesized to bind CML19 with high affinity. We identified the fragment Arg<sup>762</sup>-Asn<sup>778</sup> with detectable sequence homology to the HsCen2 binding region

of human counterpart XPC; the corresponding peptide (RAD4p, R<sup>762</sup>WYQLSSILTRERLKN<sup>778</sup>) was synthesized and used for experiments.

First, the interaction with RAD4p was examined by native PAGE on mixtures of CML19 and increasing amounts of RAD4p in the presence of Ca<sup>2+</sup> or EGTA. When RAD4p and CML19 were mixed in the presence of Ca<sup>2+</sup>, an extra band appeared on the native gel likely representing the CML19/RAD4p complex. At a 1:1 protein:peptide molar ratio, free CML19 was no longer present (Fig. 4A), suggesting the formation of a well-defined 1:1 complex. Notably, under the same experimental conditions this band was absent in the EGTA-containing sample, suggesting that the binding is primarily Ca<sup>2+</sup>-dependent (Fig. 4B).

The interaction between CML19 and RAD4p was further confirmed by following changes in the fluorescence emission of Trp763 of RAD4p induced by CML19 binding (Fig. 5). This approach was possible thanks to the absence of Trp in CML19. Fluorescence titration gave information on stoichiometry, binding strength, and binding mode. In the presence of Ca<sup>2+</sup> the fluorescence spectrum of the peptide Trp residue changed substantially upon complex formation, being characterized by a significant blue shift (20 nm, from 353 to 333 nm) and a 2.5-fold increase in intensity (Fig. 5A, inset). Solvent-exposed Trp side chains in a polar environment typically possess a fluorescence emission maximum at 350 nm, and therefore these data clearly indicate that RAD4p interacts with CML19 and the interaction leads to the embedding of Trp in a non-polar environment.

The fluorescence titration curve revealed a fast increase in fluorescence intensity that confirmed binding with a 1:1 stoichiometry and gave a  $K_d$  of  $54 \pm 14$  nM for CML19/RAD4p complex in the presence of Ca<sup>2+</sup> (Fig. 5A). The RAD4p displayed a different behavior when interacting with CML19 in the presence of EGTA, as the fluorescence increase was smaller (1.6-fold increase) and less blue-shifted (12 nm, from 353 to 341 nm) than in the presence of Ca<sup>2+</sup> (Fig. 5B, inset). Importantly, the fluorescence response in the CML19 titration performed in the presence of EGTA yielded a  $K_d$  value of  $8 \pm 1$   $\mu$ M, thus indicating an affinity for CML19/RAD4p complex that is about 150-fold lower than in the presence of Ca<sup>2+</sup> (Fig. 5B). The enhancement of the peptide binding affinity by about two orders of magnitude in the presence of Ca<sup>2+</sup> clearly shows that Ca<sup>2+</sup> activates CML19 for the binding to the target.

The ability of CML19 to bind RAD4p was also analyzed by far-UV CD spectroscopy in order to describe the conformational properties of the CML19/RAD4p interaction. It is known that CaM-binding peptides are induced into an amphipathic helix when binding to Ca<sup>2+</sup>-saturated CaM [51–54]. The spectrum of RAD4p alone in the presence of Ca<sup>2+</sup> was typical for a highly unordered structure, as is normal for linear peptides of this size (Fig. 6A, solid line), while the Ca<sup>2+</sup>-bound CML19 exhibited a CD spectrum with the intense negative bands at 222 nm and 208 nm characteristic of a well-folded protein with  $\alpha$ -helical content (Fig. 6A, dashed line; see also Fig. 2A). Notably, addition of RAD4p to CML19 at 1:1 molar ratio in the presence of Ca<sup>2+</sup> caused a significant increase in the intensity of the dichroic signal (~50%) with conservation of the relative intensity of the different bands (Fig. 6A, dotted line). A modest increase in ellipticity could be related to conformational rearrangement of CML19, although structural studies on EF-hand domains-peptide target complexes have shown that CaM/Ca<sup>2+</sup> does not gain an  $\alpha$ -helical structure upon binding to its target peptides, undergoing only small tertiary changes [54–56]. Therefore, a relevant contribution to the  $\alpha$ -helix CD bands comes from the peptide, which changes from a random coil to  $\alpha$ -helical structure following interaction with CML19. An increase in the peptide-to-protein ratio (2:1) (data not shown) did not result in notable CD changes, in accordance with the 1:1 stoichiometry observed in fluorescence and native PAGE. It is worth noting that the CD spectra taken in the

presence of EGTA led to different results, with a modest change in the far-UV CD signal upon addition of RAD4p to CML19 (Fig. 6B), confirming that CML19 binds to RAD4p primarily in a Ca<sup>2+</sup> dependent mode to form a 1:1 complex in solution.

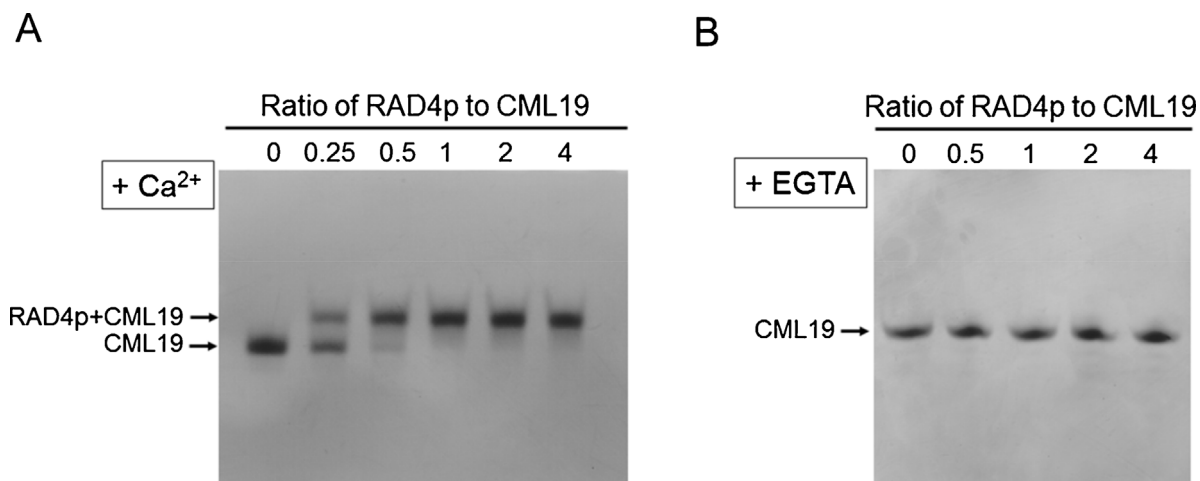
Information on the structural and dynamic properties of the complex formed by Ca<sup>2+</sup>-bound CML19 and RAD4p were also obtained from NMR experiments. Upon addition of the peptide to the Ca<sup>2+</sup>-bound CML19 the NMR spectrum noticeably changed. In particular, comparison of <sup>1</sup>H-<sup>15</sup>N HSQC spectra recorded on the Ca<sup>2+</sup>-CML19 in the absence and presence of RAD4p showed that some peaks experience chemical shift variation, and new peaks appear (Fig. 7). The RAD4p binding-induced changes in NMR spectrum confirmed that, in agreement with native PAGE, CD, and fluorescence experiments, the interaction between the protein and peptide results in well-organized, unique, and stable structure. Notably, detailed analysis of the NMR spectra revealed that the resonances belonging to glycines Gly41 (EF-1) and Gly77 (EF-2) of CML19 remain unperturbed upon protein-peptide interaction, meaning that the CML19 N-terminal domain is not critically involved in RAD4p binding. On the contrary, the other two peaks corresponding to the Gly114 of the EF-3 and likely to Gly150 of EF-4, are characterized by a clear chemical shift variation and an increase in intensity. These observations strongly suggest that the changes in the NMR spectrum originate from modifications in the C-terminal part of CML19. This finding is in accordance with the notion that one of the main centrin-specific binding features is related to the principal role played by the C-terminal domain in target binding.

#### 4. Discussion

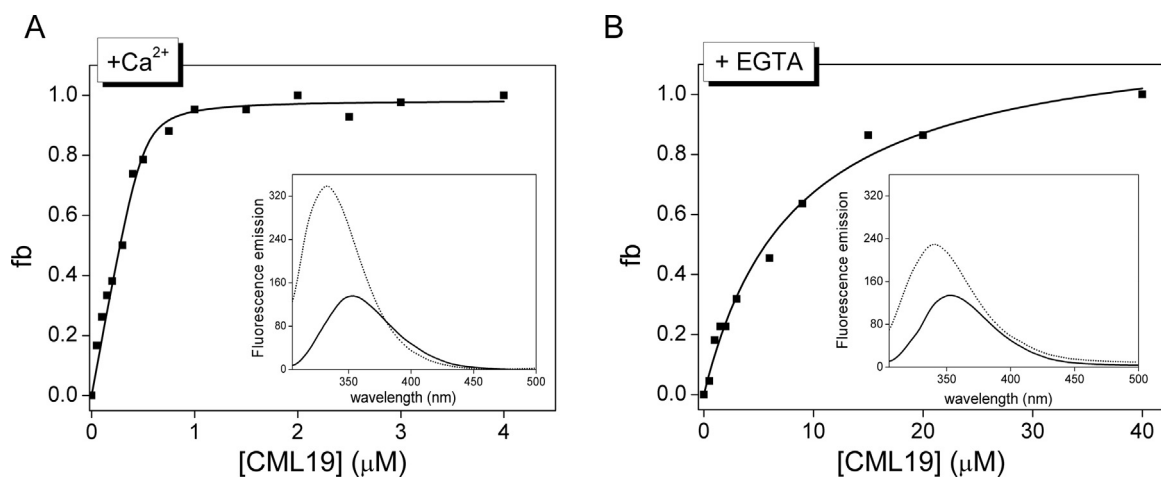
The purpose of this study was to provide a detailed molecular description of the binding of Ca<sup>2+</sup> to CML19 and subsequent protein conformational changes, and to underline the significance of the interaction between CML19 and its physiological binding partner RAD4 by identifying the CML19 binding site on RAD4. The interaction between CML19 and RAD4 has been believed to be dependent on Ca<sup>2+</sup> [19]. Thus, understanding of the effect of metals on CML19 conformation is crucial in order to elucidate its biological function. Herein, we present conclusive evidence that CML19 displays characteristic properties of a Ca<sup>2+</sup> sensor protein and is able to bind to a specific C-terminal region of RAD4 primarily in Ca<sup>2+</sup> dependent manner.

Proteins from the EF-hand superfamily are characterized by a wide range of metal binding properties, involving differences in the number of active binding sites, ion affinity, cooperativity, and Ca<sup>2+</sup>/Mg<sup>2+</sup> discrimination. In particular, centrins display a variety of Ca<sup>2+</sup> binding properties. For instance, human centrin 1 and centrin 2 (HsCen1 and HsCen2) have two C-terminal Ca<sup>2+</sup> binding sites, with modest and low affinity ( $K_d$  of 10  $\mu$ M for site 4 and  $K_d > 100$   $\mu$ M for site 3), whereas the N-terminal domain displays non-physiological ion binding sites ( $K_d > 1$  mM) [41,57,58]; Mg<sup>2+</sup> does not influence Ca<sup>2+</sup> binding. Centrin 3 (HsCen3) possesses one Ca<sup>2+</sup>/Mg<sup>2+</sup> mixed site with high-affinity ( $K_{dCa} = 3$   $\mu$ M,  $K_{dMg} = 11$   $\mu$ M) at the N-terminus and two low-affinity ( $K_d = 140$   $\mu$ M) Ca<sup>2+</sup>-specific sites at the C-terminus [43]. *Scherfella dubia* centrin (SdCen) possesses two high affinity Ca<sup>2+</sup>-binding sites in the N-terminal domain ( $K_d = 3$   $\mu$ M) and one in the C-terminal domain with moderate affinity ( $K_d = 20$   $\mu$ M) [44]. *Chlamydomonas reinhardtii* centrin (CrCen) displays four functional EF-hands with  $K_d$  of 1.2  $\mu$ M and 160  $\mu$ M for Ca<sup>2+</sup> sites in the N-terminal and C-terminal domains, respectively [59].

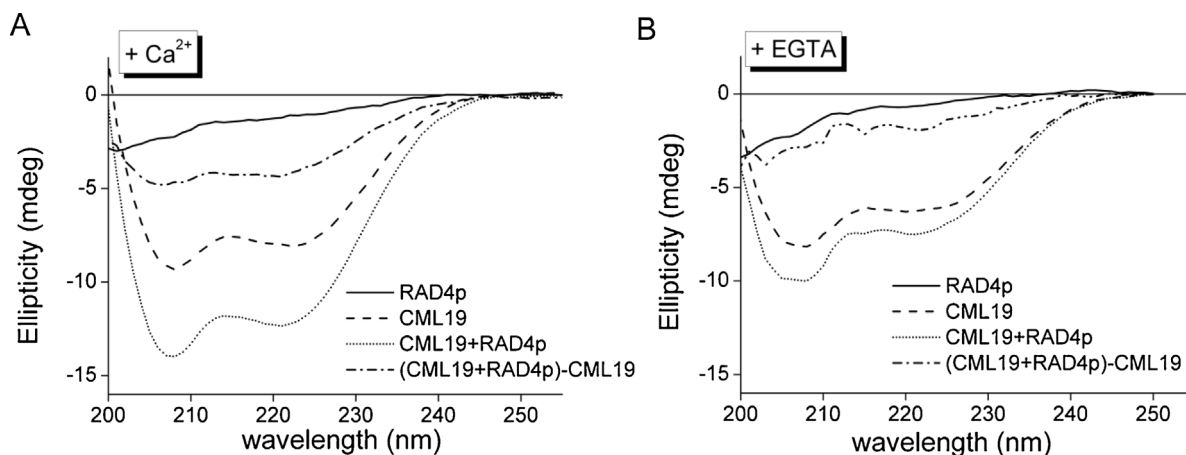
Our study clearly demonstrated that CML19 from *Ara-bidopsis thaliana* possesses four Ca<sup>2+</sup>-specific binding sites and, similar to algae centrins SdCen and CrCen [35,44], the N-terminal domain of CML19 contains the higher affinity



**Fig. 4.** Non-denaturing-polyacrylamide gel mobility shift characterization of the interaction RAD4p/CML19 in the presence of (A) 5 mM  $\text{CaCl}_2$  and (B) 5 mM EGTA. 20  $\mu\text{M}$  CML19 was incubated with increasing concentrations of RAD4p (5–80  $\mu\text{M}$ ) in 100 mM Tris-HCl, 4 M urea pH 7.5.



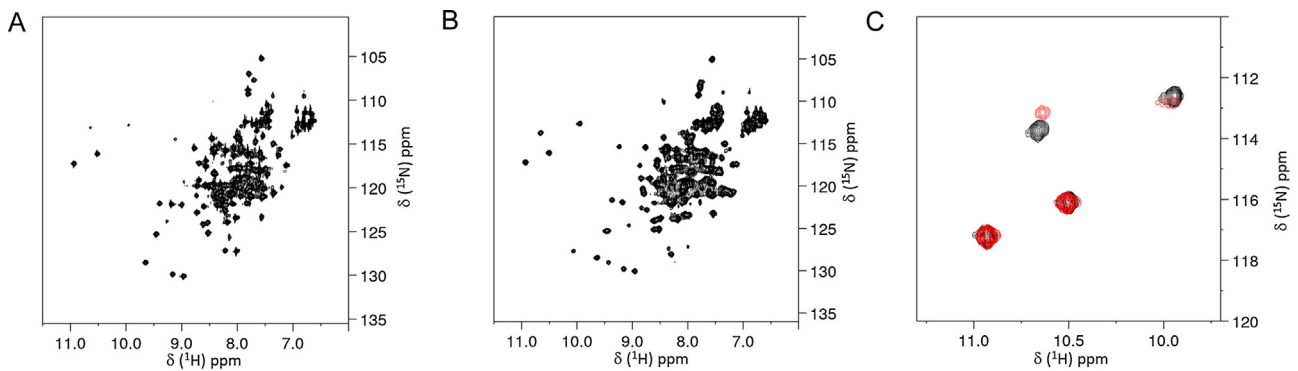
**Fig. 5.** Representative fluorescence titration of 1  $\mu\text{M}$  RAD4p with CML19 in the presence of (A) 5 mM  $\text{CaCl}_2$  and (B) 5 mM EGTA. Insets, Trp fluorescence emission spectra of RAD4p alone (solid line) and RAD4p/CML19 complex (dotted line) in the presence of (A) 5 mM  $\text{CaCl}_2$  and (B) 5 mM EGTA. The buffer was 50 mM HEPES, 150 mM KCl, at pH 7.5.



**Fig. 6.** Far-UV CD spectra changes in secondary structure for CML19/RAD4p complex in the presence of  $\text{CaCl}_2$  (A) and EGTA (B). Far-UV CD spectra of 5  $\mu\text{M}$  RAD4p alone (black line), 5  $\mu\text{M}$  CML19 alone (dashed line), and RAD4p/CML19 complex (dotted line) at 1:1 molar ratio. The CD spectrum resulting from subtraction of the spectrum of CML19/RAD4p complex from that of CML19 alone was also shown (dash dotted line).

$\text{Ca}^{2+}$  sites ( $K_{d1} = 3.5 \pm 0.8 \mu\text{M}$ ,  $K_{d2} = 2.4 \pm 0.4 \mu\text{M}$ ), while the C-terminal region comprises the lower affinity  $\text{Ca}^{2+}$  sites ( $K_{d3} = 71 \pm 19 \mu\text{M}$ ,  $K_{d4} = 270 \pm 58 \mu\text{M}$ ). Moreover, the protein

shows strict  $\text{Ca}^{2+}$  specificity as it is not able to bind  $\text{Mg}^{2+}$  at  $\text{Ca}^{2+}$ -binding sites. Notably, the N-terminal (without the first variable and disordered 19 residues) and C-terminal regions of



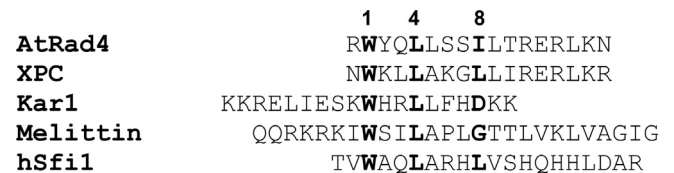
**Fig. 7.** Two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of  $\text{Ca}^{2+}$ -CML19 in the absence (A) and in the presence of RAD4p at a 1:1 molar ratio (B). (C) Zoom of the glycine downfield peaks of the overlaid  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra recorded on  $\text{Ca}^{2+}$ -CML19 before (red) and after the addition of RAD4p (black).

CML19 are ~61% and ~48% identical in their primary sequence with CrCen, and ~61% and ~58% identical with SdCen, respectively. Therefore, CML19 appears to possess molecular properties more similar to algal centrin compared to centrin from the same phylogenetic subfamily such as HsCen1 and HsCen2, suggesting that its functional mode may be similar and/or common to those of algal centrin.

Association of  $\text{Ca}^{2+}$  at the four EF-hands of CML19 provides the protein with important structural stability, as judged by changes in secondary and tertiary structure and hydrodynamic behavior. In particular, binding of  $\text{Ca}^{2+}$  induced a structural rearrangement that results in improvement in the quality of the NMR spectrum and increased molar ellipticity. Moreover,  $\text{Ca}^{2+}$  binding to CML19 decreased protein susceptibility to proteases. Notably, for both apo- and holo-CML19 the  $R_s$  resulted in overestimation of its molecular mass. In particular, the apo-form of CML19 revealed a molecular mass that was 1.7-fold higher than that calculated. Since globular proteins were used for column calibration, this defined overestimation of molecular mass suggests that CML19 is likely to be elongated molecule, perhaps with a domain architecture similar to CaM.  $\text{Ca}^{2+}$  binding decreased the apparent molecular mass, confirming that the protein undergoes a conformational change upon  $\text{Ca}^{2+}$  binding. This change is also accompanied by the exposure of hydrophobic regions on the surface, as observed in ANS fluorescence, which is likely instrumental for  $\text{Ca}^{2+}$ -dependent interactions with the target proteins [51,52,56,60]. Indeed, the complex formation of  $\text{Ca}^{2+}$  sensors with their physiological targets is usually based on the hydrophobic interactions between the hydrophobic surface of  $\text{Ca}^{2+}$  sensor proteins and crucial hydrophobic anchoring residues of the interacting targets.

Centrin are thought to interact with more than one cellular target through a hydrophobic triad,  $\text{W}^1\text{L}^4\text{L}^8$ . Considering a direct interaction between CML19 and RAD4 [19], using different bioinformatics tools we identified a CML19-binding site located at the C-terminus of RAD4 and through *in vitro* binding experiments we analyzed the interaction between a 17-residue peptide ( $\text{Arg}^{762}$ - $\text{Asn}^{778}$ ), representing this site, and CML19. The CML19-binding motif is well conserved among centrin-binding motifs in different species with three amino acids at position 1, 4, and 8 that constitute the major anchoring determinant (Fig. 8).

We found that CML19 binds one RAD4p per molecule, as for many CaM/target peptide complexes, in a  $\text{Ca}^{2+}$ -dependent manner. Indeed, the binding has an affinity of ~54 nM in the presence of  $\text{Ca}^{2+}$ , which is significantly reduced (~150-fold) in the absence of the ion, therefore reflecting  $\text{Ca}^{2+}$  dependence of RAD4p binding to CML19. The  $\text{Ca}^{2+}$ -dependent binding to CML19 also markedly changes Trp fluorescence of the RAD4 peptide, suggesting that Trp of the target becomes highly buried in a hydrophobic pocket of CML19, as for HsCen2 [61], CrCen [35], or Cdc31 [62]. This denotes



**Fig. 8.** Sequences of centrin-binding motifs. The bulky hydrophobic side chains anchoring the peptide into the hydrophobic pocket of the protein are in bold. Notably, the sequences of hSfi1 and melittin are disposed in a reverse orientation (8,4,1) as compared to the centrin-binding motif in XPC. AtRAD4 is the target of *Arabidopsis thaliana* CML19, XPC and hSfi1 are targets of human centrin, and Kar1 is a target of yeast centrin. Melittin is a model target peptide to study the interactions of  $\text{Ca}^{2+}$ -binding proteins with peptides.

a key role of the target Trp in the physical interaction with different centrin. The critical function of the Trp moiety in the hydrophobic triad W848, L851, and L855 was clearly highlighted by the significant decrease of binding affinity for HsCen2/XPC peptide complex upon  $\text{W}^1\text{A}$  substitution [31]. The  $\text{L}^4\text{A}$  substitution had a less dramatic effect, while substitution of  $\text{L}^8\text{A}$  did not cause any changes in binding affinity [31], suggesting that the hydrophobic dyad  $\text{W}^1\text{-L}^4$  represents the core of the centrin binding motif in XPC. Comparison with other centrin complexes [62,63] supports the primary importance of the  $\text{W}^1\text{-L}^4$  dyad, as divergences in the sequences of other targets are more pronounced in the case of  $\text{L}^8$ . Interestingly, the replacement of  $\text{W}^1$  in the homolog XPC abolishes its interaction with HsCen2 and prevents translocation of HsCen2 towards the nucleoplasm [31], therefore indicating that the same molecular mechanism is at the base of the physical binding to XPC and the nuclear translocation of HsCen2. Since CML19 was shown to be localized in the nucleus rapidly upon UV-C treatment [19], the major role of RAD4 may be in recruitment and re-localization of CML19 in the nucleus, similarly to the hypothesized re-localization of HsCen2 in the nucleoplasm driven by XPC [31].

In most centrin the target-binding motifs appear to be located primarily in the C-terminal EF-hands, contrary to most CaM/peptide complexes where two domains surround the target  $\alpha$ -helix. In accordance with this, the C-terminal EF-4 of CML19 has been found to be essential for the binding to RAD4 [19] and our NMR data clearly suggest that the N-terminal domain of the protein is not critically involved in formation of the CML19/RAD4p complex.

In general, it is believed that binding of  $\text{Ca}^{2+}$  to centrin promotes recognition of the target peptide by the protein, and that the peptide binding sites with low-affinity become activated in the presence of  $\text{Ca}^{2+}$  [64]. However, the interactions of EF-hand subdomains with  $\text{Ca}^{2+}$  and centrin binding partners appear to be complex and vary significantly among species [64]. While the C-terminal region of the centrin characterized to date exhibits more-conserved  $\text{Ca}^{2+}$  binding properties, the N-terminal



domains have undergone major modifications during evolution. For instance, while HsCen2 lost the ability to bind Ca<sup>2+</sup> in its N-terminal, the N-terminal portion of HsCen3 is still able to bind Ca<sup>2+</sup>, allowing for conformational changes, but only for transition from a molten globule to a compact state, without exposure of the hydrophobic region [43].

The presence of two high affinity Ca<sup>2+</sup>-specific sites in the N-terminal domain of CML19 and the finding that the N-terminal domain is not critically involved in the RAD4p binding give rise to the possibility, as for the green algal centrin CrCen and SdCen, that the N- and C-terminal domains of CML19 may interact simultaneously with different target proteins. Nonetheless, it is not known if additional binding motifs are present in natural targets of CML19. Clearly, investigation of the CML19 N- and C-terminal domains in terms of Ca<sup>2+</sup> binding and interaction with different types of natural peptide targets, including CaM-specific targets, as well as the experimental determination of the structure of Ca<sup>2+</sup>-CML19/RAD4p complex are an urgent priority to obtain deeper insights into the molecular basis for CML19 activity.

## Funding

This work was supported in part by grants from University of Verona (FUR2015) to AA and PD.

## Acknowledgments

We thank the “Centro Piattaforme Tecnologiche” of the University of Verona for providing access to the NMR spectrometer and nanoITC calorimeter.

## References

- [1] A. Perochon, D. Aldon, J.P. Galaud, B. Ranty, Calmodulin and calmodulin-like proteins in plant calcium signaling, *Biochimie* 93 (12) (2011) 2048–2053.
- [2] B. Ranty, D. Aldon, V. Cotellet, J.-P. Galaud, P. Thuleau, C. Mazars, Calcium sensors as key hubs in plant responses to biotic and abiotic stresses, *Front. Plant Sci.* 7 (2016) 327.
- [3] H. Zeng, L. Xu, A. Singh, H. Wang, L. Du, B.W. Poovaiah, Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses, *Front. Plant Sci.* 6 (2015) 600.
- [4] K.H. Edel, J. Kudla, Increasing complexity and versatility: how the calcium signaling toolkit was shaped during plant land colonization, *Cell Calcium* 57 (3) (2015) 231–246.
- [5] O. Batistic, J. Kudla, Analysis of calcium signaling pathways in plants, *Biochim. Biophys. Acta* 1820 (8) (2012) 1283–1293.
- [6] X. Zhu, C. Dunand, W. Snedden, J.-P. Galaud, CaM and CML emergence in the green lineage, *Trends Plant Sci.* 20 (8) (2015) 483–489.
- [7] J.L. Gifford, M.P. Walsh, H.J. Vogel, Structures and metal-ion-binding properties of the Ca<sup>2+</sup>-binding helix-loop-helix EF-hand motifs, *Biochem. J.* 405 (2) (2007) 199–221.
- [8] E.A. Permyakov, R.H. Kretsinger, Index, Calcium Binding Proteins, John Wiley & Sons, Inc., 2010, pp. 567–573.
- [9] K.W. Bender, W.A. Snedden, Calmodulin-related proteins step out from the shadow of their namesake, *Plant Physiol.* 163 (2) (2013) 486–495.
- [10] C. Cheval, D. Aldon, J.P. Galaud, B. Ranty, Calcium/calmodulin-mediated regulation of plant immunity, *Biochim. Biophys. Acta* 1833 (7) (2013) 1766–1771.
- [11] B.W. Poovaiah, L. Du, A. Singh, H. Wang, H. Zeng, L. Xu, Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses, *Front. Plant Sci.* 6 (2015).
- [12] B. Huang, A. Mengersen, V.D. Lee, Molecular cloning of cDNA for caltractin, a basal body-associated Ca<sup>2+</sup>-binding protein: homology in its protein sequence with calmodulin and the yeast CDC31 gene product, *J. Cell Biol.* 107 (1) (1988) 133–140.
- [13] J. Zhu, S.E. Bloom, E. Lazarides, C. Woods, Identification of a novel Ca(2+)-regulated protein that is associated with the marginal band and centrosomes of chicken erythrocytes, *J. Cell Sci.* 108 (Pt 2) (1995) 685–698.
- [14] V.D. Lee, B. Huang, Molecular cloning and centrosomal localization of human caltractin, *Proc. Natl. Acad. Sci. U. S. A.* 90 (23) (1993) 11039–11043.
- [15] L. Madeddu, C. Klotz, J.P. Le Caer, J. Beisson, Characterization of centrin genes in Paramecium, *Eur. J. Biochem.* 238 (1) (1996) 121–128.
- [16] J. Molinier, C. Ramos, O. Fritsch, B. Hohn, CENTRIN2 modulates homologous recombination and nucleotide excision repair in Arabidopsis, *Plant Cell* 16 (6) (2004) 1633–1643.
- [17] R. Nishi, W. Sakai, D. Tone, F. Hanaoka, K. Sugawara, Structure-function analysis of the EF-hand protein centrin-2 for its intracellular localization and nucleotide excision repair, *Nucleic Acids Res.* 41 (14) (2013) 6917–6929.
- [18] R. Nishi, Y. Okuda, E. Watanabe, T. Mori, S. Iwai, C. Masutani, K. Sugawara, F. Hanaoka, Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein, *Mol. Cell. Biol.* 25 (13) (2005) 5664–5674.
- [19] L. Liang, S. Flury, V. Kalck, B. Hohn, J. Molinier, CENTRIN2 interacts with the Arabidopsis homolog of the human XPC protein (AtRAD4) and contributes to efficient synthesis-dependent repair of bulky DNA lesions, *Plant Mol. Biol.* 61 (1–2) (2006) 345–356.
- [20] A. Astegno, M.C. Bonza, R. Vallone, V. La Verde, M. D’Onofrio, L. Luoni, B. Molesini, P. Dominici, Arabidopsis calmodulin-like protein CML36 is a calcium (Ca<sup>2+</sup>) sensor that interacts with the plasma membrane Ca<sup>2+</sup>-ATPase isoform ACA8 and stimulates its activity, *J. Biol. Chem.* 292 (36) (2017) 15049–15061.
- [21] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [22] A. Astegno, V. La Verde, V. Marino, D. Dell’Orco, P. Dominici, Biochemical and biophysical characterization of a plant calmodulin: role of the N- and C-lobes in calcium binding, conformational change, and target interaction, *Biochim. Biophys. Acta* 1864 (3) (2016) 297–307.
- [23] R. Vallone, V. La Verde, M. D’Onofrio, A. Giorgetti, P. Dominici, A. Astegno, Metal binding affinity and structural properties of calmodulin-like protein 14 from Arabidopsis thaliana, *Protein Sci.* 25 (8) (2016) 1461–1471.
- [24] A. Astegno, G. Capitani, P. Dominici, Functional roles of the hexamer organization of plant glutamate decarboxylase, *Biochim. Biophys. Acta, Proteins Proteomics* 1854 (9) (2015) 1229–1237.
- [25] V. La Verde, P. Dominici, A. Astegno, Determination of hydrodynamic radius of proteins by size exclusion chromatography, *Bio-protocol* 7 (8) (2017) (4/20/2017).
- [26] V. Marino, A. Astegno, M. Pedroni, F. Piccinelli, D. Dell’Orco, Nanodevice-induced conformational and functional changes in a prototypical calcium sensor protein, *Nanoscale* 6 (1) (2014) 412–423.
- [27] A. Allegrini, A. Astegno, V. La Verde, P. Dominici, Characterization of C-S lyase from *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC BAA-365 and its potential role in food flavour applications, *J. Biochem.* 161 (4) (2017) 349–360.
- [28] A. Astegno, A. Giorgetti, A. Allegrini, B. Cellini, P. Dominici, Characterization of C-S Lyase from *C. diphtheriae*: a possible target for new antimicrobial drugs, *BioMed Res. Int.* 2013 (2013) 701536.
- [29] A. Astegno, E. Maresi, M. Bertoldi, V. La Verde, A. Paiardini, P. Dominici, Unique substrate specificity of ornithine aminotransferase from *Toxoplasma gondii*, *Biochem. J.* 474 (6) (2017) 939–955.
- [30] J.R. Thompson, Z.C. Ryan, J.L. Salisbury, R. Kumar, The structure of the human centrin 2-xeroderma pigmentosum group C protein complex, *J. Biol. Chem.* 281 (27) (2006) 18746–18752.
- [31] J.B. Charbonnier, E. Renaud, S. Miron, M.H. Le Du, Y. Blouquit, P. Duchambon, P. Christova, A. Shosheva, T. Rose, J.F. Angulo, C.T. Craescu, Structural, thermodynamic, and cellular characterization of human centrin 2 interaction with xeroderma pigmentosum group C protein, *J. Mol. Biol.* 373 (4) (2007) 1032–1046.
- [32] L.P. Kozlowski, J.M. Bujnicki, MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins, *BMC Bioinf.* 13 (1) (2012) 1–11.
- [33] B.M. Waters, Moving magnesium in plant cells, *New Phytol.* 190 (3) (2011) 510–513.
- [34] F.J. Maathuis, Physiological functions of mineral macronutrients, *Curr. Opin. Plant Biol.* 12 (3) (2009) 250–258.
- [35] H. Hu, J.H. Sheehan, W.J. Chazin, The mode of action of centrin: Binding of Ca<sup>2+</sup> and a peptide fragment of kar1p to the C-terminal domain, *J. Biol. Chem.* 279 (49) (2004) 50895–50903.
- [36] S. Mukherjee, P.M. Mohan, K.V. Chary, Magnesium promotes structural integrity and conformational switching action of a calcium sensor protein, *Biochemistry* 46 (12) (2007) 3835–3845.
- [37] P. Aravind, K. Chandra, P.P. Reddy, A. Jeromin, K.V.R. Chary, Y. Sharma, Regulatory and structural EF-hand motifs of neuronal calcium sensor-1: Mg<sup>2+</sup> modulates Ca<sup>2+</sup> binding, Ca<sup>2+</sup>-induced conformational changes, and equilibrium unfolding transitions, *J. Mol. Biol.* 376 (4) (2008) 1100–1115.
- [38] J.L. Gifford, M. Jamsheidiha, J. Mo, H. Ishida, H.J. Vogel, Comparing the calcium binding abilities of two soybean calmodulins: towards understanding the divergent nature of plant calmodulins, *Plant Cell* 25 (11) (2013) 4512–4524.
- [39] A.P. Yarniuk, L.T. Nguyen, T.T. Hoang, H.J. Vogel, Metal ion binding properties and conformational states of calcium- and integrin-binding protein, *Biochemistry* 43 (9) (2004) 2558–2568.
- [40] J.N. Wingard, J. Chan, I. Bosanac, F. Haeseleer, K. Palczewski, M. Ikura, J.B. Ames, Structural analysis of Mg<sup>2+</sup> and Ca<sup>2+</sup> binding to CaBP1, a neuron-specific regulator of calcium channels, *J. Biol. Chem.* 280 (45) (2005) 37461–37470.
- [41] A. Yang, S. Miron, P. Duchambon, L. Assairi, Y. Blouquit, C.T. Craescu, The N-terminal domain of human centrin 2 has a closed structure, binds calcium with a very low affinity, and plays a role in the protein self-assembly, *Biochemistry* 45 (3) (2006) 880–889.
- [42] H. Wiech, B.M. Geier, T. Paschke, A. Spang, K. Grein, J. Steinkotter, M. Melkonian, E. Schiebel, Characterization of green alga, yeast, and human centrin. Specific subdomain features determine functional diversity, *J. Biol. Chem.* 271 (37) (1996) 22453–22461.

- [43] J.A. Cox, F. Tirone, I. Durussel, C. Firanesco, Y. Blouquit, P. Duchambon, C.T. Craescu, Calcium and magnesium binding to human centrin 3 and interaction with target peptides, *Biochemistry* 44 (3) (2005) 840–850.
- [44] L. Radu, I. Durussel, L. Assairi, Y. Blouquit, S. Miron, J.A. Cox, C.T. Craescu, *Scherffelia dubia* centrin exhibits a specific mechanism for Ca(2+)-controlled target binding, *Biochemistry* 49 (20) (2010) 4383–4394.
- [45] S.M. Gagne, S. Tsuda, M.X. Li, M. Chandra, L.B. Smillie, B.D. Sykes, Quantification of the calcium-induced secondary structural changes in the regulatory domain of troponin-C, *Protein Sci.* 3 (11) (1994) 1961–1974.
- [46] B.R. Sorensen, J.T. Eppel, M.A. Shea, *Paramecium* calmodulin mutants defective in ion channel regulation associate with melittin in the absence of calcium but require it for tertiary collapse, *Biochemistry* 40 (4) (2001) 896–903.
- [47] B.R. Sorensen, M.A. Shea, Calcium binding decreases the stokes radius of calmodulin and mutants R74A, R90A, and R90G, *Biophys. J.* 71 (6) (1996) 3407–3420.
- [48] M. Garrigos, S. Deschamps, A. Viel, S. Lund, P. Champeil, J.V. Moller, M. le Maire, Detection of Ca(2+)-binding proteins by electrophoretic migration in the presence of Ca2+ combined with 45Ca2+ overlay of protein blots, *Anal. Biochem.* 194 (1) (1991) 82–88.
- [49] E. de Castro, C.J. Sigrist, A. Gattiker, V. Bulliard, P.S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch, N. Hulo, ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins, *Nucleic Acids Res.* 34 (Web Server issue) (2006) W362–5.
- [50] A. Popescu, S. Miron, Y. Blouquit, P. Duchambon, P. Christova, C.T. Craescu, Xeroderma pigmentosum group C protein possesses a high affinity binding site to human centrin 2 and calmodulin, *J. Biol. Chem.* 278 (41) (2003) 40252–40261.
- [51] M. Ikura, G.M. Clore, A.M. Gronenborn, G. Zhu, C.B. Klee, A. Bax, Solution structure of a calmodulin–target peptide complex by multidimensional NMR, *Science* 256 (5057) (1992) 632–638.
- [52] W.E. Meador, A.R. Means, F.A. Quijcho, Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin–peptide complex, *Science* 257 (5074) (1992) 1251–1255.
- [53] S.M. Roth, D.M. Schneider, L.A. Strobel, M.F. Van Berkum, A.R. Means, A.J. Wand, Characterization of the secondary structure of calmodulin in complex with a calmodulin-binding domain peptide, *Biochemistry* 31 (5) (1992) 1443–1451.
- [54] A. Astegno, E. Maresi, V. Marino, P. Dominici, M. Pedroni, F. Piccinelli, D. Dell’Orco, Structural plasticity of calmodulin on the surface of CaF2 nanoparticles preserves its biological function, *Nanoscale* 6 (24) (2014) 15037–15047.
- [55] S.C. Popescu, G.V. Popescu, S. Bachan, Z. Zhang, M. Seay, M. Gerstein, M. Snyder, S.P. Dinesh-Kumar, Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays, *Proc. Natl. Acad. Sci. U. S. A.* 104 (11) (2007) 4730–4735.
- [56] K.P. Hoeflich, M. Ikura, Calmodulin in action: diversity in target recognition and activation mechanisms, *Cell* 108 (6) (2002) 739–742.
- [57] E. Matei, S. Miron, Y. Blouquit, P. Duchambon, I. Durussel, J.A. Cox, C.T. Craescu, C-terminal half of human centrin 2 behaves like a regulatory EF-hand domain, *Biochemistry* 42 (6) (2003) 1439–1450.
- [58] I. Durussel, Y. Blouquit, S. Middendorp, C.T. Craescu, J.A. Cox, Cation- and peptide-binding properties of human centrin 2, *FEBS Lett.* 472 (2–3) (2000) 208–212.
- [59] S. Veeraraghavan, P.A. Fagan, H. Hu, V. Lee, J.F. Harper, B. Huang, W.J. Chazin, Structural independence of the two EF-hand domains of caltractin, *J. Biol. Chem.* 277 (32) (2002) 28564–28571.
- [60] H. Ishida, H.J. Vogel, Protein–peptide interaction studies demonstrate the versatility of calmodulin target protein binding, *Protein Pept. Lett.* 13 (5) (2006) 455–465.
- [61] A. Yang, S. Miron, L. Mouawad, P. Duchambon, Y. Blouquit, C.T. Craescu, Flexibility and plasticity of human centrin 2 binding to the xeroderma pigmentosum group C protein (XPC) from nuclear excision repair, *Biochemistry* 45 (11) (2006) 3653–3663.
- [62] S. Li, A.M. Sandercock, P. Conduit, C.V. Robinson, R.L. Williams, J.V. Kilmartin, Structural role of Sfi1p-centrin filaments in budding yeast spindle pole body duplication, *J. Cell Biol.* 173 (6) (2006) 867–877.
- [63] H. Hu, W.J. Chazin, Unique features in the C-terminal domain provide caltractin with target specificity, *J. Mol. Biol.* 330 (3) (2003) 473–484.
- [64] Y. Zhang, C.Y. He, Centrins in unicellular organisms: functional diversity and specialization, *Protoplasma* 249 (3) (2012) 459–467.

# A Method for Metal/Protein Stoichiometry Determination Using Thin-Film Energy Dispersive X-ray Fluorescence Spectroscopy

Silvia Fruncillo,<sup>†,‡,§</sup> Matteo Trande,<sup>§</sup> Christopher F. Blanford,<sup>†,||</sup> Alessandra Astegno,<sup>§,||</sup> and Lu Shin Wong<sup>\*,†,‡,§,||</sup>

<sup>†</sup>Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, United Kingdom

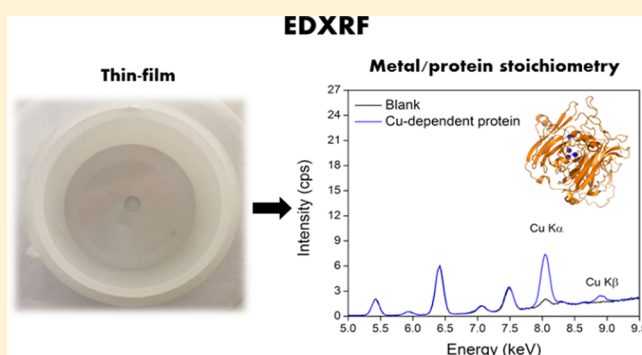
<sup>‡</sup>Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

<sup>§</sup>Department of Biotechnology, University of Verona, Strada Le Grazie 15, Verona 37134, Italy

<sup>||</sup>Department of Materials, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

## Supporting Information

**ABSTRACT:** A convenient approach is proposed for the quantitation of elemental cofactors in proteins and the determination of metal/protein stoichiometry, on the basis of energy dispersive X-ray fluorescence spectroscopy (EDXRF). The analysis of proteins containing the metals Cu, Fe, Zn, and Ca and also the nonmetallic element P is shown as a demonstration of the generality of the method. In general, the reported method gives limit of detection (LOD) and limit of quantification (LOQ) values in the low ppm range and requires only a few microliters of protein sample at micromolar concentrations. Moreover, sample preparation does not require any digestion steps before the analysis. The expected metal/protein stoichiometry was observed for each protein analyzed, highlighting the precision and accuracy of the method in all the tested cases. Furthermore, it is shown that the method is compatible with multimeric proteins and those with post-translational modifications such as glycosylation.



Many proteins contain metal cofactors that are crucial for their structure or biochemical function. It is estimated that approximately one-quarter to one-third of all proteins require metals for their function.<sup>1</sup> Throughout nature, there is a large variation in the type of metal present, from the relatively small alkali earth metals (e.g., Ca and Mg) to heavier transition metals (e.g., Fe and Cu), as well as their stoichiometry within the protein. There is therefore a demand for convenient methods to identify and quantify the inorganic species that are present in proteins.

The most commonly used techniques for metal cofactor identification and quantification are derived from inorganic chemical analysis. They include inductively coupled plasma mass spectrometry (ICP-MS), which typically gives a limit of detection (LOD) at ppb levels; inductively coupled plasma optical emission spectrometry (ICP-OES), with a LOD at the ppm level; and atomic absorption spectroscopy (AAS), with a LOD between that of ICP-MS and ICP-OES.<sup>2,3</sup> The concentration of the element of interest can then be correlated to the protein concentration (typically measured by a separate method) to give the metal/protein stoichiometry. However, despite their suitability for the study of metallic elements and the excellent detection limits, these techniques require sample predigestion and a relatively large amount of analyte (typically mL volumes at the relevant concentration). Protein

stoichiometry can also be determined by whole protein (“native”) mass spectrometry (MS), though this method does not typically give absolute quantification and requires access of MS instrumentation that is not always readily available.<sup>4,5</sup>

In the area of materials engineering, X-ray fluorescence spectroscopy (XRF) is a commonly used technique for the elemental analysis of inorganic materials.<sup>6</sup> The principle of XRF is based upon the detection of X-rays emitted from sample atoms upon excitation, typically by irradiation with X-rays of higher energy. The emitted X-rays are of energies that are characteristic for each element, therefore enabling the identification and quantification of the atomic composition of the sample. XRF-based methods can be applied to a wide range of samples, including solids, powders, liquids, and suspensions, and does not require any sample pretreatment prior to analysis.

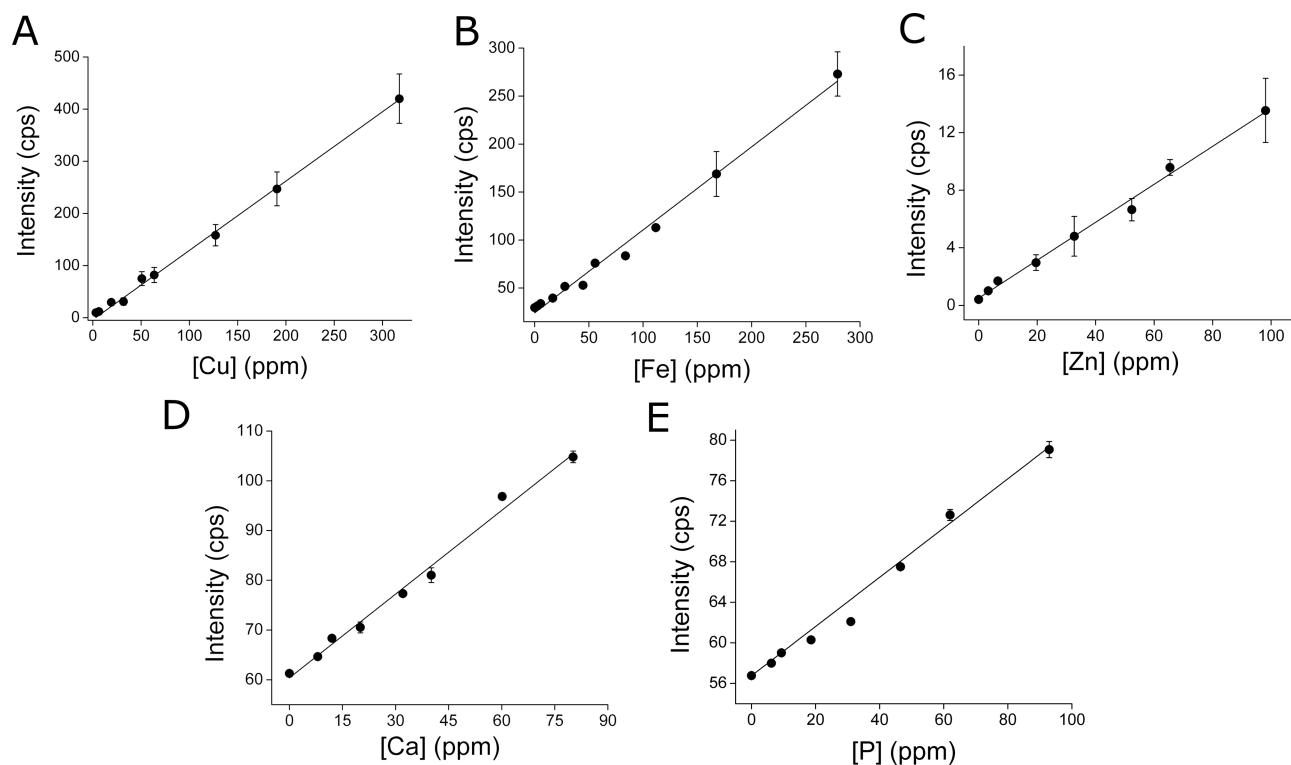
Total-reflection XRF (TXRF) is an XRF-based technique that has been used for the qualitative and quantitative analysis of metal cofactors in proteins.<sup>7–9</sup> It is able to provide simultaneous analysis of several elements with LOD values in the ppm–ppb range, in short analysis times (typically 2–30

Received: July 22, 2019

Accepted: August 19, 2019

Published: August 19, 2019





**Figure 1.** Calibration plots of EDXRF peak area against element concentration for (A) Cu at 0–317 ppm, (B) Fe at 0–279 ppm, (C) Zn at 0–98 ppm, (D) Ca at 0–80 ppm, and (E) P at 0–93 ppm. Error bars represent the standard error of the mean (SEM).

min per analysis), and with only a small amount of sample (a few  $\mu\text{g}$  or  $\mu\text{L}$ ). A related technique, microfocused particle induced X-ray emission (microPIXE), is also able to give a similar performance, though it requires highly specialized equipment that is not widely available.<sup>10</sup>

An XRF technique that is more widely available is energy dispersive X-ray fluorescence (EDXRF). EDXRF has higher detection limits than TXRF and microPIXE but employs a less complex equipment design that has much lower instrumentation and maintenance costs, while retaining the characteristics of X-ray spectroscopy in terms of the qualitative and quantitative analysis of nearly all elements, in any physical form, as well as the simplicity of sample preparation and operation.<sup>11,12</sup> More recently, advances in EDXRF equipment design (e.g., the introduction of filters to reduce background signals, the ability to tailor voltage and current settings to improve the sensitivity to specific elements, and the use of silicon drift detectors) have given improved signal-to-noise ratios, spectral resolution, throughput, and reliability.<sup>6</sup>

By adapting the general principles of protein thin-film TXRF, we report herein a generally applicable method that employs EDXRF for the rapid, convenient, and affordable quantitation of inorganic cofactors in proteins and for accurate determination of element/protein stoichiometry within a protein complex. The method is demonstrated on a range of proteins containing Cu, Fe, Zn, Ca, and the biologically relevant nonmetal P. These proteins include examples of prokaryotic and eukaryotic origin and multimeric proteins as well as proteins bearing post-translational modifications.

## RESULTS AND DISCUSSION

**Experimental Design and Instrument Calibration.** It was envisaged that by employing thin-film analysis it would be

possible to minimize the amount of sample material required and avoid distortions of the spectra associated with thick samples that could lead to less accurate quantification.<sup>13</sup> To further maximize the signal-to-noise ratio, the analyte solutions were made up in a TRIS-acetate buffer, which had previously been reported to give good results with TXRF.<sup>7</sup> This buffer is particularly suitable as it contains no elements with an atomic number  $>11$ , thus avoiding the introduction of any peaks that may overlap with the signal of interest and minimizing any intra-element absorption or enhancement effects.

For calibration, thin-film standards for each element were produced from salt solutions of known concentrations by drying an 8  $\mu\text{L}$  droplet of the solution. EDXRF spectra were then recorded, and calibration curves were obtained by plotting the peak area versus the original concentration of the solution (Figure 1). These plots gave good linear correlations, indicating that film thickness or small variations in the size of the film had a negligible effect on the quantification.<sup>13</sup> Since very low concentrations of each element were analyzed, the spectra were recorded for 1600 s in order to maximize the LOD while maintaining a reasonable duration for the proposed analysis. The calculated LOD and limit of quantification (LOQ) values were found to be in the low ppm range (Table 1), equivalent to nmol–pmol quantities of the element.

The light elements Ca and P emit photons of relatively low energy that are poorly transmitted through air;<sup>2</sup> thus to mitigate this effect, the sample chamber was purged with helium prior to recording the spectra. Even so, these elements showed the highest LOD and LOQ values. Moreover, P, Ca, and Fe generally showed the highest levels of background signal compared to all the other tested elements (also evidenced by the high baseline in the calibration fit), likely

**Table 1. Best Linear Fit Calibrations for Each Element**

element	linear best fit equation <sup>a</sup>	R <sup>2</sup> <sup>b</sup>	LOD (ppm)	LOQ (ppm)
Cu	$I_{\text{Cu}} = 1.31[\text{Cu}] + 0.71$	0.997	1.1	3.5
Fe	$I_{\text{Fe}} = 0.86[\text{Fe}] + 24.16$	0.991	0.9	3.2
Zn	$I_{\text{Zn}} = 0.13[\text{Zn}] + 0.50$	0.993	0.8	2.7
Ca	$I_{\text{Ca}} = 0.56[\text{Ca}] + 60.39$	0.989	4.4	14.6
P	$I_{\text{P}} = 0.25[\text{P}] + 56.21$	0.986	5.4	18.8

<sup>a</sup> $I_{\text{element}}$  represents the signal intensity in the EDXRF spectrum for the element (in cps); [element] is the concentration of the element (in ppm). <sup>b</sup>R<sup>2</sup> is the correlation coefficient for a linear regression.

due to their presence as common contaminants in Mylar films and other materials commonly used during EDXRF analysis such as polypropylene and polyester films. Nevertheless, these data indicate that the LOD and LOQ are sufficiently low to allow for the proposed protein analysis.

**Protein Analysis.** Elemental analysis was undertaken on 10 proteins containing Cu, Fe, Zn, Ca, and the nonmetallic element P (Table 2). Thin films were prepared from protein

**Table 2. Element/Protein Stoichiometry**

entry	protein <sup>a</sup>	element	range of protein concentrations tested (μM)	obtained E/P <sup>b</sup> (±SEM)	expected E/P
1	CotA	Cu		3.97 ± 0.10	
2	MvBOx	Cu	35–120	4.13 ± 0.12	4 <sup>14,15</sup>
3	CytC	Fe		1.03 ± 0.05	
4	AHb1	Fe	200–500	0.98 ± 0.04	1 <sup>16–18</sup>
5	HRP	Fe		0.99 ± 0.05	
6	ADH1	Zn	35–140	2.18 ± 0.11	2 <sup>19</sup>
7	CaM1	Ca		4.01 ± 0.20	
8	CML7	Ca	140–400	3.85 ± 0.08	4 <sup>20–22</sup>
9	CML19	Ca		3.60 ± 0.12	
10	HRP	Ca	250–500	1.85 ± 0.12	2 <sup>17</sup>
11	UbiX	P	600–800	1.30 ± 0.20	1 <sup>23</sup>

<sup>a</sup>Full name and origin of each protein is given in the SI. <sup>b</sup>E/P represents the element/protein stoichiometry. The mean values ± SEM from at least six independent samples are presented.

solutions of known concentration, and their EDXRF spectra were recorded (Figure 2). By applying the calibration line equation in Table 1, the concentration of the metal in the original analyte solution was determined. Subsequently, using the previously measured protein concentration (expressed as the concentration of the monomer), the element/protein stoichiometry could then be calculated (eq S1).

For all the proteins analyzed, the ratios of the element to protein were in good agreement with the values obtained through other techniques (see refs in Table 2). For robust stoichiometry determination, it was found that the concentration of the metal should ideally lie between the LOQ and the highest calibrant concentration used in the construction of the calibration curve. Thus, heavier elements and proteins with multiple binding sites required lower amounts of protein for successful analysis.

The stoichiometry values obtained by this method are not dependent on the quaternary structure of the protein since the calculations rely only on the protein monomer concentration. For example, AHb1 forms a homodimer<sup>18,24</sup> with one Fe atom per monomer, and this method gives a metal/monomer ratio of 1:1 (Entry 4, Table 2), while ADH1 forms a homotetramer<sup>19</sup> with two Zn atoms per monomer, giving a

metal/monomer ratio of 2:1 (Entry 6, Table 2). The analysis of individual elements was also not affected by the presence of other elements or post-translational modifications. For example, commercial batches of eukaryotic proteins CytC and HRP include varying degrees of acetylation and glycosylation, respectively,<sup>16,17,25</sup> but analyses of both proteins gave the expected Fe stoichiometry of 1:1 (Entries 3 and 5, Table 2). Simultaneous quantification of multiple elements from a single protein is possible, as for HRP, which contains both a catalytic Fe and two structural Ca atoms (Entries 5 and 10, Table 2).

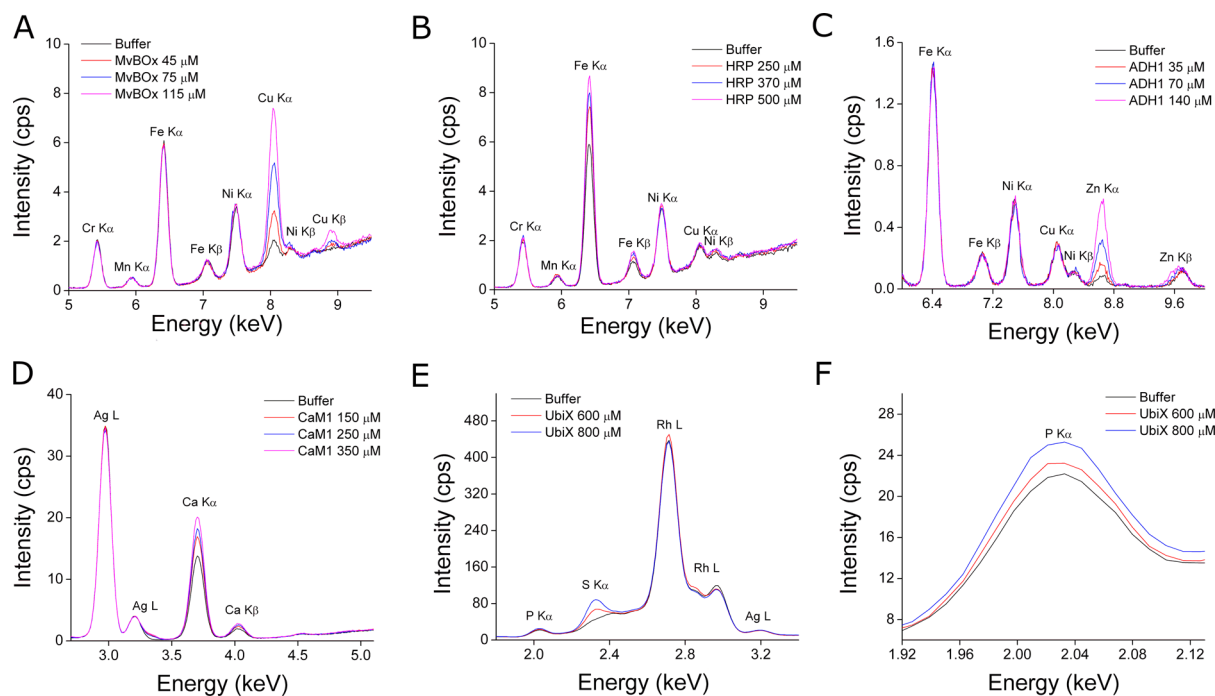
EDXRF can also be used for the quantification of nonmetal elements, provided the photon energies are within the detectable range. As an example, the protein UbiX contains a single riboflavin-5'-phosphate (flavin mononucleotide, FMN) cofactor, which itself contains a single P atom. Here, the EDXRF-determined P/protein ratio was consistent with previous literature,<sup>23</sup> although the SEM was relatively large due to the low sensitivity of the instrument toward light elements (Figure 2E). In comparison, elements such as Fe and Cu gave appreciable differences in spectral intensities even with differences in the concentration of the analyte solutions as low as 5 μM. Nevertheless, this method can be used to confirm or exclude the presence of cofactor binding in proteins of interest. In the case of CytC, this protein has two possible phosphorylation sites,<sup>25</sup> but no signal for P was detectable above the background even at the highest protein concentration that was analyzed, thus indicating that the CytC used in this study was in fact not phosphorylated.

Protein quantification represents a crucial component of the stoichiometry calculations. In theory, S can be detected and quantified by EDXRF, so it could be used for simultaneous protein quantification by measuring the S present in Cys and Met residues of the protein. However, because S is a relatively light element giving only weak signals and overlapping signals are present in the spectral region between 2.2 and 2.6 keV caused by the Rh anode, quantification would only be possible when large amounts of S were present. For example, in the case of UbiX that has 14 S atoms, at a protein concentration of 600 μM (equating to an S concentration of 269 ppm), it was clearly possible to detect the presence of S (Figure 2E). However, the LOD was extremely high (estimated at >180 ppm), making simultaneous analysis with other cofactors inconvenient at best and impossible in cases where this value was higher than the solubility of the protein of interest.

## CONCLUSIONS

In the present work, a simple and accurate thin-film EDXRF method has been developed and validated for the determination of metal stoichiometry within protein complexes. This approach was demonstrated with a range of proteins containing Cu, Fe, Zn, and Ca as well as with one protein containing the nonmetal P. The method gives LOD values between 1 and 6 ppm and LOQ values between 3 and 20 ppm with a minimal consumption of sample (pmol) and does not require predigestion prior to analysis. It is generalizable across proteins of different molecular weights, oligomerization states, and types and numbers of bound metals as well as post-translational modifications (e.g., glycosylation and acetylation). The quantitation of multiple elements within a single protein is also possible.

Though only a selected range of biologically relevant elements was investigated using existing EDXRF equipment,



**Figure 2.** Illustrative EDXRF spectra of protein samples at different concentrations. Overlaid spectra (A) of Cu-dependent MvBOx, (B) of Fe-dependent HRP, (C) of Zn-dependent ADH1, (D) of Ca-dependent CaM1, (E) of P-containing UbiX, and (F) showing a magnification of the P peak from the spectra shown in (E). The Rh signal observed in (E) originates from the Rh anode of the instrument. The spectral acquisition settings used for each element are indicated in Table S2.

in principle, this type of spectroscopy allows the quantification of any element between Na and U, provided the appropriate instrument setup is available. By combining the fact that only minimal sample preparation is required and EDXRF instruments are relatively widely available, this method represents a competitive alternative to other methods for the quantitation of elements in proteins. It is intended that the method reported here can be employed for the analysis of novel synthetic proteins containing transition metals such as Rh or Ru,<sup>26</sup> the validation of currently uncharacterized metalloproteins where the metal binding capability is only predicted through sequence-based bioinformatic analysis, or the quality assurance of recombinantly produced metalloproteins.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.9b03319.

Experimental methodology; Table S1: characteristics of proteins analyzed with EDXRF; Table S2: EDXRF instrument settings (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Tel: +44 161 306 8939. E-mail: l.s.wong@manchester.ac.uk.

### ORCID

Silvia Fruncillo: 0000-0003-4953-8349

Christopher F. Blanford: 0000-0002-0112-7818

Alessandra Astegno: 0000-0002-7341-0970

Lu Shin Wong: 0000-0002-7437-123X

## Author Contributions

S.F. and L.S.W. designed the research. S.F. and M.T. performed the experiments. S.F., M.T., A.A., and L.S.W. analyzed the data. S.F., M.T., C.F.B., A.A. and L.S.W. wrote and edited the manuscript.

## Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

S.F. acknowledges the University of Manchester (UK) and the Agency for Science and Technology Research (Singapore) for an A\*STAR Research Attachment Programme PhD studentship. C.F.B. and L.S.W. acknowledge funding from the N8 Agrifood Programme (UK). M.T. and A.A. acknowledge funding from the University of Verona (Italy) (FUR 2018 to A.A.). The authors acknowledge the use of the Department of Materials X-ray Diffraction Suite at the University of Manchester and the technical support, advice, and assistance provided by Dr. John E. Warren.

## ■ REFERENCES

- (1) Waldron, K. J.; Robinson, N. J. *Nat. Rev. Microbiol.* **2009**, *7*, 25–35.
- (2) Rio-Echevarria, I. M.; Ponti, J.; Bogno, A.; Gilliland, D.; Altissimo, M.; Pascolo, L.; Ceccone, G.; Gianoncelli, A. *X-Ray Spectrom.* **2019**, *48*, 94–101.
- (3) Garcia, J. S.; de Magalhaes, C. S.; Arruda, M. A. Z. *Talanta* **2006**, *69*, 1–15.
- (4) Ro, S. Y.; Schachner, L. F.; Koo, C. W.; Purohit, R.; Remis, J. P.; Kenney, G. E.; Liauw, B. W.; Thomas, P. M.; Patrie, S. M.; Kelleher, N. L.; Rosenzweig, A. C. *Nat. Commun.* **2019**, *10*, 2675.
- (5) Wongkongkathep, P.; Han, J. Y.; Choi, T. S.; Yin, S.; Kim, H. I.; Loo, J. A. J. *Am. Soc. Mass Spectrom.* **2018**, *29*, 1870–1880.
- (6) Margui, E.; Zawisza, B.; Sitko, R. *TrAC, Trends Anal. Chem.* **2014**, *53*, 73–83.

- (7) Szoboszlai, N.; Polgari, Z.; Mihucz, V. G.; Zaray, G. *Anal. Chim. Acta* **2009**, *633*, 1–18.
- (8) Mertens, M.; Rittmeyer, C.; Kolbesen, B. O. *Spectrochim. Acta, Part B* **2001**, *56*, 2157–2164.
- (9) Wittershagen, A.; Rostam-Khani, P.; Zickermann, V.; Zickermann, I.; Gemeinhardt, S.; Ludwig, B.; Kolbesen, B. O. *Fresenius' J. Anal. Chem.* **1998**, *361*, 326–328.
- (10) Garman, E. F.; Grime, G. W. *Prog. Biophys. Mol. Biol.* **2005**, *89*, 173–205.
- (11) Viksna, A.; Lindgren, E. S.; Standzenieks, P.; Jacobsson, J. *X-Ray Spectrom.* **2004**, *33*, 414–420.
- (12) Fiedler, H. D.; Drinkel, E. E.; Orzechowicz, B.; Leopoldino, E. C.; Souza, F. D.; Almerindo, G. I.; Perdona, C.; Nome, F. *Anal. Chem.* **2013**, *85*, 10142–10148.
- (13) Markowicz, A. *Pramana* **2011**, *76*, 321–329.
- (14) Cracknell, J. A.; McNamara, T. P.; Lowe, E. D.; Blanford, C. F. *Dalton Trans.* **2011**, *40*, 6668–6675.
- (15) Durao, P.; Chen, Z.; Fernandes, A. T.; Hildebrandt, P.; Murgida, D. H.; Todorovic, S.; Pereira, M. M.; Melo, E. P.; Martins, L. O. *JBIC, J. Biol. Inorg. Chem.* **2008**, *13*, 183–193.
- (16) Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 585–595.
- (17) Gajhede, M.; Schuller, D. J.; Henriksen, A.; Smith, A. T.; Poulos, T. L. *Nat. Struct. Biol.* **1997**, *4*, 1032–1038.
- (18) Kumar, N.; Astegno, A.; Chen, J.; Giorgetti, A.; Dominici, P. *Int. J. Mol. Sci.* **2016**, *17*, 640.
- (19) Raj, S. B.; Ramaswamy, S.; Plapp, B. V. *Biochemistry* **2014**, *53*, 5791–5803.
- (20) Astegno, A.; La Verde, V.; Marino, V.; Dell'Orco, D.; Dominici, P. *Biochim. Biophys. Acta, Proteins Proteomics* **2016**, *1864*, 297–307.
- (21) La Verde, V.; Trande, M.; D'Onofrio, M.; Dominici, P.; Astegno, A. *Int. J. Biol. Macromol.* **2018**, *108*, 1289–1299.
- (22) Trande, M.; Pedretti, M.; Bonza, M. C.; Di Matteo, A.; D'Onofrio, M.; Dominici, P.; Astegno, A. *J. Inorg. Biochem.* **2019**, *199*, 110796.
- (23) White, M. D.; Payne, K. A. P.; Fisher, K.; Marshall, S. A.; Parker, D.; Rattray, N. J. W.; Trivedi, D. K.; Goodacre, R.; Rigby, S. E. J.; Scrutton, N. S.; Hay, S.; Leys, D. *Nature* **2015**, *522*, 502.
- (24) Mukhi, N.; Dhindwal, S.; Uppal, S.; Kumar, P.; Kaur, J.; Kundu, S. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 1944–1956.
- (25) Kalpage, H. A.; Bazylanska, V.; Recanati, M. A.; Fite, A.; Liu, J.; Wan, J. M.; Mantena, N.; Malek, M. H.; Podgorski, I.; Heath, E. I.; Vaishnav, A.; Edwards, B. F.; Grossman, L. I.; Sanderson, T. H.; Lee, I.; Huttemann, M. *FASEB J.* **2019**, *33*, 1540–1553.
- (26) Jarvis, A. G.; Obrecht, L.; Deuss, P. J.; Laan, W.; Gibson, E. K.; Wells, P. P.; Kamer, P. C. J. *Angew. Chem., Int. Ed.* **2017**, *56*, 13596–13600.