

Georgia State University

ScholarWorks @ Georgia State University

Chemistry Dissertations

Department of Chemistry

11-28-2007

Exploring the Role of Calcium Ions in Biological Systems by Computational Prediction and Protein Engineering

Yubin Zhou

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_diss

 Part of the [Chemistry Commons](#)

Recommended Citation

Zhou, Yubin, "Exploring the Role of Calcium Ions in Biological Systems by Computational Prediction and Protein Engineering." Dissertation, Georgia State University, 2007.

doi: <https://doi.org/10.57709/1059260>

This Dissertation is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

Exploring the Role of Calcium Ions in Biological Systems by Computational
Prediction and Protein Engineering

by

Yubin Zhou

Under the Direction of Professor Jenny J. Yang

ABSTRACT

Ca^{2+} , a “signal for death and life”, is closely involved in the regulation of numerous important cellular events. Ca^{2+} carries out its function through its binding to Ca^{2+} -receptors or Ca^{2+} -binding proteins. The EF-hand protein, with a helix-loop-helix Ca^{2+} -binding motif, constitutes one of the largest protein families. To facilitate our understanding of the role of Ca^{2+} in biological systems (denoted as “calciomics”) using genomic information, an improved pattern search method (<http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm>) for the identification of EF-hand and EF-like Ca^{2+} -binding proteins was developed. This fast and robust method allows us to analyze putative EF-hand proteins at the genome-wide level and further visualize the evolutionary scenario of the EF-hand protein family.

This prediction method further enables us to locate a putative viral EF-hand Ca^{2+} -binding motif within the rubella virus nonstructural protease that cleaves the nonstructural protein precursor into two active replicase components. A novel grafting approach has been used to probe the metal-binding properties of

this motif by engineering the predicted 12-residue Ca^{2+} -coordinating loop into a non- Ca^{2+} -binding scaffold protein, CD2 domain 1. Structural and conformational studies were further performed on a purified, bacterially-expressed NS protease minimal metal-binding domain spanning the EF-hand Ca^{2+} -binding motif. It was revealed that Ca^{2+} binding induced local conformational changes and increased thermal stability. Furthermore, functional studies were carried out using RUB infectious cDNA clone and replicon constructs. Our studies have shown that the Ca^{2+} binding loop played a structural role in the NS protease and was specifically required for optimal stability under physiological conditions.

In addition, we have predicted and characterized a calmodulin-binding domain in the gap junction proteins connexin43 and connexin44. Peptides encompassing the CaM binding motifs were synthesized and their ability to bind CaM was determined using various biophysical approaches. Transient expression in HeLa cells of two mutant Cx43-EYFP constructs without the putative CaM-binding site eliminated the Ca^{2+} -dependent inhibition of gap junction permeability. These results provide the first direct evidence that CaM binds to a specific region of the ubiquitous gap junction protein Cx43 and Cx44 in a Ca^{2+} -dependent manner, providing a molecular basis for the well-characterized Ca^{2+} -dependent inhibition of Cx43-containing gap junctions.

KEY WORDS: Calcium, EF-hand, Calmodulin, Prediction, CD2, Rubella virus, Protein engineering, Pattern search, Gap junction, Connexin

EXPLORING THE ROLE OF CALCIUM IONS IN BIOLOGICAL SYSTEMS BY
COMPUTATIONAL PREDICTION AND PROTEIN ENGINEERING

by

Yubin Zhou

A Dissertation Submitted in Partial Fulfillment of the Requirement of the Degree
of

Doctor of Philosophy
in the College of Arts and Sciences
Georgia State University

2007

Copyright by

Yubin Zhou

2007

EXPLORING THE ROLE OF CALCIUM IONS IN BIOLOGICAL SYSTEMS BY
COMPUTATIONAL PREDICTION AND PROTEIN ENGINEERING

by

Yubin Zhou

Committee Chair: Jenny J. Yang

Committee: Teryl K. Frey

Charles F. Louis

Giovanni Gadda

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2007

Acknowledgments

All of the work in this dissertation is carried out under the direction of Prof. Jenny J. Yang. I am deeply indebted to Prof. Jenny J. Yang, whose keen help, inspiring suggestions and stimulating encouragement helped me during the course of my research. Her infectious enthusiasm and unlimited zeal for science have been major driving forces through my graduate career. I would like to express my deepest gratitude to Prof. Teryl K. Frey and Prof. Charles F. Louis, who gave me their full support and guidance in research. This dissertation would not have been possible without the expert guidance of these three esteemed advisors. Many people on the faculty, staff and colleagues of the Department of Chemistry assisted and encouraged me in various ways during my graduate years. I gladly express my gratitude to them for all their help, support, interest and valuable hints. Especially I am obliged to Dr. Yiming Ye for initiating the project and his helps in molecular cloning, to Dr. Wei Yang and Dr. Hsiau-Wei Lee for their hand-to-hand teaching in NMR, to Dr. Shunyi Li for his help in protein purification, to Dr. Wen-Pin Tzeng and Dr. Monica Lurtz for their contribution to the cell-based assays. I would also extend my gratitude to Michael Kierberger for carefully proof-reading all the chapters of this dissertation and provided many stylistic suggestions. This work is supported by a fellowship from the Molecular Basis of Disease program and grants from National Institutes of Health. Finally, my very special thanks go out to my parents Wentao Zhou and Ruiju Zhou, my wife Yun Huang and my son Joseph Zhou, whose patient love and inspiration enabled me to complete this work.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xv
CHAPTER	1
1. Introduction	1
1.1. The roles of Ca^{2+} in biological systems and calciomics	1
1.2. Properties of EF-hand Ca^{2+} -binding proteins	5
1.3. Prokaryotic EF-hand-like Ca^{2+} -binding proteins	7
1.4. Viral Ca^{2+} -binding proteins	8
1.5. Calmodulin as a ubiquitous Ca^{2+} sensor protein in eukaryotic cells	9
1.6. A grafting approach to probe metal binding properties of continuous Ca^{2+} - binding sites	12
1.7. The objectives of this dissertation	13
1.8. The significance of this dissertation	16
2. Materials and methods	19
2.1. Molecular cloning and plasmid construction	19
2.2. Protein engineering	19
2.3. Expression and purification of proteins	20
2.4. Mass spectrometry	25

2.5.	Circular dichroism spectroscopy	25
2.6.	Stopped-flow measurements	27
2.7.	Fluorescence spectroscopy	28
2.8.	Isothermal titration calorimetry	31
2.9.	Surface plasmon resonance measurement	31
2.10.	Nuclear magnetic resonance spectroscopy	32
2.11.	Cell, infectious clone and site-directed mutagenesis	33
2.12.	<i>In Vitro</i> transcription, transfection, blotting	34
2.13.	Dye transfer assay	36
2.14.	Bioinformatic tools and homology structure modeling	37
2.15.	Determination of zinc amount by colorimetric PAR assay	40
3.	Prediction of prokaryotic EF-hand and EF-hand like Ca^{2+}-binding proteins	41
3.1.	Pattern development	43
3.2.	Canonical EF-hand motif	43
3.3.	Pseudo EF-hand motif	47
3.4.	Properties of EF-hand-like motif	52
3.5.	EF-hand proteins in the bacterial genomes	54
3.6.	Possible roles of single EF-hands	58
3.7.	Evolutionary perspectives on EF hand proteins	73
3.8.	Summary	77
4.	Ca^{2+}, Ca^{2+}-binding protein and virus infection	78
4.1.	Effects of viral infections on Ca^{2+} homeostasis	80

4.2. Viral Ca^{2+} -binding proteins	101
4.3. Ca^{2+} dependent virus-host interactions	122
4.4. Summary	130
5. Probing the metal binding properties of rubella virus	
nonstructural protease	132
5.1. Molecular biology of rubella virus (RUB)	132
5.2. Identification of a Ca^{2+} -binding domain in the RUB NS protease	134
5.2.1. Prediction of the EF-hand Ca^{2+} -binding motif and homology modeling of the RUB NS protease	134
5.2.2. Probing the metal binding properties of RUB NS protease by grafting	137
5.2.3. Obtaining the metal binding affinities of grafted Ca^{2+} -binding loop by FRET and NMR	140
5.2.4. Metal selectivity of the isolated EF-hand Ca^{2+} -binding motifs from RUB NS protease	142
5.2.5. Oligomeric states of the grafted EF-hand motif	144
5.2.6. Using the minimal metal binding domain RUBCa for metal- binding and conformational studies	147
5.2.7. Mutating the potential Ca^{2+} -binding ligands in infectious cDNA clones and replicons	151
5.2.8. Summary	155
5.3. Conformational and functional studies on a Zn^{2+} -binding cysteine-rich domain of RUB NS protease	162

5.3.1. Determination of Zn^{2+} contents in RUBCa and its mutants	164
5.3.2. Zn^{2+} -induced conformational changes	165
5.3.3. Effects of Cys-to-Ser mutations on protease cleavage	169
5.3.4. Secondary and tertiary structure of RUBCa and its mutants	171
5.3.5. A possible iron-containing domain in RUBCa	171
5.3.6. Summary	175
6. Prediction and identification of calmodulin binding sites in gap junction protein connexins	176
6.1. Calmodulin as intracellular Ca^{2+} sensor	176
6.1.1. Binding of metal ions to CaM: affinity, cooperativity and specificity	177
6.1.2. Structure of CaM: conformational plasticity and structural malleability	184
6.1.3. Target recognition of CaM: diversity and recurrence	187
6.1.4. Prediction and screening of potential CaM target sequences	193
6.2. Gap junction and calmodulin	196
6.2.1. Prediction of calmodulin binding sites in connexins	198
6.2.2. Secondary and tertiary structure change induced by peptide binding	200
6.2.3. Monitoring the CaM-Cx interaction by SPR	206
6.2.4. Structural changes induced by Cx binding	208
6.2.5. Peptide binding affinities and thermodynamics of the CaM-Cx interaction	210

6.2.6. Effect of ionic strength and pH on the CaM-Cx interaction	215
6.2.7. Effects of Cx binding on the metal binding properties of CaM	217
6.2.8. Physiological effects of knocking out the putative CaM binding site	222
6.2.9. Indication of <i>in vivo</i> interaction	225
6.2.10. Indication of potential CaM-Cx binding mode	228
6.2.11. <i>In vivo</i> functional analysis of the putative Cx43 CaM-binding site	229
6.3. Summary	231
7. Conclusions and major discoveries	232
Publications and manuscripts	234
Appendices	236
References	247

LIST OF TABLES

Table 1.1. List of characterized CaM target proteins	11
Table 3.1. Summary of patterns to predict EF-hand proteins	44
Table 3.2. List of the pseudo EF-hand proteins	50
Table 3.3. EF-hand-like Ca^{2+} -binding proteins with known structure	53
Table 3.4. List of potential prokaryotic EF-hand-containing proteins	60
Table 4.1. Virus-induced alterations on cellular Ca^{2+} -homeostasis	82
Table 4.2. Roles of Ca^{2+} in the life cycle of virus and viral Ca^{2+} -binding proteins	104
Table 4.3. Putative EF-hand and EF-hand-like Ca^{2+} -binding motifs predicted in virus genomes	116
Table 4.4. Interactions between cellular Ca^{2+} -binding proteins and viruses	131
Table 5.1. The metal-binding affinities of the engineered protein CD2.RUBCa and the minimal metal-binding domain RUBCa	158
Table 6.1. The interhelical angle and distance in apo-CaM and Ca^{2+} -CaM	187
Table 6.2. List of CaM-target complexes deposited in Protein Data Bank	190
Table 6.3. The binding affinities of Cx peptides to CaM	219
Table 6.4. Effect of Cx peptides binding on the Ca^{2+} binding affinity of CaM	219

LIST OF FIGURES

Fig. 1.1. Ca^{2+} as a signal of life and death	2
Fig. 1.2. Tunable Ca^{2+} -binding affinities and 3D representation of typical Ca^{2+} -binding proteins	4
Fig. 1.3. Ligand residues and coordination geometry of typical helix-loop-helix EF-hand Ca^{2+} binding motifs	6
Fig. 1.4. The sophisticated and fine-tuned CaM-target network	10
Fig. 1.5. Schematic representation of the novel grafting approach	14
Fig. 2.1. The cloning, expression and purification scheme of GST fusion proteins	22
Fig. 2.2. SDS-PAGE and MS spectrum of purified CaM	24
Fig. 3.1. Consensus sequence of canonical and pseudo EF-hand motifs drawn based on profiles HMM using LogoMat-M	45
Fig. 3.2. Statistical analysis of prediction performance by developed patterns	48
Fig. 3.3. Multiple sequence alignment of the pseudo EF-hand proteins	51
Fig. 3.4. Phylogenetic analysis of prokaryotic proteins containing multiple EF-hand motifs	57
Fig. 3.5. Phylogenetic analysis of the EF-hand protein family	76
Fig. 4.1. Examples of virus-induced perturbations on Ca^{2+} homeostasis	79

Fig. 4.2. Examples of viral Ca^{2+} -binding proteins with determined 3D structures	103
Fig. 5.1. Genomic organization of RUB and processing of viral proteins	133
Fig. 5.2. A putative EF-hand Ca^{2+} -binding motif in RUB NS protease	136
Fig. 5.3. Grafting the predicted EF-hand Ca^{2+} -binding loop into CD2.D1 and formation of metal-protein complex	138
Fig. 5.4. Obtaining metal-binding affinities using aromatic residue-sensitized Tb^{3+} -FRET and 1D ^1H NMR	141
Fig. 5.5. Metal selectivity of the engineered protein CD2.RUBCa	143
Fig. 5.6. Determination of oligomeric states of the grafted EF-hand motif using PFG NMR	146
Fig. 5.7. Metal-ion titration of the minimal metal-binding domain RUBCa	148
Fig. 5.8. Ca^{2+} -induced conformational changes and thermal unfolding of the putative metal-binding domain RUBCa	150
Fig.5.9. Effects of mutations of the potential Ca^{2+} -coordination ligands on RUB replication	152
Fig.5.10. Replicon RNA synthesis and P200 cleavage in transfected cells	154
Fig.5.11. Prediction of a putative EF-hand Ca^{2+} -binding motif in the nsP1 of alphaviruses	157
Fig.5.12. Cartoon representation of a cysteine-rich domain in RUB NS protease	163
Fig.5.13. Determination of protein-bound Zn^{2+} amounts by PAR assay.	166
Fig.5.14. Zn^{2+} -induced conformational changes	168

Fig.5.15. Effect of mutations of cysteines on P200 precursor cleavage	170
Fig.5.16. Far UV CD and intrinsic Trp fluorescence spectra of RUBCa and its cysteine mutants	172
Fig.5.17. Absorbance spectra and metal content of RUBCa	174
Fig. 6.1. The EF-hand β -scaffold structure formed by paired adjacent EF- hand Ca^{2+} binding sites, the N-terminal domain of CaM	179
Fig. 6.2. Schematic diagram of Ca^{2+} binding to two coupled EF-hand motifs and binding energetics	182
Fig. 6.3. The Ca^{2+} -induced conformational changes of CaM	188
Fig. 6.4. Membrane topology and predicted calmodulin-binding sequences in Cx43 and Cx44	199
Fig. 6.5. Far UV CD spectra of the peptides Cx43 ₁₃₆₋₁₅₈ and Cx44 ₁₂₉₋₁₅₀ with different concentrations of TFE	201
Fig. 6.6. Circular Dichroism studies of the interaction between Cx43 ₁₃₆₋₁₅₈ and CaM	203
Fig. 6.7. Circular Dichroism studies of the interaction between Cx44 ₁₂₉₋₁₅₀ and CaM	204
Fig. 6.8. Monitoring the interaction of calmodulin with Cx peptides by SRP	207
Fig. 6.9. Interaction between CaM and Cx peptides revealed by NMR	209
Fig. 6.10. Interaction of Cx peptides with dansyl-CaM.	212
Fig. 6.11. ITC microcalorimetric traces and the derived isotherms of CaM titrated with Cx44 ₁₂₉₋₁₅₀	214
Fig.6.12. Effects of salt and pH on the binding affinity of Cx44 ₁₂₉₋₁₅₀	216

Fig.6.13. Ca^{2+} titration of CaM and CaM-Cx peptides complexes	218
Fig.6.14. Stopped-flow traces for the EGTA-induced Ca^{2+} dissociation from CaM and the CaM-Cx44 ₁₂₉₋₁₅₀ complex	221
Fig.6.15. Functional effect of a sustained elevation in $[\text{Ca}^{2+}]_i$ on CaM binding-deficient Cx43-YFP mutants	224

LIST OF ABBREVIATIONS

ANS	1-anilino-8-naphthalene sulfonate
CaBP	Calcium binding protein
CaM	Calmodulin
CD	Circular dichroism
CD2.D1	Domain 1 of Cluster of Differentiation 2
Cx	connexin
DTT	1,4-dithio-DL-threitol
EGTA	Ethylene glycol tetraacetic acid
ER	Endoplasmic reticulum
ESI	Electrospray ionization
FRET	Fluorescence resonance energy transfer
FPLC	Fast performance liquid chromatography
GST	Glutathione-S-transferase
HSQC	Heteronuclear single quantum coherence
ICP	Inductively coupled plasma
IPTG	Isopropyl β -D-thiogalactoside
ITC	Isothermal titration calorimetry
K_d	Dissociation constant
k_{off}	Off rate
LB	Luria-Bertani medium
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
NSP	Nonstructural protein
NOESY	Nuclear Overhauser enhancement spectroscopy
SPR	Surface plasmon resonance
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
RUB	Rubella virus
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TFE	Trifluoroethanol
TOCSY	Total correlation spectroscopy
UV	Ultraviolet

1. Introduction

1.1. The roles of Ca^{2+} in biological systems and calciomics

Ca^{2+} , a “signal for life and death”, acts as a universal and versatile messenger in cellular signal transduction, and functions as a pivotal regulator of the cell life cycle including cell division, differentiation and apoptosis (3). In recent years, accumulating evidence indicates that Ca^{2+} even play its unique roles in the lower form of lives, such as bacteria and viruses (4-6). A new term, calciomics, has been coined to describe the sophisticated roles of Ca^{2+} in biological systems in the post-genomic era.

The versatile speed, amplitude and spatial-temporal patterning of Ca^{2+} in eukaryotic cells controls vital biological processes by precisely modulating the activity of a repertoire of signaling components including cellular receptors, ion channels, pumps, exchangers, Ca^{2+} buffers, Ca^{2+} effectors, Ca^{2+} -sensitive enzymes and transcription factors in different cellular compartments (Fig. 1.1). The extracellular space and the internal Ca^{2+} stores serve as two sources of cytosolic Ca^{2+} signals. The temporal Ca^{2+} signaling pattern is mainly observed as Ca^{2+} oscillation in the cytosol, whereas the spatial pattern of Ca^{2+} is reflected by Ca^{2+} spikes and Ca^{2+} waves due to the dynamic changes of Ca^{2+} concentration in different compartments following signal stimulation. In general, the cellular ionized Ca^{2+} gradient follows the order of extracellular space ($[\text{Ca}^{2+}]_o$: $\sim 10^{-3}$ mM) > sarcoplasmic reticulum (SR)/endoplasmic reticulum (ER) ($[\text{Ca}^{2+}]_{\text{ER}}$: $\sim 10^{-3}$ mM) >

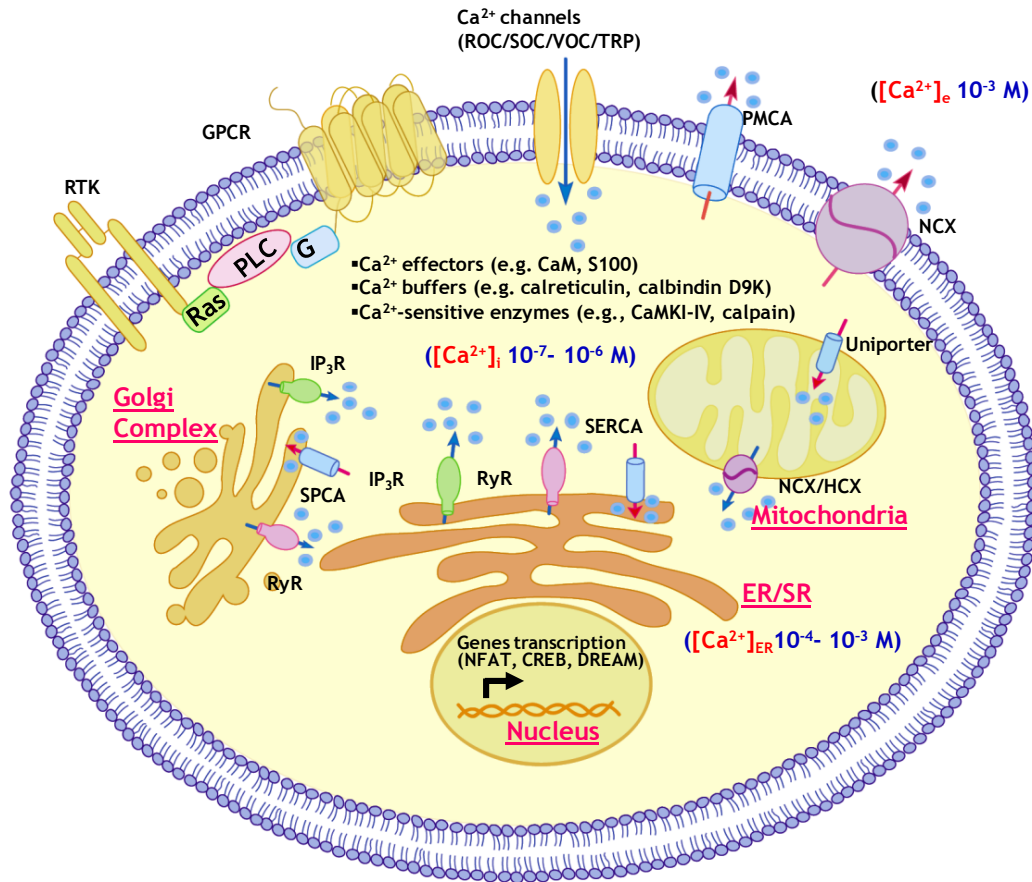


Figure 1.1. The choreography of Ca^{2+} signaling in eukaryotic cells. Upon extracellular stimulation, the free cytosolic Ca^{2+} ($[\text{Ca}^{2+}]_c$) rapidly increases due to the entry of extracellular Ca^{2+} across the plasma membrane via Ca^{2+} channels, such as voltage-operated channels (VOC), receptor-operated channels (ROC), transient receptor potential ion-channel (TRP) and store-operated channels (SOC), or by the release of Ca^{2+} from internal stores (e.g., endoplasmic reticulum (ER), Golgi complex, and lysosomes) through inositol-1,4,5-triphosphate receptors (IP_3R) and ryanodine receptors (RyR) due to activation of membrane receptors (G protein coupled receptor (GPCR) and receptor tyrosine kinase (RTK)) and the subsequent synthesis of IP_3 . At the resting state, $[\text{Ca}^{2+}]_c$ is maintained at submicromolar range by extruding Ca^{2+} outside of the plasma membrane via plasma membrane Ca^{2+} -ATPase (PMCA) and $\text{Na}^+/\text{Ca}^{2+}$ exchanger (NCX), or by pumping Ca^{2+} back into internal stores through sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase (SERCA) or secretory pathway Ca^{2+} -ATPase (SPCA). The Ca^{2+} signals are delivered by affecting the activity of Ca^{2+} buffers, Ca^{2+} effectors and Ca^{2+} -regulated enzymes. The signals can also have “long-term” effects by modulating the activity of transcriptional factors including nuclear factor of activated T cells (NFAT), cyclic AMP response element-binding proteins (CREB) and downstream regulatory element modulator (DREAM).

cytosol ($[Ca^{2+}]_c$: $\sim 10^{-7}$ M to $\sim 10^{-5}$ μ M) > other internal calcium stores such as the mitochondrion ($[Ca^{2+}]_o$: $\sim 10^{-7}$ M) and nucleus ($[Ca^{2+}]_n$: $\sim 10^{-7}$ M). Extracellular Ca^{2+} , sensed by the extracellular Ca^{2+} -sensing receptor, is believed to maintain the long-term Ca^{2+} homeostasis by replenishing the internal calcium stores. In contrast, the internal calcium stores are directly responsible for the changes in cytosolic Ca^{2+} concentration through the activity of two principal Ca^{2+} release channels, e.g., the ryanodine (RyR) receptor and the inositol 1,4,5-triphosphate (IP_3) receptor (Fig. 1.1).

Following signal stimulation or alteration in the membrane potential, the cytosolic Ca^{2+} concentrations can be elevated by 100 fold from 10^{-7} M to 10^{-5} M (7). Ca^{2+} carries out its functions by binding to specific Ca^{2+} receptors or Ca^{2+} -binding proteins (CaBPs) with varying affinities (Fig. 1.2A). According to the role Ca^{2+} ions or the proteins play in a biological context, most Ca^{2+} binding proteins may fall into one of three categories: trigger or sensor proteins (e.g., calmodulin) (8), buffer proteins (e.g., calbindin D_{9K} and parvalbumin) (9), or Ca^{2+} -stabilized proteins (e.g., thermolysin) (10) (Figs. 1.2B-D). The Ca^{2+} -binding sites may be divided into continuous or discontinuous ones. In the continuous Ca^{2+} -binding sites, the Ca^{2+} -coordinating ligand residues are from a short continuous stretch of amino acids. A discontinuous Ca^{2+} -binding pocket, however, is formed by ligands from a number of residues that are separated in the primary sequence but are in close spatial proximity in spatial arrangement. Among the continuous Ca^{2+} -binding sites, the helix-loop-helix EF-hand Ca^{2+} -binding proteins can be found in

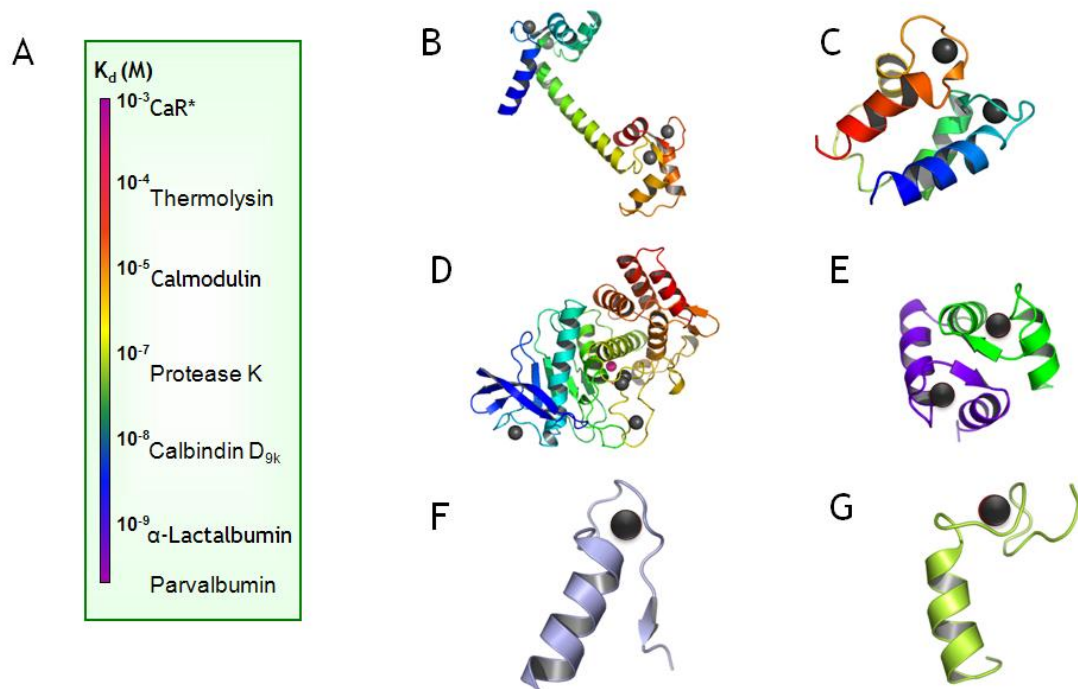


Figure 1.2. Tunable Ca^{2+} -binding affinities and 3D representation of typical Ca^{2+} -binding proteins. (A), The Ca^{2+} -binding affinities of Ca^{2+} -binding proteins may vary by 10^5 , depending on their diversified function and location in the cellular compartments. (B), 3D cartoon of prototypical EF-hand protein calmodulin (PDB code: 3cln). (C), 3D representation of calbindin D_{9K} or S100G (PDB code: 1b1g). (D) 3D structure of thermolysin (PDB code: 1tlx). (E), 3D structure of prokaryotic CaM-like protein calerythrin (PDB code: 1nya). (F-G), 3D cartoon of EF-hand-like Ca^{2+} -binding proteins: alginate binding protein (F; pdb: 1kw) and dockerin (G; PDB code: 1daq). Ca^{2+} ion is shown as black sphere.

each category and constitute more than 50% of all well-characterized Ca^{2+} -binding proteins (6,11).

1.2. Properties of EF-hand Ca^{2+} -binding proteins

The helix-loop-helix EF-hand moiety is one of the most common motifs in proteins of bacteria, archaea, and eukaryotes (6,11). Since the delineation of the EF-hand motif in 1973, the family of EF-hand proteins has expanded to include at least 66 subfamilies thus far (11-13). By binding to Ca^{2+} , this motif may undergo conformational changes enabling Ca^{2+} -modulated functions, as seen in the trigger or sensor proteins calmodulin and troponin C (8,14), or may buffer the concentration of Ca^{2+} to maintain local Ca^{2+} homeostasis as reported in buffering proteins such as parvalbumin (12) and calbindin D_{9k} (9). The coordination of Ca^{2+} in EF-hand motifs is fulfilled by adopting a pentagonal bipyramidal geometry with seven oxygen atoms from the side-chain carboxyl or hydroxyl groups, the main chain carbonyls, and a bridged water. Though EF-hands have been found abundantly in eukaryotes and bacteria (4-6,15-17), this Ca^{2+} -binding motif has seldom been reported in viruses.

EF-hand motifs are divided into two major types: the canonical EF-hands as seen in calmodulin (CaM) and the prokaryotic CaM-like protein calerythrin (Fig. 1.2E), and the pseudo EF-hands exclusively found in the N-termini of S100 and S100-like proteins (Fig. 1.2B). The major difference between these two groups lies in the Ca^{2+} -binding loop: the 12-residue canonical EF-hand loop binds Ca^{2+} mainly via side-chain carboxylates or carbonyls (loop sequence

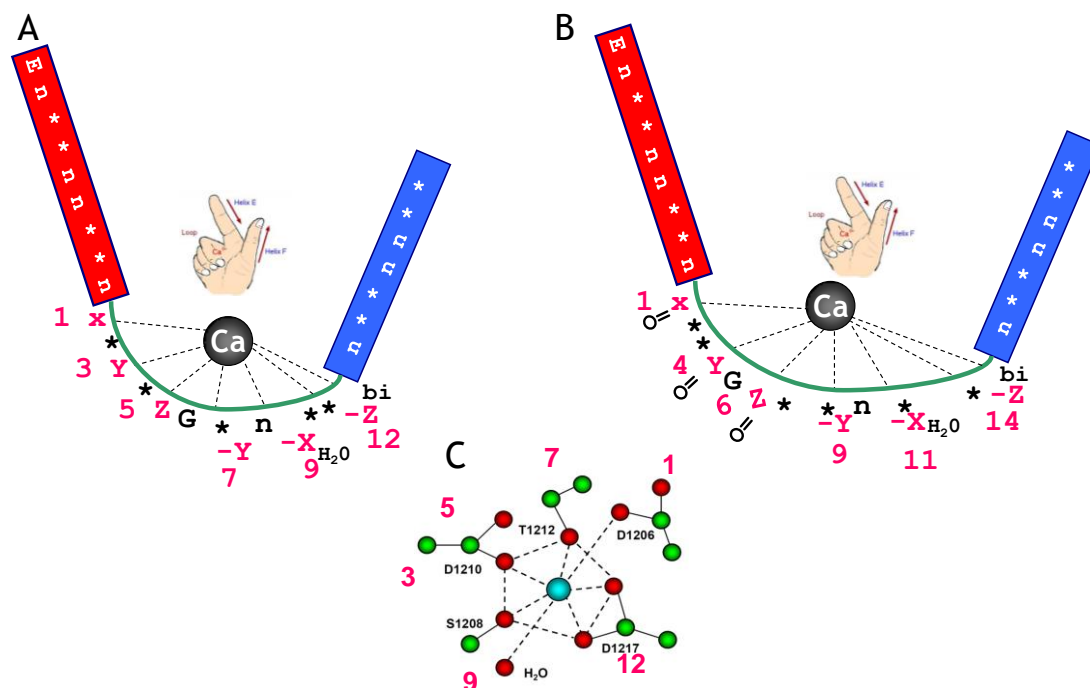


Figure 1.3. Ligand residues and coordination geometry of typical helix-loop-helix EF-hand Ca^{2+} binding motifs. (A), Cartoon illustration of the canonical EF-hand Ca^{2+} -binding motif. (B), Cartoon illustration of the pseudo EF-hand Ca^{2+} -binding motif. (C), the pentagonal bipyramidal geometry adopted by the Ca^{2+} binding site of the canonical EF-hand motif.

positions 1, 3, 5, 12) (Fig. 1.3A), whereas the 14-residue pseudo EF-hand loop chelates Ca^{2+} primarily via mainchain carbonyls (positions 1, 4, 6, 9) (Fig. 1.3B). The residue at the $-X$ axis coordinates the Ca^{2+} ion through a bridged water molecule. The EF-hand loop has a bidentate ligand (Glu or Asp) at axis $-Z$. Among all the structures reported to date, the majority of EF-hand motifs are paired either between two canonical or one pseudo and one canonical motifs. For proteins with odd numbers of EF-hands, such as the penta-EF-hand calpain, EF-hand motifs were coupled through homo- or hetero-dimerization (18-22).

1.3. Prokaryotic EF-hand-like Ca^{2+} -binding proteins

Recently, EF-hand-like proteins with diversified flanking structural elements around the Ca^{2+} -binding loop have been reported in bacteria (Figs. 1.2F-G) (5,23,24). Several lines of evidence indicate that these prokaryotic EF-hand-like proteins are widely-implicated in Ca^{2+} signaling and homeostasis in bacteria (4,5,17,25). They contain flexible lengths of Ca^{2+} -binding loops that differ from the EF-hand motifs. However, their coordination properties resemble classical EF-hand motifs. For example, the semi-continuous Ca^{2+} -binding site in D-galactose-binding protein (GBP) contains a nine-residue loop (aa 134-142). The Ca^{2+} ion is coordinated by seven protein oxygen atoms, five of which are from the loop mimicking the canonical EF-loop whereas the other two are from the carboxylate group of a distant Glu (aa 205). Another example is a novel domain named Excalibur (extracellular Ca^{2+} -binding region) isolated from *Bacillus subtilis*. This domain has a conserved 10-residue Ca^{2+} -binding loop strikingly

similar to the canonical 12-residue EF-hand loop (23). The diversity of the structure of the flanking region is illustrated by the discovery of EF-hand-like domains in bacterial proteins. For example, a helix-loop-strand instead of the helix-loop-helix structure is observed in periplasmic galactose-binding protein (*Salmonella typhimurium*, 1gcg) (24) or alginate-binding protein (*Sphingomonas* sp., 1kwh) (Fig. 1.2F) (26); the entering helix is missing in protective antigen (*Bacillus anthracis*, 1acc) (27) or dockerin (*Clostridium thermocellum*, 1daq) (Fig. 1.2G) (28). Previous studies in our laboratory have also shown that the single Ca^{2+} -binding loops from CaM are capable of binding Ca^{2+} either alone or with the flanking helices when they are inserted into a non- Ca^{2+} -binding host protein CD2 domain 1 with β -strand structure (1,2). The four EF-loops of CaM in the host protein have dissociation constants (K_d) ranging from 34 μM to 814 μM (1). NMR studies revealed that the grafted EF-loop is directly involved in chelating Ca^{2+} (29).

1. 4. Viral Ca^{2+} -binding proteins

Being extremely adept at hijacking the host cellular machinery, viruses have been extensively reported to interfere with the Ca^{2+} signaling pathways or Ca^{2+} -dependent processes, and thereby, achieve their optimal infectivity to produce progenies (30). The interplays between virus and Ca^{2+} , in general, fall into three major categories: 1) viral particles or viral proteins directly or indirectly disturb the Ca^{2+} homeostasis by altering membrane permeability and/or manipulating key components of the Ca^{2+} -signaling repertoire; 2) a number of

important structural or nonstructural viral proteins directly bind to Ca^{2+} for structural integrity or optimal functions; 3) virus-host interactions that require cellular Ca^{2+} -sensitive proteins or processes.

Ca^{2+} -binding motifs are present in virus proteins and are involved in virion assembly and stability (31-43), virion-associated activities such as cell fusion (44,45), and neuraminidase activity (46,47). To date, almost all of the reported viral CaBPs are structural proteins.

1.5. Calmodulin as a ubiquitous Ca^{2+} sensor protein in eukaryotic cells

CaM is a small (148 amino acids; MW: 16.7 kDa) and acidic (pI: ~4.0) EF-hand Ca^{2+} -binding protein. It was first discovered in the brain and heart as cyclic nucleotide phosphodiesterase activator protein (PAF) (48). It consists of two globular and autonomous domains, each of which contains two helix-loop-helix EF-hand motifs. Through its reversible or irreversible binding to Ca^{2+} , the resultant conformational changes and the interaction with target proteins, CaM is capable of transducing the intracellular Ca^{2+} signal changes into a myriad of divergent cellular events, such as cell proliferation, cell differentiation and apoptosis (3).

Upon signal stimulation or alteration in the membrane potential, the cytosolic Ca^{2+} concentrations might be elevated by 100 fold from 10^{-7} M to 10^{-5} M. The elevation of $[\text{Ca}^{2+}]_c$ can be effectively sensed by CaM since the Ca^{2+} binding affinity of CaM ranges from 0.2 μM to 2 μM . Ca^{2+} signaling is a fast process and takes place in milliseconds (3). To convert this transient Ca^{2+} signals into more sustained physiological processes, CaM undergoes conformational

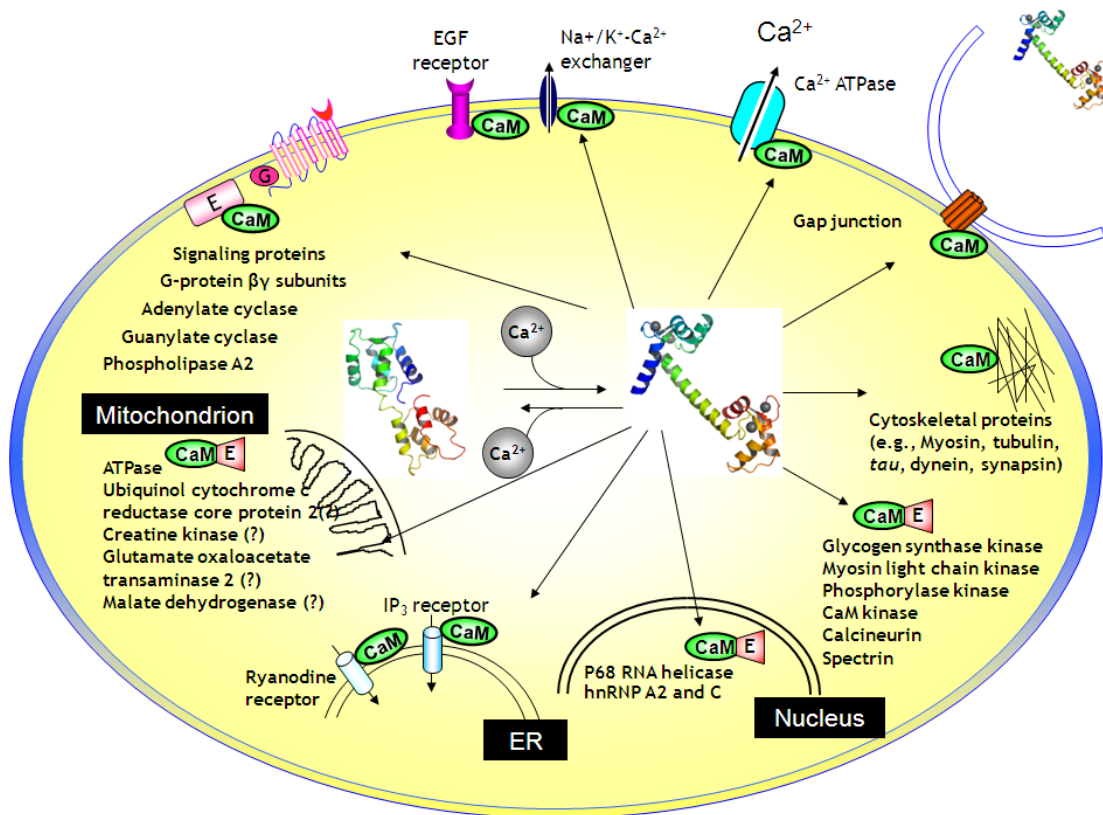


Figure 1.4. The sophisticated and fine-tuned CaM-target network. Through Ca^{2+} -dependent interaction with targets, CaM relays the Ca^{2+} signaling to a multitude of target enzymes or proteins, which are distributed in different subcellular compartments, to carry out diversified cellular functions.

Table 1.1. List of characterized CaM target proteins.

Category	Target Protein
Receptors, ion channels, exchangers, and pumps	SK channels EGFR Na ⁺ /K ⁺ Ca ²⁺ exchanger connexin Ca ²⁺ ATPase IP3 receptor L-type Ca ²⁺ channel Ca ²⁺ -pump PMCA RyR receptor
Signaling proteins	Adenylate cyclase I G-protein coupled receptor NOS I, III cAMP phosphodiesterase
Kinases and phosphatases	CaMK I-IV CaMKKs calcineurin phosphorylase kinase phosphofructokinase NAD kinase glycogen synthase kinase
Cytoskeleton and motility	MLCK caldesmon spectrin myosin MARCKS dystrophin synthrophin actinin IQGAP1 P190 tau synapsin
Gene expression and regulation	p68 RNA helicase hnRNP p62
Others	cyclin E neuromodulin neurogranin

changes and interacts with a variety of protein kinases and phosphatases, both of which are capable of covalently modifying downstream effectors by adding or removing a phosphate group (49-53). The time scale of these processes usually lasts for seconds or even minutes. In addition, to more efficiently modulate the signaling cascades, Ca^{2+} -CaM (sometimes apo-CaM) targets to enzymes or proteins (e.g., adenylate cyclases, phosphodiesterase, nitric oxide synthase, G-protein coupled receptor) involved in the conversion of other secondary messengers, such as cAMP and cGMP (49-53). These enzymes or proteins themselves do not respond to changes in Ca^{2+} concentration. Thus, CaM relays the Ca^{2+} signaling to more than 100 target enzymes or proteins to carry out corresponding cellular functions (Fig. 1.4; Table 1.1) (54). More importantly, the Ca^{2+} signal itself can be modulated by CaM through its interaction with membrane receptors, ion channels and pumps (55,56). CaM resides primarily in the cytosol and targets to a wide range of proteins (54). The sophisticated and fine-tuned CaM-target network leads to the occurrence of divergent cellular responses to transient Ca^{2+} signals in different types of cells. Our findings, as shown in Chapter 6, suggest that gap junction proteins can be added to the expanding list of CaM targets.

1.6. A grafting approach to probe metal binding properties of continuous Ca^{2+} - binding sites

A grafting approach has been developed in our laboratory to validate the Ca^{2+} -binding capability of any predicted EF-hand motifs, and to further analyze

the factors governing the metal binding properties in an isolated EF-hand motif (1,57). In this approach, any single Ca^{2+} -binding site can be engineered into a scaffold protein, obviating the expression and purification of the intact proteins and the complexity of cooperativity. Thus, it is possible to dissect key local factors that contribute to the intrinsic Ca^{2+} -binding affinity and conformational properties.

Through previous study in our laboratory, we find that the domain 1 of the rat cell adhesion protein CD2 is an excellent scaffold protein (1,57). It can be used to make an accurate measurement of metal-binding affinity using aromatic residue-sensitized Tb^{3+} fluorescence resonance energy transfer (Tb^{3+} -FRET) and a competition assay utilizing Trp residues in CD2 (Fig. 1.5). CD2 retains its native structure with the inserted Ca^{2+} -binding sites, which allows for the measurement of the intrinsic Ca^{2+} -binding affinity with a minimized contribution of protein conformational change and metal-metal interaction (1,2,29,57). The metal-binding affinity of the grafted EF-loop is independent of the host protein environment. Using this grafting approach, we have reported the first estimation of the intrinsic Ca^{2+} affinities of the four EF-hand loops of CaM and their associated cooperativity (1). Further, this grafting approach can be applied to identify and verify continuous Ca^{2+} -binding sites in naturally-occurring proteins.

1.7. The objectives of this dissertation

The objectives of this research are to predict and investigate EF-hand Ca^{2+} -binding proteins using computational tools and protein engineering. The main focus of this research is to develop robust tools or methods to predict and

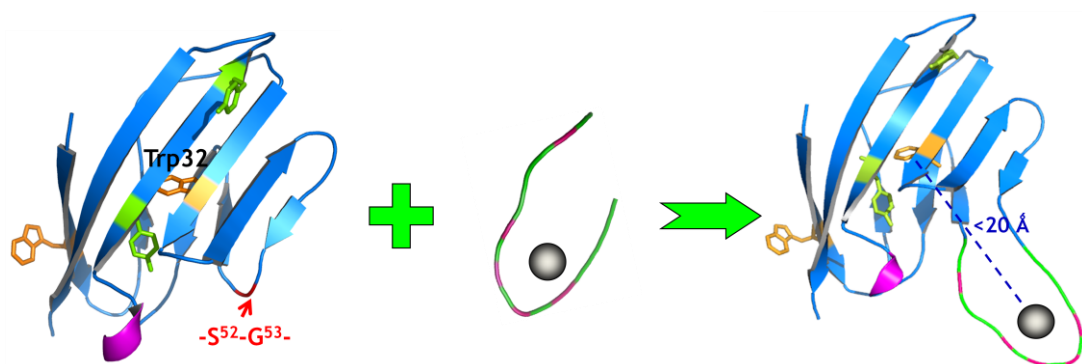


Figure 1.5. Schematic representation of the grafting approach. Any continuous Ca^{2+} -binding sequence can be inserted into a scaffold protein CD2 domain 1 (CD2.D1) between S^{52} and G^{53} . The host protein CD2.D1 remains stable at pH 4-11 and is highly tolerant to mutations and insertions. Triple-glycine (-GGG-) linkers are added on each side to render increased flexibility. The aromatic residue Trp32, which is buried between two layers of β -sheets, is within 20 Å of the potential metal binding pocket. This unique design makes it possible to directly monitor Tb^{3+} binding utilizing fluorescence resonance energy transfer (Tb^{3+} -FRET).

analyze the EF-hand Ca^{2+} -binding motif as well as its interaction with target proteins. To reflect the implication of Ca^{2+} in all biological systems, representatives were particularly selected from prokaryotes (prokaryotic genomes), virus (Rubella virus nonstructural protease) and eukaryotes (CaM). Specifically, this research focuses on three key objectives, as follows:

A. Developing a robust and reliable prediction method for EF-hand Ca^{2+} -binding proteins and analyzing prokaryotic EF-hand proteins

To facilitate the understanding of the role of Ca^{2+} in biological systems using genomic information, we will establish a web searching platform for the prediction of EF-hand and EF-hand-like Ca^{2+} -binding proteins. Improvements will be introduced to currently existing patterns to increase prediction accuracy. Systemic analysis of prokaryotic and viral EF-hand proteins will be made to grasp the evolutionary scenario of EF-hand motifs.

B. Investigating the biological role of a putative viral EF-hand motif

The biological role of a predicted viral EF-hand Ca^{2+} -binding domain within the rubella virus (RUB) nonstructural (NS) protease will be investigated by three approaches. 1) A grafting approach will be applied to investigate the metal binding properties of the isolated EF-hand motif and its mutant. 2) A minidomain approach will be used to analyze the metal induced conformational changes. 3) The physiological relevance of the EF-hand Ca^{2+} -binding motif to virus replication and posttranslational processing will be assessed using rubella virus infectious clone and its replicon systems.

C. Elucidating the molecular basis underlying the calmodulin-connexin interaction in eukaryotic cells

We propose a model in which the regulation of gap junctions is mediated by calmodulin (CaM) in a Ca^{2+} -dependent manner. The molecular basis for the interaction of CaM with putative CaM-binding sequences in the connexins will be revealed using various biophysical methods, including high resolution NMR.

1.8. The significance of this dissertation

The significance and high impact of this research is multifaceted. **First**, our development of easily accessible computational tools for the prediction of EF-hand or EF-hand-like Ca^{2+} -binding protein will benefit the whole metalloprotein field. To our knowledge, our server (<http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm>) will be the first of its kind to provide such service for scientists. **Second**, the comprehensive prediction and detailed analysis on putative prokaryotic and viral EF-hand and EF-hand like Ca^{2+} -binding proteins will contribute to our understanding of the versatile roles Ca^{2+} ions play in various biological systems; **Third**, our study on the viral EF-hand Ca^{2+} -binding protein provides a novel approach to probe individual Ca^{2+} -binding site at simplified and more readily achievable levels, and concurrently overcomes the barriers encountered in the expression of intact proteins; **Fourth**, the biological relevance of the EF-hand Ca^{2+} -binding motif within the rubella virus (RUB) nonstructural (NS) protease will be elucidated; **Fifth**, the confirmation of Ca^{2+} -dependent CaM-Cx interaction model will facilitate understanding of

mechanisms underlying gap junction regulation, and thus provide a molecular basis of potential therapeutic intervention of diseases related to malfunction of gap junctions, such as cataracts and heart diseases.

Chapter 2 in this dissertation summarizes all the methods used in this study, including molecular cloning, protein engineering, protein expression and purification, spectroscopic techniques (UV, circular dichroism, fluorescence, NMR), biocomputational analyses of genomes, homology structure modeling, and biochemical assays. All mathematical equations used in this study are also listed in Chapter 2.

Chapter 3 explains the development of robust and complete patterns (or motif signatures) for the prediction of EF-hand Ca^{2+} -binding proteins. It also includes our systemic genome-wide analysis of putative prokaryotic EF-hand proteins and discussion of possible evolutionary scenarios of the EF-hand protein family.

Chapter 4 summarizes most of the known mechanisms underlying virus-mediated alterations in Ca^{2+} homeostasis and Ca^{2+} -dependent virus-host interactions. The altered Ca^{2+} -signaling is closely associated with virus entry, viral gene expression, virion maturation and release. In addition, a total of 93 EF-hand motifs in viral proteins that are worthy of further studies are predicted.

Chapter 5 is devoted to the identification and confirmation of a viral EF-hand Ca^{2+} -binding, and Zn^{2+} -binding domain in the rubella virus nonstructural protease. The unique and important biological relevance of this motif with the virus infectivity is characterized.

Chapter 6 in this dissertation is focused on the elucidation of the molecular basis underlying the interaction between gap junction protein connexins and calmodulin, a prototypical EF-hand Ca^{2+} -binding protein. Predicted CaM-target sequences in the connexins are synthesized and their ability to bind CaM is determined using a range of biophysical approaches.

Chapter 7 is a succinct summary of major discoveries of this dissertation.

Appendices list preliminary results obtained from related side projects that remain incomplete. It includes: 1) analyzing key factors governing the metal binding events by inserting well-characterized Ca^{2+} -binding sequences from cellular proteins into CD2.D1 with the grafting approach; 2) predicting putative CaM binding site(s) in rubella virus nonstructural protease.

2. Materials and methods

2.1. Molecular cloning and plasmid construction

The sequence encoding the putative Ca^{2+} -binding domain (RUBCa, aa 1143-1252 of the NS-ORF) with 5' BamH I and 3' EcoR I sites was amplified from the RUB infectious cDNA clone Robo502 (58,59) by using standard PCR methods. The PCR product was subsequently inserted into the BamH I-EcoR I-digested pGEX-2T vector (GE Healthcare) to produce the plasmid pGEX-2T-RUBCa.

2.2. Protein engineering

The predicted 12-residue Ca^{2+} -binding loop was directly inserted between S⁵² and G⁵³ (denoted as pGEX2T-CD2.RUBCa) by PCR using an established protocol (60). The mutant (denoted as pGEX2T-CD2.RUBCa-AA) with double mutations D⁵A and D¹²A (numbered according to the loop position of inserted Ca^{2+} -binding motif) was produced using standard PCR methods. All sequences were verified by automated sequencing on an ABI PRISM-377 DNA sequencer (Applied Biosystems) in the Advanced Biotechnology Core Facilities of Georgia State University.

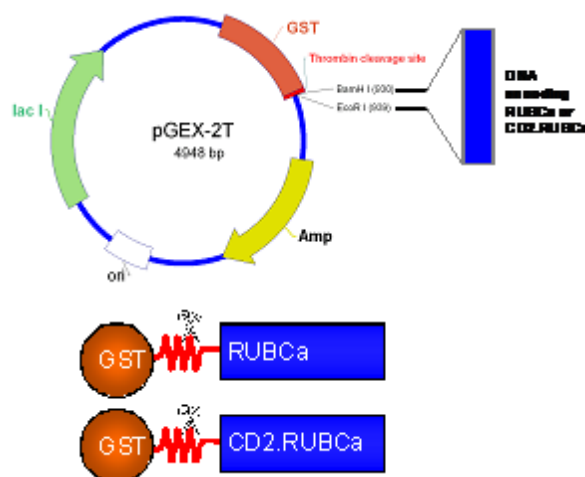
2.3. Expression and purification of proteins

CD2.RUBCa and its mutant. The engineered proteins CD2.RUBCa and its mutant CD2.RUBCa.AA were expressed as GST fusion proteins in *Escherichia coli* BL21 (DE3) transformed with the plasmid pGEX2T-CD2.RUBCa or pGEX2T-CD2.RUBCa.AA in LB medium with 100 mg/L of ampicillin and grown at 37 °C. 100 μ M of isopropyl- β -D-thiogalactopyranoside (IPTG) was added when the O.D.₆₀₀ reached 0.6 to induce protein expression for another 3 to 4 hours at 37 °C. The cultures were centrifuged at 7,000 rpm at 4 °C with a Sorvall centrifuge equipped with a SG-3 rotor. The harvested cell pellets were resuspended in lysis buffer consisting of 0.2% sarcosine, 1mM DTT, 1 mM AEBSF in PBS, pH 7.4. The resuspended solution was sonicated 6 times with each time 20 seconds with 80% duty. After centrifugation at 17,000 rpm with a Sorvall centrifuge equipped with a SS-34 rotor at 4 °C, the clarified supernatant was passed through a flow-through affinity column loaded with 4-5 mL slurry GS4B beads (GE Healthcare). After a minimum of 10 bead-volume of washing with PBS, on-column cleavage was performed to remove the GST tag by adding 20 units of thrombin to each column. The elutant containing the target protein was filtered through 0.45 μ M filter and injected into Superdex 75 gel filtration column. The eluted proteins were pooled together and further purified by Hitrap SP cation exchange chromatography. The molecular weight of CD2.RUBCa and its mutant were also confirmed by MALDI-TOF-MS. The concentration was determined by using the absorption at 280 nm with an extinction coefficient of 11,700 M⁻¹ cm⁻¹ (61).

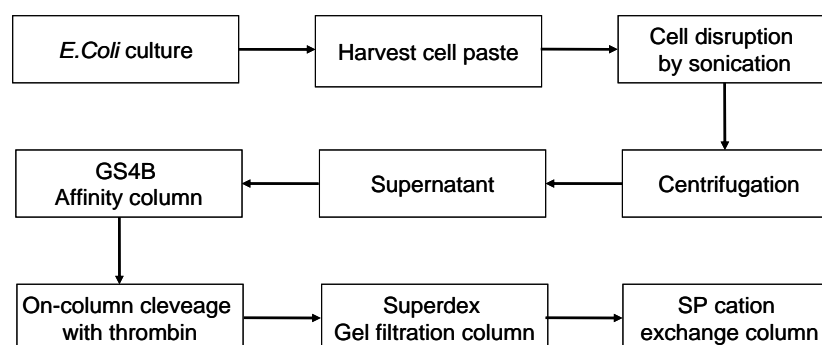
RUBCa (Fig. 2.1). The RUBCa was expressed as a GST fusion protein in *Escherichia coli* BL21 (DE3) transformed with the plasmid pGEX-2T-RUBCa in LB medium with 100 mg/L of ampicillin and grown at 37 °C. 100 μ M of isopropyl- β -D-thiogalactopyranoside (IPTG) and 50 μ M of ZnCl_2 were added when the O.D.₆₀₀ reached 0.7 to induce protein expression for another 3 to 4 hours. The proteins were purified following the protocols for GST-fusion protein purification (60) using glutathione sepharose 4B beads (GE Healthcare). The protein was cleaved from its GST tag on beads by taking advantage of the thrombin cleavage site and eluted. The elutants containing RUBCa were further purified using Superdex 75 and Hitrap SP columns (GE Healthcare). The molecular weight of RUBCa was confirmed by MALDI-TOF-MS in the Advanced Biotechnology Core Facilities of Georgia State University. The concentration of RUBCa was measured by its absorption at 280 nm with an extinction coefficient of 19,630 $\text{M}^{-1} \text{cm}^{-1}$, which was calculated according to previously described methods (62).

CaM. Recombinant rat CaM was expressed in *Escherichia coli* strain BL21(DE3)pLysS transformed with the plasmid pET7-CaM that harbors the synthetic CaM gene (63). pET7-CaM transformed cells were grown in LB medium to obtain unlabeled CaM. ^{15}N -labeled CaM was expressed in SV minimal medium using 0.5 g/L $^{15}\text{NH}_4\text{Cl}$ (Cambridge Isotope Laboratories, MA, USA) as the sole nitrogen source. Bacterially-expressed CaM was purified by phenyl-Sepharose (Sigma, MO, USA) chromatography as previously described. The purity of CaM was examined by mass spectrometry or SDS-PAGE (Fig. 2.2). The concentration of CaM was determined by using the ϵ_{276} of 3,030 $\text{M}^{-1} \text{cm}^{-1}$ (64).

A



B



C

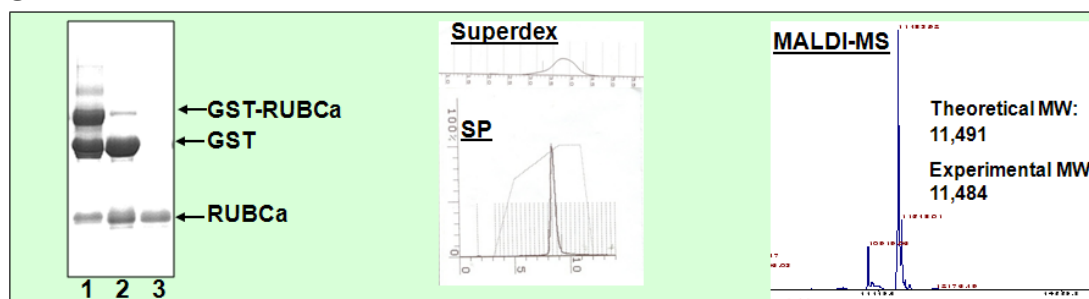


Figure 2.1. The cloning, expression and purification scheme of GST fusion proteins. (A), Target genes are inserted into EcoRI and BamHI digestion sites of pGEX-2T vector which harbors a GST tag that could be removed by thrombin. (B), A typical purification scheme of GST fusion proteins. (C), SDS-PAGE, elution profiles and MS spectrum of the purified protein RUBCa. The SDS-PAGE (left) shows the cleavage of GST fused RUBCa with thrombin for 1 h (lane 1) and 3 h (lane 2), as well as the elutants containing target proteins (Lane 3). The elutants are further purified to homogeneity by gel filtration (Superdex) and cation exchange (SP) chromatography (middle). The molecular weight is finally confirmed by MALDI-TOF-MS (right).

Dansyl CaM was prepared according to the method of Johnson *et al* with slight modifications (65). Briefly, rat CaM was dansylated in the dark by mixing 1 ml protein (1 mM) with a 5-fold molar excess of dansyl chloride (dissolved in 1:1 acetone/ethanol) in 10 mM Mops, 100 mM KCl, 1 mM CaCl₂, pH 7.0 for 16 h at 4 °C. The reaction mixture was then extensively dialyzed against 10 mM Tris, 100 mM KCl at pH 7.4 to remove the residual free dansyl chloride. The modification of CaM by dansyl chloride was confirmed by ESI-MS with an increase of +233 in the molecular mass. The bound dye concentration was determined by using the ϵ_{335} of 3980 M⁻¹ cm⁻¹ (65). An average of ~0.8 mol of the dansyl chromophore was incorporated into per mol of CaM.

Peptides. The peptides Cx43₁₃₆₋₁₅₈ (Ac-KYGIEEHGKVKMRGGLLRTYIIS-NH₂) and Cx44₁₂₉₋₁₄₉ (Ac-VRDDRGGKVRIAGALLRTYVFN-NH₂) were synthesized by Sigma-Genosys (Sigma, USA) and purified by preparative reversed-phase HPLC with purity over 95%. A randomized control peptide (Ac-LGGEYLVTMESKIHKGKRIGYR-NH₂), with the same composition of amino acids as Cx43₁₃₆₋₁₅₈ but arranged in a different order, was similarly synthesized. The molecular weight of the synthetic peptide was determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. To mimic its protein environment and eliminate extra charges, the designed peptide was blocked at its N-terminus with an acetyl group and at its C-terminus with an amide group.

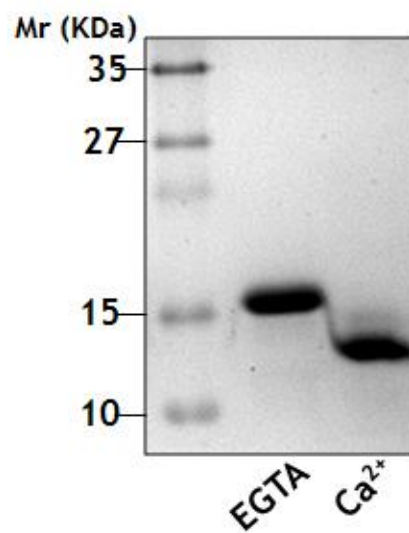
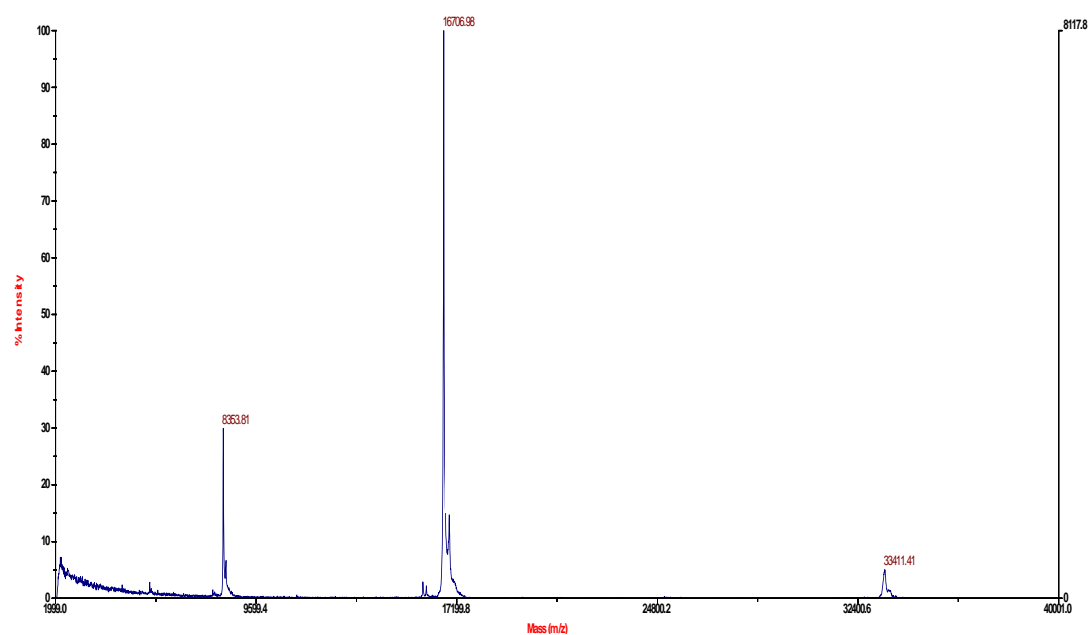
A**B**

Figure 2.2. SDS-PAGE (A) and MS (B) spectrum of purified CaM. Upon binding to Ca^{2+} , the CaM undergoes conformational changes by switching from a globular shape at its apo-form to an elongated dumbbell shape, and therefore, moves faster in the electric field. The measured molecular weight (16706.98 Da) matches well with the calculated theoretical molecular mass (16707 Da).

2.4. Mass spectroscopy

Electrospray ionization mass spectrometry (ESI-MS). The detection of metal-protein complex by ESI-MS was performed by using a Waters Micromass Q-Tof micro instrument. The data were acquired in positive ion mode by syringe pump infusion of the protein solutions at a flow rate of 7 $\mu\text{L}/\text{min}$. The protein sample stock (~ 1 mM) in 10 mM Tris, pH 7.4 was diluted 100 folds into water. Metal ions were added in 5 molar excess to the protein concentration to observe specific binding.

Inductively coupled plasma mass spectrometry (ICP-MS). All the glassware, plastiware and Teflon containers used in the preparation of samples were pretreated with 2% HNO_3 (optima grade, Fisher Scientific). All buffers were pretreated with chelex-100 (Bio-Rad) to remove the background Ca^{2+} ions. 20 to 30 μM of the refolded proteins were acidified with 2% HNO_3 and analyzed by ICP-MS (Finnigan Element 2). The dialyzed buffer was used as a blank, and the residual Ca^{2+} background of 2-5 μM was subtracted from the measurement of protein samples.

2.5. Circular dichroism spectroscopy

CD2.RUBCa and RUBCa. The CD spectra of proteins were recorded in a Jasco-810 spectropolarimeter at ambient temperature using a quartz cell of 10 mm path length with protein concentrations ranging from 2 to 5 μM . All spectra were obtained as the average of at least eight scans with a scan rate of 100 nm/min. The calculation of secondary structure elements was performed by using

DICHROWEB, an online server for protein secondary structure analyses (66). The thermal denaturation was studied using a 1 mm quartz cell with protein concentration of 15-30 μM in 10 mM Tris-HCl (pH 7.4), 10 mM KCl with 1 mM EGTA or 1 mM CaCl_2 . The ellipticity was measured from 190 to 260 nm and converted to mean residue molar ellipticity ($\text{deg cm}^2 \text{dmol}^{-1} \text{res}^{-1}$). To obtain the thermal transition point, the signal changes at 222 nm were fitted using the equation described previously (67).

CaM and CaM-Cx complex. Circular dichroism spectra were acquired in the far UV (190-260 nm) or near UV region (250-340 nm) on a Jasco-810 spectropolarimeter at room temperature using a 1 cm path length quartz cuvette. All spectra were an average of 10-20 scans. The background signals from the corresponding buffers were subtracted from the sample signals. The far UV CD spectra of the peptide in different percentages of trifluoroethanol (TFE) were obtained using 10 mM of the peptide in 10 mM Tris-HCl, 10 or 100 mM KCl, pH 7.4. In the peptide titration of CaM experiment, 2-5 μL aliquots of the peptide stock solution (150 μM in 10 mM Tris-HCl, 10 or 100 mM KCl, pH 7.4) was gradually added into a 2-mL solution containing 1-2 μM CaM in the same buffer with 5 mM CaCl_2 or 5 mM EGTA. The signals from the peptide itself were subtracted. All the measurements were carried out in at least triplicate. The binding constants of the synthetic peptide to CaM were obtained with a 1:1 binding model by fitting the following equation (Eq. 1):

$$f = \frac{([CaM]_T + [P]_T + K_d) - \sqrt{([CaM]_T + [P]_T + K_d)^2 - 4[CaM]_T[P]_T}}{2[CaM]_T} \quad (1)$$

where f is the fractional change of CD signals at 222 nm, K_d is the dissociation constant for the peptide, and $[P]_T$ and $[CaM]_T$ are the total concentrations of the synthetic peptide and CaM, respectively. The secondary structure contents of the peptides or proteins were calculated with the online secondary structure prediction server DICHROWEB that integrates analysis algorithms such as SELCON, CONTINLL, CDSSTR and K2D (66). Near UV CD spectra were recorded with protein/peptide concentration of 80-100 μ M in 10 mM Tris-HCl, 100 mM KCl, pH 7.4 with 5 mM Ca^{2+} or 5 mM EGTA.

2.6. Stopped-flow measurements

Stopped-flow experiments were performed in a Jasco-810 spectropolarimeter equipped with a BioLogic stopped-flow apparatus at 25 °C. The instrument dead-time is ~2.5 ms at a drive force of about 0.4 MPa. The optical cuvette pathlength is 0.5 cm. All the experiments were performed in 50 mM Tris-HCl, 100 mM KCl, pH 7.4 with at least 15 traces recorded. The changes of CD signal at 222 nm were monitored with 1 ms sampling interval in the time range of 0-2 s. EGTA-induced Ca^{2+} dissociation was studied by mixing 100 μ L of 2 μ M CaM-peptide mixture (with 0.1 mM Ca^{2+}) in Syringe 1 with an equal volume of 10 mM EGTA in Syringe 2 within 15 ms. The acquired data were fitted by a single exponential model.

2.7. Fluorescence spectroscopy

CD2.RUBCa and *RUBCa*. Fluorescence emission spectra were measured with a 1-cm pathlength cell on a PTI lifetime fluorometer at ambient temperature. For intrinsic tryptophan fluorescence, spectral measurements were carried out at protein concentrations of 2-4 μM in 20 mM PIPES-10 mM KCl at pH 6.8 with slit widths of 4 and 8 nm for excitation and emission, respectively. The emission spectra were collected from 300 to 400 nm with an excitation wavelength of 282 nm. 8-Anilino-1-naphthalene sulfonic acid (ANS) fluorescence emission spectra were recorded from 400 to 600 nm with an excitation wavelength of 390 nm. Protein samples (5 μM) with 1 mM CaCl_2 or 1 mM EGTA were added into the solution containing 40 μM ANS, 20 mM PIPES-10 mM KCl at pH 6.8.

Tyr/Trp-sensitized Tb^{3+} fluorescence energy transfer (Tb^{3+} -FRET) experiments were conducted as described previously (68,69). For the metal competition studies, the solution containing 40 μM of Tb^{3+} and 1.5 μM of protein was set as the starting point. The stock solution of metal ions with the same concentration of Tb^{3+} and protein was gradually added into the initial mixture. The fluorescence intensity was normalized by subtracting the contribution of the baseline slope using logarithmic fitting. The Tb^{3+} -binding affinity of protein was obtained by fitting normalized fluorescence intensity data using (Eq. 2):

$$f = \frac{([P]_T + [M]_T + K_d) - \sqrt{([P]_T + [M]_T + K_d)^2 - 4[P]_T[M]_T}}{2[P]_T} \quad (2)$$

where f is the fractional change, K_d is the dissociation constant for Tb^{3+} , and $[P]_T$ and $[M]_T$ are the total concentrations of protein and Tb^{3+} , respectively.

The Ca^{2+} competition data were first analyzed using the apparent dissociation constant obtained by equation 1. By assuming that the sample is almost saturated with Tb^{3+} at the starting point of the competition, the Ca^{2+} -binding affinity is further obtained by using the equation (Eq. 3)

$$K_{d,\text{Ca}} = K_{app} \times \frac{K_{d,\text{Tb}}}{K_{d,\text{Tb}} + [M_{\text{Tb}}]} \quad (3)$$

where $K_{d,\text{Ca}}$ and $K_{d,\text{Tb}}$ are the dissociation constants of Ca^{2+} and Tb^{3+} , respectively. K_{app} is the apparent dissociation constant. All the measurements were conducted in triplicate.

CaM. Steady-state fluorescence spectra were recorded using a QM1 fluorescence spectrophotometer (PTI) with a xenon short arc lamp at ambient temperature. Tyrosine fluorescence was monitored using excitation at 277 nm and emission at 307 nm with 2-4 nm bandpasses. The Ca^{2+} binding constants were determined by titrating the CaM (8 mM) or 1:1 CaM-peptide mixture (8 μM) in 1 mM EGTA, 100 mM KCl, 50 mM Tris-HCl, pH 7.4 with 1-5 μL aliquots of 10 mM Ca^{2+} stock solution in the same buffer containing equal concentrations of CaM and peptide. The pH change (0.03 to 0.04) was minimal during the titration process. To obtain the accurate Ca^{2+} concentrations during titration, Ca^{2+} concentration at each point was determined with the Ca^{2+} dye Oregon Green 488 BAPTA-5N (0.2 μM ; $K_d = 20 \mu\text{M}$; Invitrogen) with an excitation wavelength of 492 nm and an emission wavelength of 520 nm. The ionized Ca^{2+} concentration was subsequently calculated according to the equation (Eq. 4):

$$[Ca^{2+}] = K_d \frac{F - F_{\min}}{F_{\max} - F} \quad (4)$$

where F is the fluorescence intensity of the dye at each titration point, F_{\min} and F_{\max} are the intensities of the Ca^{2+} -free and the Ca^{2+} -saturated dyes, respectively. The Ca^{2+} titration of CaM data was fit to the nonlinear Hill equation (Eq. 5):

$$f = \frac{[M]^n}{K^n + [M]^n} \quad (5)$$

Where, f is the relative signal change observed during the experiment; $[M]$ is the concentration of free Ca^{2+} ; K refers to dissociation constants of Ca^{2+} ; and n is the Hill coefficient.

For dansyl-CaM fluorescence measurement, 1 mL solution containing 1-2 μ M dansyl-CaM in 10 mM Tris-HCl, 100 mM KCl, pH 7.4 with 5 mM Ca^{2+} or 5 mM EGTA was titrated with 5-10 μ L aliquots of the peptide stock solution (10 μ M) in the same buffer. The fluorescence spectra were recorded using an excitation of 335 nm and an emission between 400 and 600 nm with the slit width set at 4-8 nm.

Tb^{3+} luminescence lifetime measurement was performed on a QM1 fluorescence spectrophotometer with a xenon flash lamp (PTI) at 25 °C. The buffer consists of 50 mM Tris, 100 mM KCl, pH 7.4. The lifetime value was obtained by fitting the acquired data to a single-component exponential decay function.

2.8. Isothermal titration calorimetry

ITC experiments were performed on a Microcal VP-ITC microcalorimeter. Samples were prepared and dialyzed in a buffer consisting of 20 mM PIPES, 100 mM KCl, 2 mM CaCl_2 , pH 6.8. All the solutions were degassed for at least 15 min prior to experiments. 4-6 μL aliquots of peptides (400-600 μM) were injected from the syringe into the reaction cell containing 25 μM CaM in the same buffer at 5 min intervals at 25 °C. The heat of dilution and mixing was measured by injecting the same amount of peptides into the reaction cell that contained the reaction buffer and subtracted. All the data were analyzed by using the Microcal Origin software. The data fit well to a single-site binding mode which provided the stoichiometric information, K_a and thermodynamic parameters including enthalpy (ΔH) and entropy (ΔS) of binding.

2.9. Surface plasmon resonance measurements

Real time binding was performed using surface plasmon resonance (SPR) at the Center for Disease Control and Prevention (Atlanta, GA, USA) on a Biacore 3000 system (Biacore AB, Uppsala, Sweden). CaM (500 nM in 10 mM sodium formate, pH 3.5) was directly immobilized onto the sensor chip CM5 using an amine coupling kit as described by the manufacturer. Synthetic peptides with varying concentrations were subsequently injected over the sensor surface at a flow rate of 50 mL/min in binding buffers (5 mM Ca, 100 mM KCl, 50 mM Tris-HCl, pH 7.4). Two minutes later, peptide-free binding buffer was injected to monitor the dissociation process. All measurements were carried out in parallel

using two cells, one with immobilized CaM and the other as blank control with the carboxylated dextran matrix deactivated. Binding of peptide to CaM-immobilized flow cells was corrected for binding to control flow cells. The sensor chip was regenerated using 10 mM glycine, pH 2.2.

2.10. Nuclear magnetic resonance spectroscopy

One-dimensional ^1H NMR spectra were recorded on a Varian 500 MHz NMR spectrometer with a spectral width of 6600 Hz. Samples of 200 to 250 μM were prepared in 20 mM PIPES-10 mM KCl, 10% D_2O at pH 6.8. La^{3+} stock solution was gradually added into the NMR sample tube. The program FELIX98 (MSI) was used to process NMR data with an exponential line broadening of 2 Hz window function and the suppression of water signal with a Gaussian deconvolution function with a width of 20.

All NMR experiments were performed using either Varian Inova 500 or 600 MHz spectrometers. NMR spectra were acquired with a spectral width of about 13 ppm in the ^1H dimension and 36 ppm in the ^{15}N dimension at 35 °C. For the (^1H , ^{15}N)-HSQC experiment, 0.5 mM ^{15}N uniformly-labeled CaM was titrated with 10-20 μL aliquots of the peptide stock solution (2.1 mM) in a buffer consisting of 10% D_2O , 100 mM KCl, 50 mM Tris-HCl with 10 mM EGTA or 10 mM Ca^{2+} . The pH values of both sample solutions were carefully adjusted to 7.4 with trace amounts of 2M KOH. NMR data were processed using the FELIX98 program (Accelrys).

For gradient diffusion experiments, spectra were collected using a

modified diffusion LED pulse sequence in 20 mM PIPES- 10 mM KCl at 25 °C as previously described (70). The diffusion constants (D) were obtained by fitting the integrated areas of arrayed spectrum with equation 6:

$$I = I_0 \exp[-(\gamma\delta G)^2(\Delta - \delta/3)D] \quad (6)$$

where γ is the gyromagnetic ratio of proton. The time between PFG pulse (Δ) and the PFG duration time (δ) were 100.5 and 5 ms, respectively. The gradient strength (G) was arrayed from 0.2 Gauss/cm to about 28.9 Gauss/cm using 50 steps. I_0 is the integrated area of selected resonances at 0 Gauss/cm and I is the integrated area of the desired resonances at each array spectrum after baseline corrections. PIPES and dioxane were used as internal references. The diffusion process of a spherical particle is governed by Stokes-Einstein's equation:

$$D = kT/6\pi\eta r_s \quad (7)$$

where k is the Boltzman constant, η is the solvent viscosity, r_s is the hydrodynamic radius.

2.11. Cells, infectious clone and site-directed mutagenesis

The infectious genomic cDNA clone Robo502 was previously described (58,59). To generate the mutated construct Robo502AA ($D^{1210}A$, $D^{1217}A$), a two-round asymmetric PCR strategy was employed to create the mutations in an amplified fragment between unique BsmI and RsrII sites at nts 3243 and 3897 of the RUB genome. The doubly digested fragment was used to replace the corresponding fragment in Robo502. The mutations were also transferred to the replicons, RUBrep-FLAG/GFP and RUBrep-HA/GFP (71), which express a P150

tagged with the FLAG or the HA epitope and the reporter protein, green fluorescent protein (GFP), in place of the SP-ORF.

2.12. *In Vitro* transcription, transfection, Northern blotting and Western blotting

Vero cells were obtained from the American Type Culture Collection and maintained at 35 °C under 5% CO₂ in Dulbecco's modified Eagle's medium (DEME, Gibco) supplemented with 5% fetal bovine serum and 10 µg/mL gentamicin. CsCl density gradient purified plasmids were linearized with EcoRI (New England Biolabs), followed by phenol-chloroform extraction and ethanol precipitation. RNA transcripts were synthesized at 37 °C in a 25-µl reaction mixture containing 40 mM Tris-HCl (pH 7.5), 1 µg DNA, 6 mM MgCl₂, 2 mM spermidine, 10 mM dithiothreitol, 1 mM of each NTP, 1 unit of RNasin (Roche Applied Science), 2 mM cap analog m⁷G(5')ppp(5')G (New England Biolabs), and 25 units of SP6 DNA-dependent RNA polymerase (Epicenter Technologies). The transcription reaction mixtures were directly used for cell transfection without DNase treatment. The transcripts were analyzed by electrophoresis of 2 µL aliquots of the reaction mixture in 0.8% agarose gel. The precise yield was determined using a UV/vis spectrometer (Shimadzu). Transfection of the RNA transcripts was done when using Vero cells at ~80% confluence in 60 mm² plates. Vero cells were washed twice with 3 mL of phosphate-buffered saline and once with 3 ml of Opti-MEMI (Gibco). 10 µL of the transcripts were subsequently mixed with 500 µL Opti-MEMI and 7 µL lipofectamine-200 (1 mg/ml). The mixture

and 1.5 mL of Opti-MEMI were added onto the plates to cover the monolayer of Vero cell. After 4 h of incubation, the transfection mixture was removed, and 4 ml of DMEM supplemented with 2% fetal bovine serum was added (transfection is designated as passage 0 or P0). Following development of significant cytopathic effect (CPE), the medium was harvested and passed twice in Vero cells (P1 and P2). Virus titers were determined by plaque assay as previously described (72,73). To determine if the D¹²¹⁰A and D¹²¹⁷A mutations were maintained following transfection and passage, plaques were picked from terminal dilution plaque assay plates of P1 culture fluid and amplified once in Vero cells. Following amplification, total infected cell RNA was extracted and the NS protease region of the genome was amplified by RT-PCR and the RT-PCR product was sequenced (74).

The expression of reporter gene GFP in living, transfected cells was monitored by a Zeiss Axioplan upright microscope with epifluorescence capability and photographed with a Zeiss Axiocam. Northern blotting was used to detect replicon-specific RNA species by following previously established protocols (71). To detect *in vivo* NS protease activity, Vero cells transfected with RUBrep-HA/GFP, RUBrepAA-HA/GFP, or RUBrepNS*-HA/GFP (containing a C¹¹⁵²A mutation of the catalytic site) were incubated at 35 °C or 39 °C. 6 hrs post-transfection, the cells were lysed and resolved by 8% SDS-PAGE, electroblotted onto nitrocellulose membranes, and probed with anti-HA monoclonal antibodies (Sigma) according to a previously established procedure (71).

2.13. Dye transfer assay

Confluent monolayers of HeLa cells grown on glass coverslips were loaded with the Ca^{2+} indicator Fura-2 AM (5 μM) in 2 ml HBSS⁺⁺ buffer (containing 1.8 mM Ca^{2+} , with added 10 mM HEPES, 5 mM NaHCO_3 , pH 7.2), then transferred to a microincubation chamber (model MSC-TD, Harvard Apparatus, Holliston, MA) as described previously (68). Imaging of intracellular Ca^{2+} was performed with a Nikon TE300 (Nikon Inc., Melville, NY) inverted microscope equipped with Nikon filter blocks for Fura-2 emission and AF594 optics (Chroma Technology Corp, Rockingham, VT), a Metaltek filter wheel (Metaltek Instruments, Raleigh, NC) housing excitation filters for Fura-2, a 75 watt xenon short arc lamp, a Hamamatsu CCD digital camera (Hamamatsu Corporation, Bridgewater, NJ), and supported on a vibration isolation table (Technical Manufacturing, Peabody, MA). $[\text{Ca}^{2+}]_i$ was measured ratiometrically ($\lambda_{340} / \lambda_{380}$) with Fura-2 throughout each experiment in the injected cell and the cells adjacent to the injected cell, and Ca^{2+} concentrations determined as described previously (75). MetaFluor software (Universal Imaging Corp., Downingtown, PA) was used for data collection. Micropipettes (borosilicate glass capillaries: 1 mm O.D., 0.75 mm I.D, 100 μm internal microfilament; Dagan Corporation, Minneapolis, MN) were pulled on a Flaming/Brown-type pipette puller (P-87; Sutter Instruments, Novato, CA). Micropipettes had tip diameters of $<1 \mu\text{m}$ and resistances of $\sim 100\text{-}300 \text{ M}\Omega$ when filled with AlexaFluor 594 (1 mM) dissolved in deionized water. The micropipette was positioned with a low-drift hydraulic micromanipulator (MW-3; Narishige, Greenvale, NY), and AlexaFluor

594 was microinjected iontophoretically using a train of 5-ms current pulses applied every 100 ms for 60 s (3 s total injection time) at ambient temperature. If the micropipette became plugged, it was replaced with a new micropipette, and the data from such a partial injection were excluded from the analysis. Current was generated with a Duo 773 (World Precision Instruments, Sarasota, FL). Current duration, magnitude, and polarity were controlled with an A310 Accupulser pulse generator (World Precision Instruments, Sarasota, FL). Digitized images of AlexaFluor594 cell-to-cell transfer were recorded 2 min following the iontophoretic injection of fluorescent dye. A sustained elevation in $[Ca^{2+}]_i$ was effected by adding 1 μ M ionomycin to the medium, then 2 min later increasing the extracellular $[Ca^{2+}]$ from 1.8 mM to 21.8 mM. All these works were carried out by Dr. Monica Lurtz.

2.14. Bioinformatic tools and homology structure modeling

Multiple sequence alignments and phylogenetic analysis. 1904 proteins with potential canonical EF-hands and 84 proteins with pseudo-EF-hands from SwissProt encompassing 66 distinct subfamilies of EF-hand proteins were included in our EF-hand databases. Typical members of each subfamily were collected to generate a sub-database for multiple sequence alignments and phylogenetic analysis. Multiple sequence alignment (MSA) was performed using the ClustalW program with a gap open penalty of 10 and gap extension penalty set at 0.5.(76) The same program was applied to generate N-J tree for further display by the TreeView program.(77)

Generation of profile HMM and patterns. Profile HMM (Hidden Markov

Models) was generated from multiple sequence alignment results using HMMER by choosing both hmmbuild and hmmcalibrate algorithms. The statistical profile is subsequently visualized as HMM logo using LogoMat-M.(78,79) The EF-hand patterns were generated by taking into account highly conserved residues within both the pseudo and canonical EF-hand motifs.

Evaluation of canonic and pseudo EF-hand pattern. The precision, sensitivity and positive predictive values (PPV) of canonical EF-hand patterns loop, eloop, loopf, and eloopf were compared to that of the pattern PS00018. A total of 170 hits, including true positive, false negative, and false positive, were randomly selected from the results of PS00018 and set as the sub-database for comparison. The newly generated pseudo EF-hand patterns, as well as the well-established S100 pattern PS00303, were used to search for possible pseudo-EF-hand Ca^{2+} binding domains against major protein sequence databases such as SwissProt, iProClass and NCBI reference sequences (RefSeq). The Ca^{2+} -binding properties of the proteins in the selected dataset have been experimentally verified and the prediction is compared with the verified information to determine the true positive, true negative, false negative, and false positive. The methods are then applied to predict proteins with unknown Ca^{2+} -binding properties in bacterial genomes. For statistical analysis, the precision, sensitivity, and PPV were determined as follows:

$$precision = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (9)$$

$$PPV = \frac{TP}{TP + FP} \quad (10)$$

(TP: true positive; TN: true negative; FN: false negative; FP: false positive)

The Ca^{2+} -binding motif signatures developed in our laboratory (<http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm>) as well as the pattern PS00018 from PROSITE (<http://us.expasy.org/cgi-bin/nicedoc.pl?PDOC00018>) were used to scan through the genome of rubella virus for any potential EF hand Ca^{2+} -binding site (6). Sequence alignments were conducted by using the program ClustalW. The secondary structure was predicted based on the consensus prediction results using programs PSIPRED, JPREP, and PHD (80-82). The homology modeling of the engineered protein and the protease was constructed using the comparative structure modeling SWISS-MODEL (83). CD2 (PDB code: 1hng) and the leader protease of foot-and-mouth disease virus (PDB code: 1qmy) were used as the templates for structure modeling based on their available high resolution structure (61,84). Prediction of Ca^{2+} -binding sites in the modeled structure was conducted by using the program GG, a computational algorithm developed in our laboratory on the basis of the geometric description, graph theory and key structural and chemical features associated with calcium binding in proteins (85).

Bioinformatic analyses and prediction of CaM-binding site in connexins.

The topology and orientation of transmembrane regions of the rodent connexin 43 and sheep Cx44 were predicted on the basis of the consensus results using four different programs including SOSUI (86), TMHMM (87), MEMSAT (88), and

HMMTOP (89). The potential CaM binding sites were predicted using the CaM target database based on several criteria, including the distribution of hydrophobic and basic residues, propensity to form helix, and hydrophobicity of the sequence, that are common to more than 100 CaM target sequences (90).

2.15. Determination of zinc amount by colorimetric PAR assay

PAR colorimetric assay with slight modification was used as described (91). Protein samples were pretreated by Chelex 100 to remove background metal ions and then digested in a total volume of 100 μ L with 100 μ g protease K (Sigma) per mL in HSD buffer (50 mM HEPES-KOH, 200 mM NaCl, 5 mM dithiothreitol [DTT]), pH 7.0 at 56 °C for 30 min. Subsequently, an identical volume of HSD containing 5 mM iodoacetamide (IAM; Sigma) and 200 μ M 4-(2-pyridylazo)resorcinol (PAR, Sigma) was added. The absorption ranging from 300 nm to 600 nm was measured by RF-1501 UV spectrometer (Shimadzu). HSD containing 0-15 μ M standard ZnCl_2 solution was used to create a standard curve. The amount of protein in the samples to be analyzed for metal content was adjusted so that the amount of metal released was in the linear range of the standard curve. Carbonic anhydrase that contained one zinc binding site is used as positive control.

3. Prediction of prokaryotic EF-hand and EF-hand like Ca^{2+} -binding proteins

Since the delineation of the EF-hand motif in 1973, the family of EF-hand proteins has thus far expanded to include at least 66 subfamilies. As noted in section 1.2, there are two types of EF-hand motifs: the canonical and pseudo EF-hands. The major difference between these motifs lies in the Ca^{2+} -binding loop: the 12-residue canonical EF-hand loop binds Ca^{2+} mainly via sidechain carboxylates or carbonyls (loop sequence positions 1, 3, 5, 12), whereas the 14-residue pseudo EF-hand loop chelates Ca^{2+} primarily via backbone carbonyls (positions 1, 4, 6, 9) (Fig. 3. 1). The residue at the $-X$ axis coordinates the Ca^{2+} ion through a bridged water molecule. The EF-hand loop has a bidentate ligand (Glu or Asp) at axis $-Z$.

With the continuing expansion of genomic information, many efforts have been made to predict the continuous EF-hand Ca^{2+} -binding proteins and to understand the role of Ca^{2+} in biological systems. Pattern (motif signature) search is one of the most straightforward ways to predict continuous EF-hand Ca^{2+} -binding sites in proteins. Based on the sequence alignment results of canonical EF-hand motifs, especially the conserved side chains directly involved in Ca^{2+} binding, a pattern PS00018 (<http://us.expasy.org/cgi-bin/nicesite.pl?PS00018>) has been generated to predict canonical EF-hand sites. Alternative patterns have also been proposed with the addition of other

conserved residues in the motif (92,93). For the pseudo EF-hand loop, however, each type of amino acid may serve as potential Ca^{2+} binding ligands because of the use of main chain, which makes prediction solely from the sequences relatively difficult. To circumvent this problem, the prediction of pseudo EF-hand sites was achieved by detecting the canonical EF-hands based on the assumption that all the pseudo EF-hands are paired by a C-terminal canonical EF-hand. The currently available pattern PS00303 from EXPASY website (<http://us.expasy.org/cgi-bin/nicedoc.pl?PDOC00275>) predicts the S100 type Ca^{2+} binding proteins by spanning the C-terminal canonical EF-hand motifs. It is worth pointing out that the prediction results obtained using this strategy do not directly provide the sequence of the pseudo EF-hand Ca^{2+} binding loop.

Toward our goal of predicting and understanding the role of Ca^{2+} in biological systems (denoted as calciomics), we report herein our progress in identifying EF-hand and EF-like motifs from the primary sequences. A series of patterns were generated by taking advantage of the metal binding properties of currently available EF-hand proteins and considering the helical structural context around the Ca^{2+} -binding loop. We modified the pattern PS00018 by allowing more choices (Glu, Gln, and Ser) at position 1 (axis X) and adding constraints at the flanking helical regions. By easing the constraints at the C-terminal canonical EF-hand, and simultaneously incorporating reserved residues in the N-terminal pseudo EF-hand, we also generated a modified pattern for the prediction of pseudo EF-hand sites. Compared with the pattern PS00303, the new pattern, reflecting conserved genomic information in both the N- and C-

terminal EF-hands, significantly improved the predictive accuracy and sensitivity. Finally we report our analysis of EF-hand proteins in bacterial genomes using the prediction method we developed. Our prediction results indicate that no pseudo EF-hand protein is found in bacteria, which provides an additional piece of evidence suggesting that pseudo EF-hand motif likely evolved later than canonical EF-hand motifs.







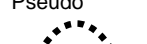



3.1. Pattern development

Based on the sequence alignment results and the statistical profile corresponding to each type of EF-hand Ca^{2+} -binding site (Fig. 3.1), several patterns reflecting the most conserved information at particular positions have been developed and summarized in Table 3.1. Among them, patterns 1-4, in addition to the commonly used pattern PS00018, can be used for the prediction of canonical EF-hand Ca^{2+} -binding sites with varying degrees of constraints on the sequences. Patterns PS00303 and PC (abbreviation of pseudo and canonical EF-hands pattern) can be used for the prediction of Ca^{2+} -binding motifs within S100 and S100-like proteins. The patterns for EF-hand-like proteins are applied to the prediction of EF-hand-like Ca^{2+} -binding motifs with the loop length ranging from 10 to 15 residues.

3.2. Canonical EF-hand motif

The widely applied pattern PS00018 has a stringent restraint at loop sequence position 1 (axis X) that only allows Asp, although Asn or Ser also

Table 3.1. Summary of patterns to predict EF-hand proteins.

Pattern number	Pattern name	Conservation positions	Motif signature
To predict Canonical EF-hand			
1	eloopf 	Canonical EF-loop and both flanking helices	x-{DNQ}-x(2)-{GP}-{ENPQS}-x(2)-{DPQR}-[DNS]-x-[DNS]-{FLIVWY}-[DNESTG]-[DNQGHRK]-{GP}-[LIVMC]-[DENQSTAGC]-x(2)-[ED]-[FLYMVIW]-x(2)-{NPS}-{DENQ}-x(3)
2	eloop 	Canonical EF-loop and the entering helix	x-{DNQ}-x(2)-{GP}-{ENPQS}-x(2)-{DPQR}-[DNS]-x-[DNS]-{FLIVWY}-[DNESTG]-[DNQGHRK]-{GP}-[LIVMC]-[DENQSTAGC]-x(2)-[ED]
3	loopf 	Canonical EF-loop and the exiting helix	[DNS]-x-[DNS]-{FLIVWY}-[DNESTG]-[DNQGHRK]-{GP}-[LIVMC]-[DENQSTAGC]-x(2)-[ED]-[FLYMVIW]-x(2)-{NPS}-{DENQ}-x(3)
4	loop 	Canonical EF-loop	[DNS]-x-[DNS]-{FLIVWY}-[DNESTG]-[DNQGHRK]-{GP}-[LIVMC]-[DENQSTAGC]-x(2)-[ED]
5	PS00018 ^a 	Canonical EF-loop	D - x - [DNS] - {ILVIFYW} - [DENSTG] - [DNQGHRK] - {GP} - [LIVMC] - [DENQSTAGC] - x(2) - [DE] - [LIVIFYW]
To predict Pseudo EF-hand			
6	PC 	Both helices of the pseudo EF-motif, both helices and the loop of the paired canonical EF-motif	[LMVITNF]-[FY]-x(2)-[YHIVF]-[SAITV]-x(5,9)-[LIVM]-x(3)-[EDS]-[LFM]-[KRQLE]-x(20,28)-[LQKF]-[DNG]-x-[DNSC]-x-[DNK]-x(4)-[FY]-x-[EKS]
7	Pseudo 	Both helices of the pseudo EF-motif	[LMVITNF]-[FY]-x(2)-[YHIVF]-[SAITV]-x(5,9)-[LIVM]-x(3)-[EDS]-[LFM]-[KRQLE]
8	PS00303 ^b 	Both helices and the loop of the paired canonical EF-motif	[LIVIFYW](2) - x(2) - [LK] - D - x(3) - [DN] - x(3) - [DNSG] - [FY] - x - [ES] - [FYVC] - x(2) - [LIVMFS] - [LIVMF]
To predict EF-hand-like			
9	Excalibur 	The 10-residue loop	D-x-D-x-D-G-x(2)-C-E
10	EF-hand-like 	The loop	D - x - [DNS] - {ILVIFYW} - [DEN] - G - {GP}-x(5, 6)-[DE]

^a <http://us.expasy.org/cgi-bin/nicesite.pl?PS00018>^b <http://au.expasy.org/cgi-bin/nicesite.pl?PS00303>

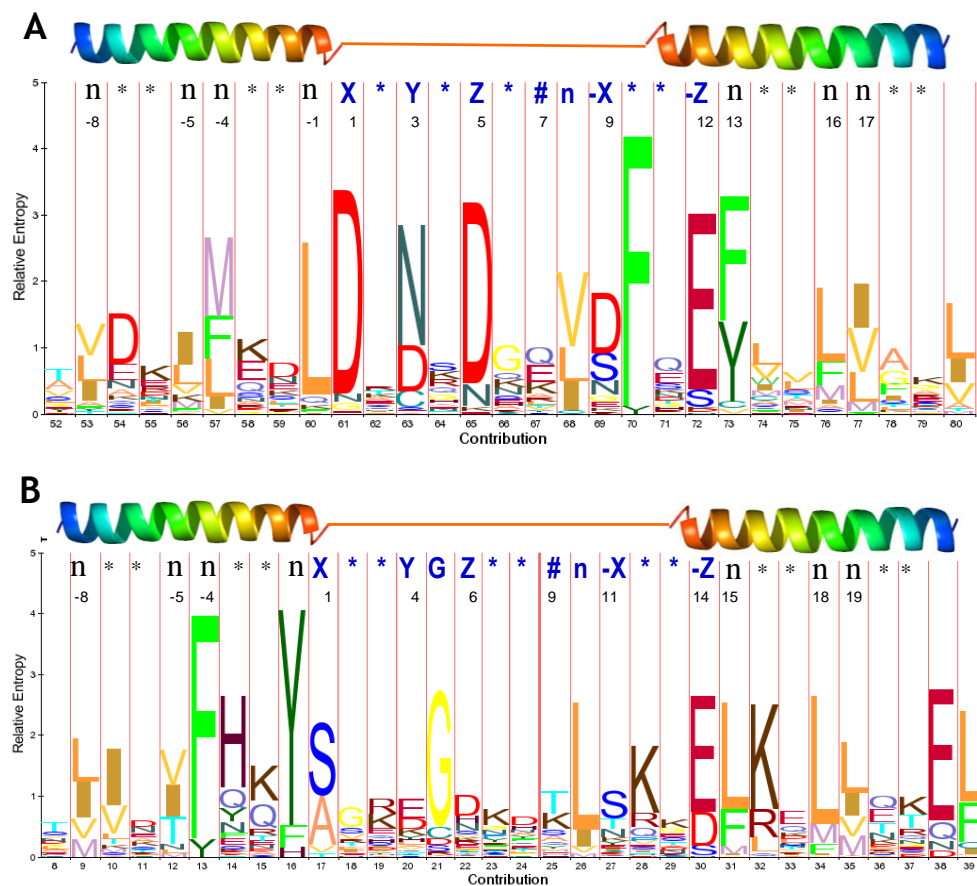


Figure 3.1. Consensus sequence of canonical EF-hand (A) and pseudo EF-hand domains (B) drawn based on profiles HMM using LogoMat-M (<http://logos.molgen.mpg.de/cgi-bin/logomat-m.cgi>). n: the hydrophobic residues within the flanking helices. #: the potential Ca^{2+} binding ligands involving the mainchain carbonyl groups.

occupy the position in a few EF-hands (Ca^{2+} and integrin binding protein 2 (Q9Z309), CaBPE63-1 (P48593), rat CaM (pdb code: 3cln)). The pattern PS00018 focuses solely on the loop region and does not reflect conserved information within the flanking regions. To improve the pattern PS00018, we incorporated the diverse features of the flanking structural contexts and developed patterns catering to different constraints on EF-hands. Based on multiple sequence alignment results on over 1,000 canonical EF-hands from the SwissProt protein sequence database, constraints on both flanking helices and the 12-residue loop were well defined in separate patterns. As shown in Table 3.1, three patterns were derived: a) $x\text{-}\{\text{DNQ}\}\text{-}x(2)\text{-}\{\text{GP}\}\text{-}\{\text{ENPQS}\}\text{-}x(2)\text{-}\{\text{DPQR}\}$ for the entering helix; b) $[\text{DNS}]\text{-}x\text{-}[\text{DNS}]\text{-}\{\text{ILVFYW}\}\text{-}[\text{DNESTG}]\text{-}[\text{DNQGHRK}]\text{-}\{\text{GP}\}\text{-}[\text{LIVMC}]\text{-}[\text{DENQSTAGC}]\text{-}x(2)\text{-}[\text{ED}]$ for the Ca^{2+} -binding loop; and c) $[\text{FLYMVIW}]\text{-}x(2)\text{-}\{\text{NPS}\}\text{-}\{\text{DENQ}\}\text{-}X(3)$ for the exiting helix. As revealed by the sequence alignment, hydrophobic residues are favored at positions -1, -4, -5 and -8 in the entering helix and at positions 13, 16, and 17 in the exiting helix (Fig. 3.1A). Hence, hydrophilic residues or residues tending to interrupt helical structure were excluded at these positions. The prediction of loop only (b), E-loop (a+b), loop-F (b+c) and E-loop-F (a+b+c) can then be achieved using the patterns in combination. This strategy provides an alternative way to perform prediction of EF-hands as well as EF-hand-like sites with deviations at the flanking regions.

Fig. 3.2A shows the statistical results of the prediction of canonical EF-hand motifs using the pattern eloopf. According to the Prosite documentation PDOC00018, the pattern PS00018 results in more than 2,000 hits in the

SwissProt database. To compare our patterns with PS00018, a total of 170 protein sequences were randomly selected from the SwissProt database. Of these, 119 are true canonical EF-hand proteins with experimental validation, while 51 are not. With these sequences as the testing database, prediction results show that patterns 1-5 have similar sensitivity while the precision and PPV of patterns 1-3 increased by 10% to 20% compared to the pattern PS00018. Fewer false positive hits were detected when using the patterns 1-3. Hence, additional constraints on the flanking regions enhance the overall accuracy of prediction and the true positive predictions.

3.3. Pseudo EF-hand motif

As listed in Table 3.2, pseudo EF-hands are mostly found in the S100 protein family and among members of the “fused gene” family, such as trichohyalin, horenin and repetin (94-99). The small, acidic S100 protein, calbindin_{D9K}, carries two distinct EF-hands: a canonical EF-hand at the C terminus and a pseudo EF-hand motif at the N-terminus. The canonical EF-hands are highly-conserved among all the S100 proteins (Fig. 3.3). However, there is significant sequence variation in the Ca²⁺-binding loop of the pseudo EF-hand (Fig. 3.1B). The pseudo EF-hand within the S100A10 even loses the capacity to bind Ca²⁺ ion due to the lack of chelating ligands (100). Therefore, one important concept that must be kept in mind is that the prediction of pseudo EF-hand does not presume the capability of binding Ca²⁺ ion.

Using the multiple sequence alignment, we analyzed the Ca²⁺-binding

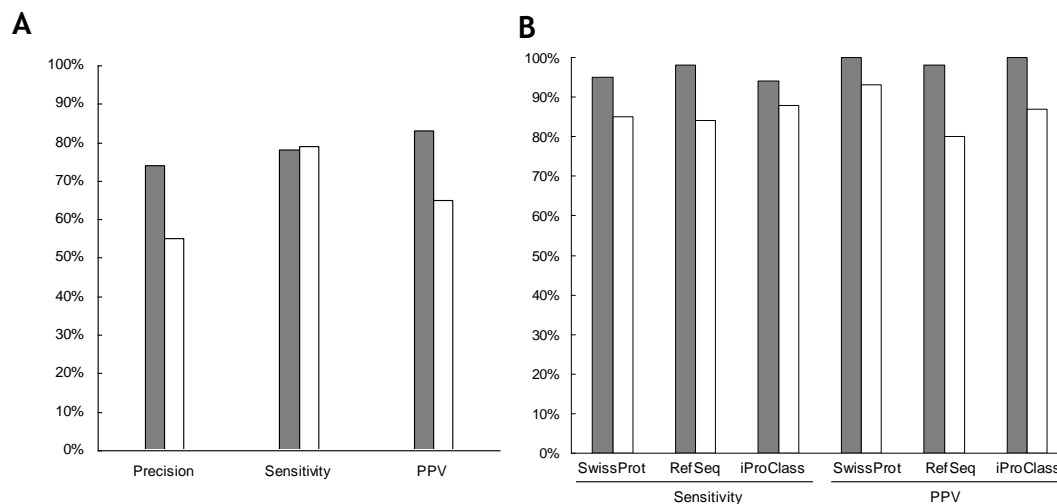


Figure 3.2. Statistical analysis on prediction performance of patterns. (A) Prediction results using patterns eloopf (gray bar) and PS00018 (open bar). The patterns were applied to search for canonical EF-hand proteins in a test database containing 170 proteins. (B) Prediction results using the pseudo EF-hand pattern PC (gray bar) and PS00303 (open bar). Both patterns were used to search for potential pseudo EF-hand proteins against major protein sequence databases SwissProt, NCBI RefSeq, and iProClass.

ligands of all pseudo EF-hands with known structures in Protein Data Bank (Table 3.2). Based on the statistical results, a profile HMM and the resultant HMM logo were built (Fig. 3.1B). Of the Ca^{2+} -binding ligands, Ser and Ala are preferred at loop position 1 (X); Glu dominates at both positions 4 (Y) and 14 (-Z); Gly and Leu are preferred at the positions 5 and 10, respectively; Asp is most frequently found at position 6 (Z); and Thr and Lys reside equally at position 9 (-Y). By integrating highly conserved residues located at the flanking helices (positions -1, -4, -5, 15, and 16), we generated a pattern (Table 3.1, pattern pseudo) for the prediction of pseudo EF-hand Ca^{2+} -binding site. The pattern PC was further developed by incorporating the conserved signature in the downstream canonical EF-hand (positions -1, 1, 3, 5, 10, and 12).

To assess the performance of the developed patterns, we applied the patterns against major protein sequence databases such as SwissProt, iProClass (including PIR, trEMBL) and NCBI reference sequences (RefSeq). Fig. 3.2B shows the comparison of the pattern PC and the pattern PS00303. A notable limitation of the pattern PS00303 is its failure to predict pseudo EF-hands within S100A13, S100A14, S100A16, S100A17 and S100P from some species due to stringent restraints at the C-terminal EF-hand motif. Moreover, since the prediction is based on the C-terminal canonical EF-hand, the prediction of PS00303 includes more false positive hits from the Ca^{2+} -binding proteins possessing only canonical EF-hands, such as calneuron 1 and Ca^{2+} -binding protein 7. In comparison with PS00303, 12.5% more true positive hits, in average, resulted with the pattern PC in the three databases. Meanwhile, the

Table 3.2. List of pseudo EF-hand proteins.

Protein	Synonyms	PDB codes	Accession Number (species*)
S100A1	S-100 protein alpha chain	1k2h	P02639 (b), P23297 (h), P56565 (m), P35467 (r), Q7LZT1 (weatherloach)
S100A2	S-100L, CAN19		P10462 (b), P29034 (h)
S100A3	S-100E	1kso	P62818 (m), P62819 (r), P33764 (h)
S100A4	Metastasin, Calvasculin, Mts1 protein, 18A2, PEL98, Placental calcium-binding protein homolog, P9K, Nerve growth factor induced protein 42A	1m31	P35466 (b), P26447 (h), Q9TV56 (d), P07091 (m), P05942 (r)
S100A5	S-100D		P63084 (m), P63083 (r), P33763(h)
S100A6	Calcyclin, Prolactin receptor associated protein, 5B10	1a03, 1cnp, 1jwd, 1k8u, 1k96, 1k9k, 1k9p, 2ncp	P14069 (m), P05964 (r), P06703 (h), P30801 (rb), Q98953 (c), O77691 (hs)
S100A7	Psoriasin, Dermal allergen BDA11, Allergen Bos d 3	1psr, 2psr, 3psr	P31151 (h), Q28050 (b)
S100A8	Calgranulin A, Neutrophil cytosolic 7 kDa protein P7, Mgrin inhibitory factor-related protein 8, Chemotactic cytokine CP-10, MRP-8	1mr8	P28782 (b), P05109 (h), P27005 (m), P50115 (r)
S100A9	Calgranulin B, Neutrophil cytosolic 23 kDa protein, Mgrin inhibitory factor-related protein 14 (MRP-14), P23, BEE22, P14, Leukocyte L1 complex heavy chain, Calprotectin L1H subunit	1irj	P28783 (b), P06702 (h), P50117 (rb), P31725 (m), P50116 (r)
S100A10	Calpactin I light chain, p10 protein, p11, Cellular ligand of annexin II, Nerve growth factor induced protein 42C	1a4p, 1bt6	P60902 (b), P60903 (h), P04163 (p), P620504 (rhesus macaque), P08207 (m), P05943 (r), P27003 (c), P27004 (African clawed frog)
S100A11	Endothelial monocyte-activating polypeptide, Calgizzarin, S100C, MLN 70, EMAP	1nsh, 1qls	P31949 (h), P24480 (rb), P50543 (m), Q6B345 (r), P31950 (p), P24479 (c)
S100A11P	Putative S100 calcium-binding protein A11 pseudogene		O60417 (h)
S100A12	Calgranulin C, CAGC, Calcium-binding protein in amniotic fluid 1, CAAF1, RAGE binding protein, Neutrophil S100 protein, p6	1e8a, 1gqm, 1odb	P79105 (b), P80310 (p), P80511 (h), O77791 (rb)
S100A13	8 kDa amlexanox-binding protein		P79342 (b), Q99584 (h), P97352 (m)
S100A14	S114		Q9HCY8 (h), Q9D2Q8 (m)
S100A15			Q86SG5 (h)
S100A16	S100F		Q96FQ6 (h)
S100A17	clone:5430400H23 product:hypothetical EF-hand/S-100/ICaBP type calcium binding protein		Q9D3P1 (m)
S100B	S-100 protein beta chain	1b4c, 1cfp, 1mho, 1dt7, 1mwn, 1psb, 1qlk, 1sym, 1uwo	P50114 (m), P04631 (r), P04271 (h), P02638 (b)
S100G	Vitamin D-dependent calcium-binding protein intestinal, Calbindin D9K, Cholecalciferin	1b1g, 1boc, 1bod, 1cb1, 1cdn, 1clb, 1d1o, 1ht9, 1ig5, 1igv, 1kcy, 1kqv, 1ksm, 1n65, 2bca, 2bcb, 3icb, 4icb	P29377 (h), P02632 (p), P02633 (b), P51964 (c), P02634 (r), P97816 (m)
S100H	Putative S100 calcium-binding protein H_NH0456N16.1		Q9UDP3 (h)
S100P		1ozo, 1j55	P25815 (h)
S100Z			Q8WVG8 (h)
Hornerin			Q86YZ3 (h), Q8VHD8 (m)
Ictacalcin			Q91061 (channel catfish)
MRP-126			P28318 (c)
Reptin			P97347 (m)
Trichohyalin			Q07283 (h), P37709 (rb), P22793 (s)

* The abbreviation for species: b, bovine; c, chicken; d, dog; h, human; hs, horse; m, mouse; r, rat; rb, rabbit; s, sheep



Figure 3.3. Multiple sequence alignment of the pseudo EF hand proteins. 25 entries representing 84 proteins with pseudo EF-hand motif from SwissProt were used for alignment. The N-terminal pseudo EF-hand and C-terminal canonical EF-hand motifs with typical helix-loop-helix structure are shown at the top (box: helical structure; line: loop region). The oxygen atoms chelating the Ca^{2+} ion adopts a pentagonal bipyramid geometry. (n: the hydrophobic residues within the flanking helices. #: the potential Ca^{2+} binding ligands involving the mainchain carbonyl groups. bi: the bidentate chelating ligands.)

false positive hits were reduced by 10.5% on average. The average sensitivity and PPV of the pattern PC were 96 and 99%, respectively, which were 10 and 13% higher than those of the pattern PS00303 (Fig. 3.2B). The pattern PC was able to identify the pseudo EF-hand motifs in at least 3 more subgroups of S100 proteins including S100A13, S100A14, and S100A16 than the pattern PS00303 was. The first half of the pattern PC (or pattern pseudo) could be of great advantage in predicting S100-like proteins with deviations in the downstream canonical Ca^{2+} -binding loop or in predicting partially-characterized, incomplete hypothetical proteins. For instance, the pseudo EF-hand motifs in the novel Ca^{2+} -binding protein p26olf (named as the protein from frog olfactory epithelium) (101,102) could not be predicted by either PS00303 or PC since the C-terminal EF-hand contains an atypical EF-hand motif with a 4-residue insertion. However, without constraints on the C-terminus, the pattern pseudo (Table 3.1) can easily detect them.

3.4. Properties of EF-hand like motif

With the overall structural geometry of the Ca^{2+} coordination remaining conserved, “EF-hand-like” motif refers to the one containing the following deviations from the canonical EF-hand: (1) the length of the Ca^{2+} -binding loop is shorter or longer than 12 residues and/or (2) the secondary structure elements of the flanking regions are not two helices. The first deviation can be represented in the motif signature by varying the length of the loop region (Patterns 9-10 in Table 3.1). However, the structural deviation in the flanking regions can hardly be

Table 3.3. EF-hand-like Ca^{2+} -binding proteins with known structure.

PDB (resolution, Å) and protein name	Deviation ^a	Sequence ^b (structure note)	Role of Ca ²⁺	Organism
1acc (2.10) Anthrax protective antigen	3	168 STSAGPTVPD <u>R</u> NDNGIPDSLEVEGYTVDVKNKR (loop-loop-helix) 200	Structural	<i>Bacillus anthracis</i>
1daq Dockerin	3	1 MSTKLYG <u>D</u> VNDDGKVNSTDAVALKRYVLR (loop-loop-helix) 29	Structural	<i>Clostridium thermocellum</i>
1gcg (1.90) Glucose/galactose receptor	4+5	30 GISINTDNADL <u>N</u> EDGRVNSTDLGILKRY (loop-loop-helix) 58	Structural	<i>Saimonella typhimurium</i>
1h71 (2.10) Alkaline protease	4+5	49 QLTRSGASWH <u>D</u> LNNDGVINLT ^Y FLTA... <u>D</u> (loop-loop-strand) 114	n/a	<i>Psuedomonas sp.</i>
1jv2 (3.10) Integrin αVβ3	1+3+4	275 YFGFSVAATD <u>I</u> NGDDYADVFIGAPLFM (strand-loop-strand) 301	Regulatory	<i>Homo sapiens</i>
1kap (1.64) Alkaline protease	3+4	446 ASKAGSLAI <u>D</u> FSGD <u>A</u> HADFAINLIGQA (strand-loop-strand) 472	n/a	<i>Psueomonas aeruginosa</i>
1kwh (2.00) Alginate-binding protein	1+4	162 TVLKAFKEK <u>D</u> PNGNGKADEVPFIDRHPDE (helix-loop-loop) 190	Regulatory	<i>Sphingomona s sp.</i>
1hwj (2.50) Glucanotransferase	1	4 YQIYVRSFRD <u>G</u> NLDGVGD <u>F</u> RGLKNAVSYL (helix-loop-helix) 32	n/a	<i>Thermotoga maritima</i>
1qut (2.44) Slh35	2	189 PSSYKQYAV <u>D</u> FGDGHINLWDPVDAIGSVANY (helix-loop-helix) 220	Structural	<i>Escherichia coli</i>

- a. 1. Shorter loop. 2. Longer loop. 3. Entering helix missing. 4. Exiting helix missing. 5. Distal located ligands
- b. Ligand residues are underlined

predicted merely from the sequences. Therefore, we conducted a retrospective search of EF-hand-like proteins in the PDB database (Table 3.3). Four classes of EF-hand-like motifs are currently observed. The first class has a shorter loop (as seen in Excalibur) that contains a conserved 10-residue DxDxDGxxCE motif. The cysteine in the sequence may facilitate the orientation of the loop toward Ca^{2+} binding by forming disulfide bonds(23). The second class has a longer loop as seen in Slf35 (PDB code: 1qut), a soluble fragment of lytic transglycosylase B from *Escherichia coli* that has a 15-residue Ca^{2+} -binding loop flanked by two helices (103,104). The third class lacks the entering helix as seen in protective antigen (PDB code: 1acc) from *Bacillus anthracis* (27) and dockerin from *Clostridium thermocellum* (PDB code: 1daq) (28). The fourth class lacks the exiting helix as seen in alginate-binding protein (PDB code: 1kwh) from *Sphingomonas sp* (26). Some EF-hand-like proteins even infringe the EF-hand paradigm by possessing two or more types of deviations (Table 3.3).

3.5. EF-hand proteins in the bacterial genomes

To understand the roles of Ca^{2+} in bacteria, we predicted putative EF-hand proteins in the bacteria genomes from the Non-Redundant REference protein database (NREF) (105). No pseudo-EF-hand motif was predicted using the pattern PC. A total of 467 EF-hand motifs in 397 entries of proteins were predicted using the pattern eloopf (Table 3.1) for the canonical EF-hand motifs (Table 3.4). There are 39 proteins that contain multiple EF-hand motifs ranging from 2 to 6. The other 358 proteins were predicted to contain mononuclear EF-

hands. The roles of Ca^{2+} in most of these proteins are yet to be characterized.

The 39 proteins with multiple EF-hand motifs, among which 16 proteins have been summarized before(4), are implicated in a variety of cellular activities, including Ca^{2+} homeostasis (106-108), chemotaxis (24,109,110), scaffold protein binding (111), resistance to acid stress (112,113) and so on. According to the sequence homology and assuming these proteins evolved from a common ancestor, they could be further classified into three major phylogenetic groups (Fig. 3.4). The first group includes calerythrin (*Saccharopolyspora erythrea*, P06495), calsymin (*Rhizobium etli*, AAG21376), putative glycosyl hydrolase (*Bacteroides fragilis*, NF02360737), α -xylosidase (*Bacteroides thetaiotaomicron*, NF01244792), and putative Ca^{2+} binding proteins from the Gram-positive bacterial genus streptomyces (*Streptomyces ambofaciens*, BAB19055; *Streptomyces coelicolor*, CAB76018, NP_628579, CAC16980). Calerythrin is the first characterized prokaryotic CaM-like protein possessing three canonical and an atypical EF-hand motif (114,115). Two of the three high-affinity sites cooperatively chelate the metal ions and the apo protein adopts a molten globule state conformation. It may function as Ca^{2+} buffer or transporter (116). Being highly homologous to calerythrin, several members in this group (BAB19055, CAB76018, NP_628579, CAC16980, and NF02549883) are expected to adopt similar Ca^{2+} -dependent structural and biological behavior (117). Another protein calsymin was implicated in symbiotic nitrogen fixation (118). It contains three repeated homologous domains, each of which possesses two EF-hand motifs. Extracellular polysaccharide-degrading enzymes, putative glycosyl hydrolase and

α -xylosidase are involved in the metabolism of the bacterial wall. The Ca^{2+} binding in the mesophilic xylanase and other members (NF02715683 and NF02518859) may protect the proteins from enzyme attack and thermal denaturation (119).

The protein functions and the role of Ca^{2+} in the second group are not well understood except for the acid shock protein (AAB69346) from the Gram-negative bacterial genus *Brucella*. This protein is actively synthesized in response to the low pH to facilitate adaptation to acidic environments (112). In addition, the putative EF-hand protein NF01724660 may link to the Ca^{2+} -induced aggregation of the sulfate-reducing bacteria *Desulfovibrio* (120).

The third group encompasses dockerin (*Clostridium acetobutylicum*, NF00465242, NF00464378), bacterial transaldolase (*Synechocystis* sp. P72797), adhesin (*Vibrio vulnificus*, NF01147763), and others with unknown functions. Dockerin is involved in the degradation of the plant cell wall by incorporating glycosyl hydrolase into the extracellular cellulose complex “cellulosome” via interaction with the cohesion domain (111). Ca^{2+} induces the folding of dockerin (121). Bacterial transaldolase, like its eukaryotic homologues, is involved in the metabolism of glucose (122). No explanation has thus far been offered for the unique presence of EF-hand motifs in this bacterial enzyme. Autotransporter adhesin, a prototype of the adhesin family, mediates the specific attachment of bacteria to target cells (123). The binding of Ca^{2+} would probably provoke the efficient interaction and facilitate the attachment.

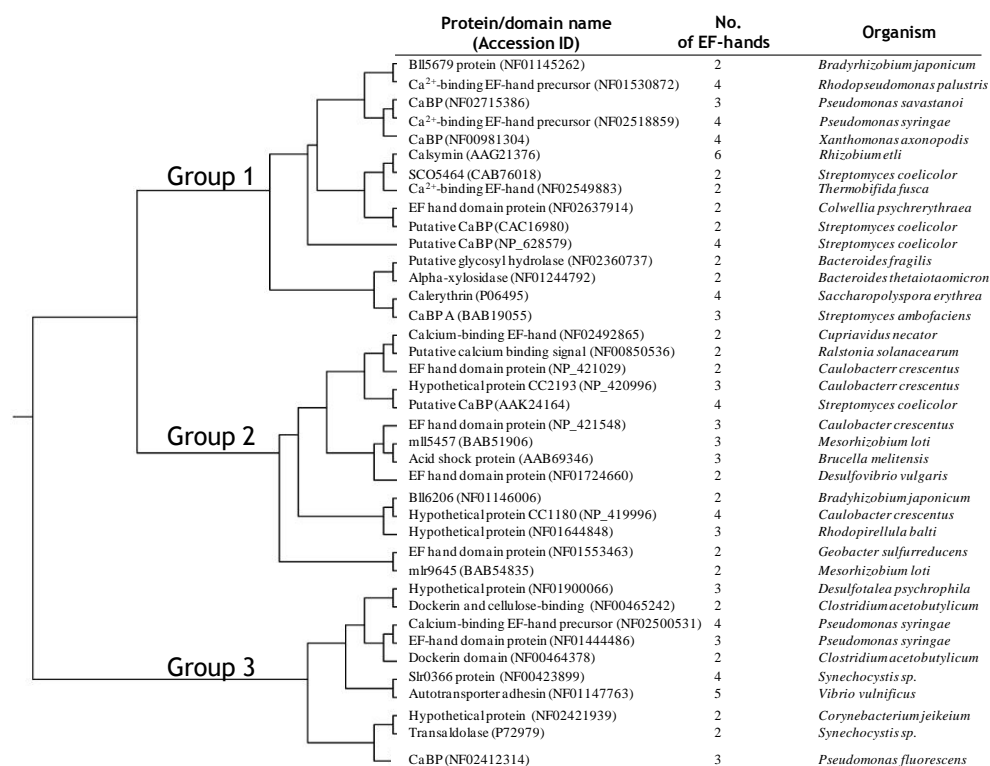


Figure 3.4. Phylogenetic analysis of prokaryotic proteins containing multiple EF-hand motifs ranging from 6 to 2. The 39 proteins can be classified into three groups (see Section 3.5 for details).

3.6. Possible roles of single EF-hands

The 358 predicted proteins containing single EF-hand motifs are in 162 complete or incomplete bacterial genomes in the PIR-NREF database (<http://pir.georgetown.edu/cgi-bin/nfspecies.pl>). They are distributed across in the majority of bacterial species (Table 3.4). These proteins are implicated in a wide range of cellular processes such as drug resistance (multiple drug resistance protein, multidrug efflux transporter), ion and nutrients transporting (K^+ -transporting ATPase B, Na^+ -solute symporter, hemin ABC transporter, cation efflux system protein), nucleic acid modification and metabolism (tRNA synthetase, ribonuclease G, exodeoxyribonuclease V gamma chain, RNA polymerase beta subunit, ATP-dependent DNA helicase, DNA polymerase tau subunit, DNA gyrase subunit A, DNA methyltransferase), transcriptional regulation (transcriptional regulator), stress response (DnaK, acid shock proteins, heat shock protein HspG), chemotaxis (CheV, histidine kinase HAMP region), energy and nutrients metabolism (GTP-binding protein, AMP nucleosidase, aminotransferase, acetyltransferase), redox reaction (flavodoxin oxidoreductase, thio-disulfide isomerase, iron-sulfur cluster binding protein, thioredoxin reductase), and cell wall modification and degradation (chitinase C, glycosyl hydrolase, exopolysaccharide synthesis protein, probably secreted sialidase, putative surface anchored protein).

Among all of these matches, ATP-binding cassette (ABC) transporter and Shr are of particular interest considering their important roles in bacterial activities and the possible implication of Ca^{2+} in the biological context. ABC

transporter couples the hydrolysis of ATP to the transport of various molecules including sugars, ions, antibiotics, and peptides across the cell membrane (124-126). Shr in *streptococcus* encodes a large hydrophilic protein (putative Fe³⁺-siderophore transport) that has no significant homologues in bacteria but shares partial homology with eukaryotic receptors such as Toll and G-protein dependent receptors. A leucine-rich repeat domain, an EF-hand domain, and two NEAT domains are identified in Shr. Shr directly binds heme-proteins such as hemoglobin, myoglobin, heme-BSA and the hemoglobin-heptoglobin complex (127,128). The presence of a nearly perfect EF-hand domain in Shr raises the possibility that Ca²⁺ may modulate its activity and represent a new type of Ca²⁺ regulated receptor involved in heme-protein binding and iron acquisition.

The single-handed EF-hand motifs were also observed in *Arabidopsis* (129). These observations raise the possibility that the ubiquitous EF-hand motif may function as an independent structural unit for Ca²⁺ binding. To date, the majority of known EF-hand motifs are coupled through the hydrophobic interaction of the flanking helices (21,22,130,131). Our previous work has shown that the isolated EF-loop III from CaM without the flanking helices in a host protein is able to bind Ca²⁺ and remains monomer in solution (70), whereas peptide fragments encompassing the helix-loop-helix EF-hand motifs were shown to dimerize in solution (132-135). We hope that our prediction could spur the exploration of the relationship between their function and Ca²⁺ binding capability.

Table 3.4. List of potential prokaryotic EF-hand-containing proteins.

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01799660	Bifunctional protein	165-193	1	<i>Acinetobacter</i> sp. ADP3
NF01800980	Hypothetical protein	134-162	1	<i>Acinetobacter</i> sp. ADP4
NF00596974	AGR_C_660p	166-194	1	<i>Agrobacterium tumefaciens</i>
NF00854152	Hypothetical protein Atu0377	147-175	1	<i>Agrobacterium tumefaciens</i>
NF00853975	Potassium-transporting ATPase B chain	555-583	1	<i>Agrobacterium tumefaciens</i>
NF00854754	Two component response regulator	273-301	1	<i>Agrobacterium tumefaciens</i>
NF00601861	AGR_C_1956p	288-316	1	<i>Agrobacterium tumefaciens</i>
NF02630316	Haloacid dehalogenase-like hydrolase:E1-E2 ATPase-associated region	556-584	1	<i>Anaeromyxobacter dehalogenans</i>
NF02630616	DNA gyrase, subunit A precursor	501-529	1	<i>Anaeromyxobacter dehalogenans</i>
NF02632976	Twin-arginine translocation pathway signal precursor	227-255	1	<i>Anaeromyxobacter dehalogenans</i>
NF02633157	TPR repeat precursor	343-371	1	<i>Anaeromyxobacter dehalogenans</i>
NF02633311	Aminotransferase, putative	153-181	1	<i>Anaeromyxobacter dehalogenans</i>
NF02145020	Prolyl-tRNA synthetase	220-248	1	<i>Anaplasma marginale</i>
NF00005240	Na ⁺ :solute symporter	266-294	1	<i>Aquifex aeolicus</i>
NF02418698	Initiation factor 2:Small GTP-binding protein domain	712-740	1	<i>Arthrobacter</i> sp. FB24
NF02419170	Regulatory protein, LysR:LysR, substrate-binding	188-216	1	<i>Arthrobacter</i> sp. FB24
NF02420554	Amino acid-binding ACT:Prephenate dehydrogenase	325-353	1	<i>Arthrobacter</i> sp. FB24
NF02183774	Hypothetical protein	129-157	1	<i>Azoarcus</i> sp. EbN2
NF02110613	AcrA/AcrE family putative membrane-fusion protein	57-85	1	<i>Azoarcus</i> sp. EbN3
NF02111329	Isoleucyl-tRNA synthetase	608-636	1	<i>Azoarcus</i> sp. EbN4
NF02599343	Hypothetical protein	333-361	1	<i>Azotobacter vinelandii</i>
NF02601050	Glycine cleavage system P-protein	678-706	1	<i>Azotobacter vinelandii</i>
NF01002531	Polysaccharide deacetylase-like protein	229-257	1	<i>Bacillus anthracis</i>
NF01003358	Multidrug resistance protein, putative	174-202	1	<i>Bacillus anthracis</i>
NF01293886	Acetyltransferase, GNAT family	61-89	1	<i>Bacillus anthracis</i>
NF00452066	hypothetical protein pxo2_15	123-151	1	<i>Bacillus anthracis</i>
NF01293170	Oxidoreductase, short-chain dehydrogenase/reductase family	195-223	1	<i>Bacillus anthracis</i>
NF01968462	Probable multidrug efflux transporter	174-202	1	<i>Bacillus cereus</i>
NF01692216	Isoleucyl-tRNA synthetase	583-611	1	<i>Bacillus cereus</i>
NF02763311	Polysaccharide deacetylase family protein, putative	229-257	1	<i>Bacillus cereus</i>
NF01693845	Acetyltransferase, GNAT family	61-89	1	<i>Bacillus cereus</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01269023	3-oxoacyl-[acyl-carrier protein] reductase	195-223	1	<i>Bacillus cereus</i>
NF02058262	3-oxoacyl-[acyl-carrier-protein] reductase	191-219	1	<i>Bacillus clausii</i>
NF00639248	Ribonuclease G (Cytosolic axial filament protein)	291-319	1	<i>Bacillus halodurans</i>
NF01942957	Hypothetical protein (D-isomer specific 2-hydroxyacid dehydrogenase, NAD binding domain)	227-255	1	<i>Bacillus licheniformis</i>
NF01943135	Putative L-2,4-diaminobutyrate decarboxylase	216-244	1	<i>Bacillus licheniformis</i>
NF00459033	Hypothetical oxidoreductase yoxD	191-219	1	<i>Bacillus subtilis</i>
NF01847653	Probable multidrug efflux transporter	174-202		<i>Bacillus thuringiensis</i>
NF01850556	Acetyltransferase, GNAT family	61-89	1	<i>Bacillus thuringiensis</i>
NF02232249	Hypothetical protein	740-768	1	<i>Bacillus thuringiensis</i>
NF01852124	Short chain dehydrogenase	195-223	1	<i>Bacillus thuringiensis</i>
NF01982311	Hypothetical protein	9-37	1	<i>Bacteroides fragilis</i>
NF02360737	Putative glycosyl hydrolase	1087-1115 1123-1151	2	<i>Bacteroides fragilis</i>
NF01983199	K ⁺ -transporting ATPase B chain	545-573	1	<i>Bacteroides fragilis</i>
NF01984037	Putative outer membrane protein	889-917	1	<i>Bacteroides fragilis</i>
NF02360881	Putative AraC-family transcriptional regulator	67-95	1	<i>Bacteroides fragilis</i>
NF01247854	Putative outer membrane protein	856-884	1	<i>Bacteroides thetaiotaomicron</i>
NF01244792	Alpha-xylosidase	1093-1121 1129-1157	2	<i>Bacteroides thetaiotaomicron</i>
NF01247957	Hypothetical protein	57-85	1	<i>Bacteroides thetaiotaomicron</i>
NF01243332	K ⁺ -transporting ATPase B chain	540-568	1	<i>Bacteroides thetaiotaomicron</i>
NF01632112	Ribonuclease G	309-337	1	<i>Bdellovibrio bacteriovorus</i>
NF01633060	Probable GTP-binding protein	120-148	1	<i>Bdellovibrio bacteriovorus</i>
NF01125245	BglX	745-773	1	<i>Bifidobacterium longum</i>
NF01125387	5'-nucleotidase family protein	412-440	1	<i>Bifidobacterium longum</i>
NF01125928	Solute binding protein of ABC transporter for branched-chain amino acids	360-388	1	<i>Bifidobacterium longum</i>
NF01356204	Putative outer membrane protein	133-161	1	<i>Bordetella bronchiseptica</i>
NF01357346	Putative exported protein	53-81	1	<i>Bordetella bronchiseptica</i>
NF01358260	Putative exported protein	193-221	1	<i>Bordetella bronchiseptica</i>
NF01360429	Potassium-transporting ATPase B chain	582-610	1	<i>Bordetella bronchiseptica</i>
NF01857138	Decorin binding protein B	94-122	1	<i>Borrelia garinii</i>
NF01347842	Potassium-transporting ATPase B chain	582-610	1	<i>Bordetella parapertussis</i>
NF01351614	Potassium-transporting ATPase B chain	582-610	1	<i>Bordetella pertussis</i>
NF01354467	Putative exported protein	198-226	1	<i>Bordetella pertussis</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01143994	Bll5001 protein	92-120	1	<i>Bradyrhizobium japonicum</i>
NF01144916	Bll3714 protein	4088-4116	1	<i>Bradyrhizobium japonicum</i>
NF01146006	Bll6206 protein	28-56	2	<i>Bradyrhizobium japonicum</i>
NF01145264	Bsr4957 protein	77-105	1	<i>Bradyrhizobium japonicum</i>
NF01138884	Potassium-transporting ATPase B chain	28-56	1	<i>Bradyrhizobium japonicum</i>
NF01145262	Bll5679 protein	567-595	1	<i>Bradyrhizobium japonicum</i>
NF02362000	Hypothetical protein	105-133	2	<i>Bradyrhizobium japonicum</i>
NF02362472	Chaperone protein DnaK	139-167	1	<i>Brucella melitensis</i>
NF00882038	Cobalt-zinc-cadmium resistance protein CZCD	187-215	1	<i>Brucella melitensis</i>
NF00882425	Transporter	173-201	1	<i>Brucella melitensis</i>
NF00729490	Acid-shock protein, putative	213-241	1	<i>Brucella melitensis</i>
NF01092664	Hypothetical protein	172-200	2	<i>Brucella melitensis</i>
NF01013862	Exodeoxyribonuclease V beta chain	63-91	1	<i>Brucella melitensis</i>
NF00001599	Cell division protein ftsY homolog	120-148	1	<i>Buchnera aphidicola</i>
NF02558178	Haloacid dehalogenase-like hydrolase:E1-E2 ATPase-associated region	191-219	1	<i>Buchnera aphidicola</i>
NF02562763	Pyridoxal-dependent decarboxylase	542-570	1	<i>Burkholderia cenocepacia</i>
NF01994075	Hypothetical protein	555-583	1	<i>Burkholderia cenocepacia</i>
NF01995267	AMP nucleosidase	150-178	1	<i>Burkholderia mallei</i>
NF01997459	K ⁺ -transporting ATPase, B subunit	100-128	1	<i>Burkholderia mallei</i>
NF01974439	Putative exported protein	136-164	1	<i>Burkholderia mallei</i>
NF01975408	Potassium-transporting ATPase b chain	556-584	1	<i>Burkholderia pseudomallei</i>
NF01976188	AMP nucleosidase	110-138	1	<i>Burkholderia pseudomallei</i>
NF02692726	Iron-sulfur cluster binding protein	547-575	1	<i>Burkholderia pseudomallei</i>
NF02628440	Cj81-040 (Fragment)	136-164	1	<i>Campylobacter jejuni</i>
NF02271643	Dihydroorotase, putative	285-313	1	<i>Campylobacter jejuni</i>
NF00555721	Potassium-transporting ATPase B chain	130-158	1	<i>Campylobacter jejuni</i>
NF00556942	Hemin ABC transporter	157-185	1	<i>Campylobacter jejuni</i>
NF02581023	Uncharacterized ACR, COG1427	539-567	1	<i>Campylobacter jejuni</i>
NF02581415	Iron-sulfur cluster binding protein	101-129	1	<i>Campylobacter lari</i>
NF02513884	Hypothetical protein	169-197	1	<i>Campylobacter lari</i>
NF02514058	Hypothetical protein	282-310	1	<i>Campylobacter upsaliensis</i>
NP_419996	hypothetical protein CC1180	155-183	1	<i>Campylobacter upsaliensis</i>
		48-76	4	<i>Caulobacter crescentus</i>
		96-124		
		120-148		
		197-225		
		221-249		

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NP_420996	hypothetical protein CC2193	36-64 77-105 191-219	3	<i>Caulobacter crescentus</i>
NP_421029	EF hand domain protein	33-61 59-87	2	<i>Caulobacter crescentus</i>
NP_421548	EF hand domain protein	294-322 319-337	2	<i>Caulobacter crescentus</i>
NF00922900	Potassium-transporting ATPase B chain	547-575	1	<i>Caulobacter vibrioides</i>
NF00923077	Hypothetical protein CC1180	96-124	1	<i>Caulobacter vibrioides</i>
NF00921487	Hypothetical protein CC2193	319-347	1	<i>Caulobacter vibrioides</i>
NF00921074	TonB-dependent receptor	282-310	1	<i>Caulobacter vibrioides</i>
NF00920950	EF hand domain protein	33-61	1	<i>Caulobacter vibrioides</i>
NF00919601	EF hand domain protein	36-64	1	<i>Caulobacter vibrioides</i>
NF00920207	dnaK-type molecular chaperone dnaK	173-201	1	<i>Caulobacter vibrioides</i>
NF00170276	Hypothetical protein CPn0045	460-488	1	<i>Chlamydophila pneumoniae</i>
NF01414521	Potassium-transporting ATPase B chain	566-594	1	<i>Chromobacterium violaceum</i>
NF01415832	Flagellar hook-associated protein	130-158	1	<i>Chromobacterium violaceum</i>
NF01417084	Acid phosphatase	38-66	1	<i>Chromobacterium violaceum</i>
NF00462703	DnaK protein	184-212	1	<i>Clostridium acetobutylicum</i>
NF00464732	General secretion pathway protein E, ATPase	4-32	1	<i>Clostridium acetobutylicum</i>
NF00463611	Membrane flavodoxin oxidoreductase	367-395	1	<i>Clostridium acetobutylicum</i>
NF00464531	dockerin domain	394-422	1	<i>Clostridium acetobutylicum</i>
NF00462598	Possible non-processive endoglucanase family 5, secreted; CelA homolog secreted; dockerin domain	418-446 450-478	2	<i>Clostridium acetobutylicum</i>
NF00464704	Probably secreted sialidase; several ASP-boxes and dockerin domain	768-796 800-828	2	<i>Clostridium acetobutylicum</i>
NF00461910	Cellulase CelE ortholog; dockerin domain	805-833 837-865	2	<i>Clostridium acetobutylicum</i>
NF00464704	dockerin and cellulose-binding domain	647-675 679-707	2	<i>Clostridium acetobutylicum</i>
NF00465242	dockerin and cellulose-binding domain	628-656 660-688	2	<i>Clostridium acetobutylicum</i>
NF00464378	dockerin domain	662-690 696-724	2	<i>Clostridium acetobutylicum</i>
NF00872981	Probable thioredoxin reductase	367-395	1	<i>Clostridium perfringens</i>
NF01191711	Encapsulation protein capA	25-53	1	<i>Clostridium tetani</i>
NF01192507	Putative cardiolipin synthetase 2	436-464	1	<i>Clostridium tetani</i>
NF01190826	RNA polymerase sigma factor for flagellar operon fliA	156-184	1	<i>Clostridium tetani</i>
NF01191088	Putative S-layer protein	438-466	1	<i>Clostridium tetani</i>
NF02636675	transcriptional regulator, LysR family	143-171	1	<i>Colwellia psychrerythraea</i>
NF02635517	hypothetical protein CPS_0266	258-286	1	<i>Colwellia psychrerythraea</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF02635672	DNA polymerase III, gamma/tau subunit	72-100	1	<i>Colwellia psychrerythraea</i>
NF02637464	EF hand domain protein	22-50	1	<i>Colwellia psychrerythraea</i>
NF02637914	EF hand domain protein	6-34 42-70	2	<i>Colwellia psychrerythraea</i>
NF02638565	hypothetical protein CPS_4268	89-117	1	<i>Colwellia psychrerythraea</i>
NF02638842	patatin-like phospholipase family protein	435-463	1	<i>Colwellia psychrerythraea</i>
NF01120262	TnpC protein	21-49	1	<i>Corynebacterium efficiens</i>
NF00925376	Conserved hypothetical secreted protein	163-191	1	<i>Corynebacterium glutamicum</i>
NF02423383	Putative surface-anchored protein	537-565	1	<i>Corynebacterium jeikeium</i>
NF02421939	Hypothetical protein	136-164 304-332	2	<i>Corynebacterium jeikeium</i>
NF02421622	ATP-dependent DNA helicase	287-315	1	<i>Corynebacterium jeikeium</i>
NF02421112	Thiol-disulfide isomerase/thioredoxin precursor	59-87	1	<i>Corynebacterium jeikeium</i>
NF02492865	Calcium-binding EF-hand	70-98 136-164	2	<i>Cupriavidus necator</i>
NF02493293	ATP-binding region, ATPase-like:Histidine kinase A, N-terminal	426-454	1	<i>Cupriavidus necator</i>
NF02493847	Fungal/archaeal/bacterial haem catalase/oxidase	518-546	1	<i>Cupriavidus necator</i>
NF02495396	Outer membrane autotransporter barrel	496-524	1	<i>Cupriavidus necator</i>
NF02497227	Histidine kinase, HAMP region:Bacterial chemotaxis sensory transducer	349-377	1	<i>Cupriavidus necator</i>
NF02497778	ATPase, E1-E2 type:Potassium-translocating P-type ATPase, B subunit	600-628	1	<i>Cupriavidus necator</i>
NF02274727	hypothetical protein DET1403	195-223	1	<i>Dehalococcoides ethenogenes</i>
NF02696634	Calcium-binding EF-hand	46-74	1	<i>Dechloromonas aromatica</i>
NF02697139	Glycine cleavage system P-protein	613-641	1	<i>Dechloromonas aromatica</i>
NF02274727	hypothetical protein DET1403	195-223	1	<i>Dehalococcoides ethenogenes</i>
NF02650500	Similar to High-affinity K ⁺ transport system ATPase chain B precursor	26-54	1	<i>Deinococcus geothermalis</i>
NF02648532	Hypothetical protein precursor	120-148	1	<i>Deinococcus geothermalis</i>
NF00437343	Ribonuclease II family protein	88-116	1	<i>Deinococcus radiodurans</i>
NF00435709	Hypothetical protein DR0685	179-207	1	<i>Deinococcus radiodurans</i>
NF00438084	Hypothetical protein DR1404	14-42	1	<i>Deinococcus radiodurans</i>
NF00435967	Hypothetical protein DR2551	191-219	1	<i>Deinococcus radiodurans</i>
NF01898379	Hypothetical protein	161-189	1	<i>Desulfotalea psychrophila</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01900066	Hypothetical protein	277-305 384-412 491-519	3	<i>Desulfotalea psychrophila</i>
NF01899559	Probable high affinity sulfate transporter (SulP)	560-588	1	<i>Desulfotalea psychrophila</i>
NF01722462	K ⁺ -transporting ATPase, B subunit	539-567	1	<i>Desulfovibrio vulgaris</i>
NF01723330	EF hand domain protein	19-47	1	<i>Desulfovibrio vulgaris</i>
NF01724660	EF hand domain protein	51-79 152-180	2	<i>Desulfovibrio vulgaris</i>
NF01239980	Glycosyl hydrolase, family 31/fibronectin type III domain protein	1162-1190	1	<i>Enterococcus faecalis</i>
NF02746267	DnaK molecular chaperone	173-201	1	<i>Erythrobacter litoralis</i>
NF02746140	Hypothetical protein	97-125	1	<i>Erythrobacter litoralis</i>
NF02746586	Hypothetical protein	93-121	1	<i>Erythrobacter litoralis</i>
NF00704786	Hypothetical protein ECs5257	4-32	1	<i>Escherichia coli</i>
NF00695695	N-acetylmuramoyl-L-alanine amidase amiB precursor	379-407	1	<i>Escherichia coli</i>
NF01134307	Rhamnulokinase	324-352	1	<i>Escherichia coli</i>
NF01744738	Putative glycosyltransferase	174-202	1	<i>Escherichia coli</i>
NF00692267	Resolvase family recombinase	6-34	1	<i>Escherichia coli</i>
NF01491630	Zn-dependent hydrolase	462-490	1	<i>Fusobacterium nucleatum</i>
NF01492715	Protease	388-416	1	<i>Fusobacterium nucleatum</i>
NF02148420	Hypothetical conserved protein	73-101	1	<i>Geobacillus kaustophilus</i>
NF01553463	EF hand domain protein	22-50 59-87	2	<i>Geobacter sulfurreducens</i>
NF01551775	EF hand domain/PKD domain protein	1723-1751	1	<i>Geobacter sulfurreducens</i>
NF01553119	Salmonella virulence plasmid 65kDa B protein	1363-1391	1	<i>Geobacter sulfurreducens</i>
NF01553013	Potassium-transporting ATPase	552-580	1	<i>Geobacter sulfurreducens</i>
NF01419166	Gli3319 protein	106-134	1	<i>Gloeobacter violaceus</i>
NF01422150	Glr3888 protein	757-785	1	<i>Gloeobacter violaceus</i>
NF01420350	Transaldolase	357-385	1	<i>Gloeobacter violaceus</i>
NF02258009	Cation efflux system protein	211-239	1	<i>Gluconobacter oxydans</i>
NF01345996	Cell division protein FtsZ	259-287	1	<i>Haemophilus ducreyi</i>
NF01345965	Hypothetical protein	60-88	1	<i>Haemophilus ducreyi</i>
NF01744248	Adhesin	286-314	1	<i>Haemophilus influenzae</i>
NF00736924	Hypothetical protein HI1594	179-207	1	<i>Haemophilus influenzae</i>
NF02647744	Exodeoxyribonuclease V gamma chain	469-497	1	<i>Haemophilus influenzae</i>
NF02647997	DNA translocase FtsK	290-318	1	<i>Haemophilus influenzae</i>
NF01331967	Bacterial cell division topological specificity factor MinE	51-79	1	<i>Helicobacter hepaticus</i>
NF01331990	Aconitate hydratase	404-432	1	<i>Helicobacter hepaticus</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF00574925	Ulcer associated adenine specific DNA methyltransferase	290-318	1	<i>Helicobacter pylori</i>
NF00572083	Hypothetical protein alpB	248-276	1	<i>Helicobacter pylori</i>
NF02103893	DNA-directed RNA polymerase beta subunit	405-433	1	<i>Idiomarina loihiensis</i>
NF02104158	RND family efflux system membrane fusion protein	62-90	1	<i>Idiomarina loihiensis</i>
NF02105366	GntP family permease	335-363	1	<i>Idiomarina loihiensis</i>
NF02106157	Outer membrane protein	535-563	1	<i>Idiomarina loihiensis</i>
NF01526365	TnpR recombinase	29-57	1	<i>Klebsiella pneumoniae</i>
NF02267084	Iron-sulfur cofactor synthesis protein	113-141	1	<i>Lactobacillus acidophilus</i>
NF02266040	Alpha-glucosidase	953-981	1	<i>Lactobacillus acidophilus</i>
NF01324779	Cysteine sulfinatase desulfurase/cysteine desulfurase	114-142	1	<i>Lactobacillus johnsonii</i>
NF01585065	Hypothetical protein	330-358	1	<i>Lactobacillus johnsonii</i>
NF01210549	Cation transporting P-type ATPase	9339-367	1	<i>Lactobacillus plantarum</i>
NF01211534	Phosphoglucosamine mutase	234-262	1	<i>Lactobacillus plantarum</i>
NF00449250	Maltose ABC transporter substrate binding protein	111-139	1	<i>Lactococcus lactis</i>
NF02043787	RNA polymerase beta' subunit	1365-1393	1	<i>Legionella pneumophila</i>
NF00628534	O-methyltransferase, SAM-dependent	132-160	1	<i>Legionella pneumophila</i>
NF02048951	DNA-directed RNA polymerase beta' subunit	1379-1407	1	<i>Legionella pneumophila</i>
NF01698417	Hypothetical protein	12-40	1	<i>Leptospira interrogans</i>
NF01085961	Potassium-transporting ATPase B chain	548-576	1	<i>Leptospira interrogans</i>
NF00554020	Preprotein translocase SecY	41-69	1	<i>Leptospira interrogans</i>
NF00844705	Potassium-transporting ATPase B chain 2	543-571	1	<i>Listeria innocua</i>
NF00813533	Lin0875 protein	48-76	1	<i>Listeria innocua</i>
NF00813673	Lin2223 protein	232-260	1	<i>Listeria innocua</i>
NF01748822	GGDEF domain protein	152-180	1	<i>Listeria monocytogenes</i>
NF01746930	Transketolase	338-366	1	<i>Listeria monocytogenes</i>
NF00812586	Lmo1869 protein	228-256	1	<i>Listeria monocytogenes</i>
NF00604651	Mlr9645 protein	129-157 152-180	2	<i>Mesorhizobium loti</i>
NF00606972	Mll5457 protein	69-97 102-130	2	<i>Mesorhizobium loti</i>
NF00610956	Potassium-transporting ATPase B chain	551-579	1	<i>Mesorhizobium loti</i>
NF02036858	DNA gyrase, A subunit	518-546	1	<i>Methylococcus capsulatus</i>
NF02038758	Hypothetical protein	209-237	1	<i>Methylococcus capsulatus</i>
NF01628109	GyrA	496-524	1	<i>Mycobacterium avium</i>
NF01628325	Hypothetical protein	76-104	1	<i>Mycobacterium avium</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01324088	Potassium-transporting ATPase B chain	570-598	1	<i>Mycobacterium bovis</i>
NF01321371	DNA gyrase subunit A	495-523	1	<i>Mycobacterium bovis</i>
NF00480640	probable DNA gyrase subunit A	916-944	1	<i>Mycobacterium leprae</i>
NF00484231	P450 heme-thiolate protein	335-363	1	<i>Mycobacterium tuberculosis</i>
NF00485681	Potassium-transporting ATPase B chain	570-598	1	<i>Mycobacterium tuberculosis</i>
NF01586715	Mannitol-permease IIBC component	43-71	1	<i>Mycoplasma mycoides</i>
NF01587410	Hypothetical protein	9-37	1	<i>Mycoplasma mycoides</i>
NF00630859	Putative type II restriction endonuclease NlaIV	119-147	1	<i>Neisseria gonorrhoeae</i>
NF01296409	DEAD/DEAH box helicase:HD domain	624-652	1	<i>Nitrosomonas europaea</i>
NF01294890	Hypothetical protein	6-34	1	<i>Nitrosomonas europaea</i>
NF01296468	Site-specific recombinase	215-243	1	<i>Nitrosomonas europaea</i>
NF02020754	Putative potassium transporter B subunit	544-572	1	<i>Nocardia farcinica</i>
NF02023298	Hypothetical protein	23-51	1	<i>Nocardia farcinica</i>
NF02024201	Putative resolvase	3-31	1	<i>Nocardia farcinica</i>
NF00818483	All5017 protein	190-218	1	<i>Nostoc sp. PCC 7120</i>
NF00819095	Alr7304 protein	3336-3364	1	<i>Nostoc sp. PCC 7120</i>
NF00817274	RND multidrug efflux transporter	87-115	1	<i>Nostoc sp. PCC 7120</i>
NF00817119	Asr1131 protein	45-73	1	<i>Nostoc sp. PCC 7120</i>
NF01061743	NADH-dependent flavin oxidoreductase	52-80	1	<i>Oceanobacillus iheyensis</i>
NF01776242	Putative pyridoxal-dependent decarboxylase	150-178	1	<i>Pectobacterium atrosepticum</i>
NF01829006	Hypothetical methylamine utilization protein	5-33	1	<i>Photobacterium profundum</i>
NF01830545	Phosphoglycerate transport system transcription regulator	250-278	1	<i>Photobacterium profundum</i>
NF01831581	Hypothetical protein	500-528	1	<i>Photobacterium profundum</i>
NF01833037	Hypothetical small protein A	21-49	1	<i>Photobacterium profundum</i>
NF01831267	Hypothetical protein	278-306	1	<i>Photobacterium profundum</i>
NF01423716	Similar to proteins involved in antibiotics biosynthesis	1476-1504	1	<i>Photorhabdus luminescens</i>
NF01426461	Hypothetical protein	106-134	1	<i>Photorhabdus luminescens</i>
NF01427112	LsrR DNA-binding protein	13-41	1	<i>Photorhabdus luminescens</i>
NF01410240	Secretion activator protein, putative	35-63	1	<i>Porphyromonas gingivalis</i>
NF02499786	carboxysome shell protein CsoS2	102-130	1	<i>Prochlorococcus marinus</i>
NF02499308	ornithine carbamoyltransferase	16-44	1	<i>Prochlorococcus marinus</i>
NF01372938	Hypothetical protein precursor	38-66	1	<i>Prochlorococcus marinus</i>
NF02498817	hypothetical protein PMN2A_0622	557-585	1	<i>Prochlorococcus marinus</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01854896	Cobalamin biosynthesis protein CobN	130-158	1	<i>Propionibacterium acnes</i>
NF01010095	Similar to Histidine biosynthesis protein	165-193	1	<i>Pseudomonas aeruginosa</i>
NF00582290	Chaperone protein htpG	404-432	1	<i>Pseudomonas aeruginosa</i>
NF00579604	TnpR protein (Fragment)	6-34	1	<i>Pseudomonas aeruginosa</i>
NF00585284	Hypothetical protein	79-107	1	<i>Pseudomonas aeruginosa</i>
NF00580099	Heat shock protein 66-kDa	187-215	1	<i>Pseudomonas aeruginosa</i>
NF00584502	Hypothetical protein	203-231	1	<i>Pseudomonas aeruginosa</i>
NF02412314	Calcium-binding protein	24-52 60-88 98-126	3	<i>Pseudomonas fluorescens</i>
NF02409575	K ⁺ -transporting ATPase, B subunit	551-579	1	<i>Pseudomonas fluorescens</i>
NF01130810	Chaperone protein hscA homolog	187-215	1	<i>Pseudomonas putida</i>
NF00593298	Resolvase	6-34	1	<i>Pseudomonas putida</i>
NF01131075	Hypothetical protein	39-67	1	<i>Pseudomonas putida</i>
NF01128446	Glycine dehydrogenase	613-641	1	<i>Pseudomonas putida</i>
NF01131319	EF hand domain protein	143-171	1	<i>Pseudomonas putida</i>
NF02715683	calcium-binding protein	101-159 134-162 180-208	3	<i>Pseudomonas savastanoi</i>
NF02713698	heat shock protein HtpG	405-433	1	<i>Pseudomonas savastanoi</i>
NF02713796	hypothetical protein PSPPH_1452	119-147	1	<i>Pseudomonas savastanoi</i>
NF02715540	type IV pilus-associated protein, putative	990-1018	1	<i>Pseudomonas savastanoi</i>
NF02713796	hypothetical protein PSPPH_1452	119-147	1	<i>Pseudomonas savastanoi</i>
NF02518859	Calcium-binding EF-hand precursor	63-91 100-128 133-161 179-207	4	<i>Pseudomonas syringae</i>
NF02518864	Heat shock protein Hsp90:ATP-binding region, ATPase-like	407-435	1	<i>Pseudomonas syringae</i>
NF02517892	Type IV pilus-associated protein, putative precursor	989-1017	1	<i>Pseudomonas syringae</i>
NF01447618	Chaperone protein htpG	405-433	1	<i>Pseudomonas syringae</i> group genomosp. 3
NF01447738	Type IV pilus-associated protein, putative	989-1017	1	<i>Pseudomonas syringae</i> group genomosp. 3
NF01447731	Hypothetical protein	143-171	1	<i>Pseudomonas syringae</i> group genomosp. 3
NF01444486	EF hand domain protein	4-32 316-344 391-419	3	<i>Pseudomonas syringae</i> group genomosp. 3

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF02500531	Calcium-binding EF-hand precursor	63-91 100-128 135-161 179-207	4	<i>Pseudomonas syringae</i>
NF01410240	Secretion activator protein, putative	35-63	1	<i>Porphyromonas gingivalis</i>
NF00853288	Potassium-transporting ATPase B chain	605-633	1	<i>Ralstonia solanacearum</i>
NF00849451	Hypothetical protein RSc1742	295-323	1	<i>Ralstonia solanacearum</i>
NF00850536	Putative calcium binding signal peptide protein	74-102, 113-141	2	<i>Ralstonia solanacearum</i>
AAG21376	Calsymin	79-107 122-150 157-185 189-217 236-264 269-297	6	<i>Rhizobium etli</i>
NF01530872	Calcium-binding EF-hand precursor	94-122 128-156	2	<i>Rhodopseudomonas palustris</i>
NF01643065	Similar to CaM-like protein-putative secreted, membrane associated or paryphoplasmic Ca(2+)-binding protein	50-78	1	<i>Rhodopirellula baltica</i>
NF01643299	Hypothetical protein	257-285	1	<i>Rhodopirellula baltica</i>
NF01649307	Hypothetical protein	103-131	1	<i>Rhodopirellula baltica</i>
NF01643065	putative secreted, membrane associated or paryphoplasmic Ca ²⁺ -binding protein	50-78	1	<i>Rhodopirellula baltica</i>
NF01644848	Hypothetical protein	214-242 340-368 417-445	3	<i>Rhodopirellula baltica</i>
NF01647263	Matrix metalloproteinase	516-544	1	<i>Rhodopirellula baltica</i>
NF01647900	Probable lipase/esterase	76-104	1	<i>Rhodopirellula baltica</i>
NF01643609	Alkaline phosphatase	1576-1604	1	<i>Rhodopirellula baltica</i>
NF01649041	Putative ABC transporter integral membrane protein	42-70	1	<i>Rhodopirellula baltica</i>
NF01644544	Probable calmodulin	347-375	1	<i>Rhodopirellula baltica</i>
NF01644969	Hypothetical protein	71-99	1	<i>Rhodopirellula baltica</i>
NF01528433	Hypothetical protein precursor	64-92	1	<i>Rhodopseudomonas palustris</i>
:NF01529160	Potassium-transporting atpase b chain, KdpB	570-598	1	<i>Rhodopseudomonas palustris</i>
NF01530032	Hypothetical protein precursor	26-54	1	<i>Rhodopseudomonas palustris</i>
NF01530456	Hypothetical protein	245-273	1	<i>Rhodopseudomonas palustris</i>
NF01530872	Calcium-binding EF-hand precursor	94-122 128-156	2	<i>Rhodopseudomonas palustris</i>
NF01531759	Glycosyl hydrolase	64-92	1	<i>Rhodopseudomonas palustris</i>
NF01532507	DUF218	94-122	1	<i>Rhodopseudomonas palustris</i>
NF02577387	DnaK protein	172-200	1	<i>Rickettsia felis</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
P06495	Calerythrin	9-37 60-88 104-132 138-164	4	<i>Saccharopolyspora erythraea</i>
NF00806470	probable glycogen debranching protein	35-63	1	<i>Salmonella enterica</i>
NF02097096	Putative glycogen debranching protein homolog	35-63	1	<i>Salmonella paratyphi</i>
NF01911344	Putative glycogen debranching protein	35-63	1	<i>Salmonella typhi</i>
NF00861463	Putative glycosyl hydrolase	35-63	1	<i>Salmonella typhimurium</i>
NF01074234	Pyruvate formate-lyase 1 activating enzyme	35-63	1	<i>Shewanella oneidensis</i>
NF01074109	RTX toxin, putative	461-489	1	<i>Shewanella oneidensis</i>
NF02205755	EF hand domain protein	30-58	1	<i>Silicibacter pomeroyi</i>
NF02206123	Type I secretion target repeat protein	7295-7323	1	<i>Silicibacter pomeroyi</i>
NF00616675	Potassium-transporting ATPase B chain	541-569	1	<i>Sinorhizobium meliloti</i>
NF00612473	Hypothetical protein SMc01708	784-812	1	<i>Sinorhizobium meliloti</i>
NF00617441	Hypothetical protein SMB21413	137-165	1	<i>Sinorhizobium meliloti</i>
NF00431107	Transposase B	519-547	1	<i>Staphylococcus aureus</i>
NF02262335	LysM domain protein	235-263	1	<i>Staphylococcus aureus</i>
NF01789691	Putative dihydrolipoamide dehydrogenase	153-181	1	<i>Staphylococcus aureus</i>
NF00433953	Similar to Zn-dependent hydrolase	465-493	1	<i>Staphylococcus aureus</i>
NF01166407	Secretory antigen SsaA	108-136	1	<i>Staphylococcus epidermidis</i>
NF00434602	HMG-CoA reductase	297-325	1	<i>Staphylococcus epidermidis</i>
NF02269390	Tn554, transposase B	519-547	1	<i>Staphylococcus epidermidis</i>
NF01165622	Hydroxymethylglutaryl-CoA reductase	297-325	1	<i>Staphylococcus epidermidis</i>
NF01167423	Hypothetical protein SE0148	85-113	1	<i>Staphylococcus epidermidis</i>
NF01165840	Hypothetical protein SE0787	465-493	1	<i>Staphylococcus epidermidis</i>
NF02607639	Capsular polysaccharide synthesis enzyme CapE	299-327	1	<i>Staphylococcus haemolyticus</i>
NF02608469	Homoserine dehydrogenase	393-421	1	<i>Staphylococcus haemolyticus</i>
NF02663555	Similar to unknown protein	465-493	1	<i>Staphylococcus haemolyticus</i>
NF01048375	Proteinase, putative	626-654	1	<i>Streptococcus agalactiae</i>
NF01049107	Hypothetical protein gbs1919	194-222	1	<i>Streptococcus agalactiae</i>
NF01049640	Hypothetical protein gbs1489	371-399	1	<i>Streptococcus agalactiae</i>
NF01113363	Hypothetical protein gbs1332	210-238	1	<i>Streptococcus agalactiae</i>
NF00439678	Putative type II restriction endonuclease	263-291	1	<i>Streptococcus thermophilus</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01280369	Putative calcium binding protein	9-37 88-137 137-165	3	<i>Streptomyces avermitilis</i>
NF01277143	Hypothetical protein	24-52	1	<i>Streptomyces avermitilis</i>
NF01281064	Putative chitinase C	413-441	1	<i>Streptomyces avermitilis</i>
NF01278333	Putative high-affinity potassium transport system	555-583	1	<i>Streptomyces avermitilis</i>
NF01277145	Putative calcium binding protein	6-34	1	<i>Streptomyces avermitilis</i>
NF00543258	Putative calcium binding protein	10-38 60-88	2	<i>Streptomyces coelicolor</i>
NF00549617	Putative calcium-binding protein	9-37 60-88 98-126 132-160	4	<i>Streptomyces coelicolor</i>
NF00550219	Putative calcium binding protein	9-37 60-88 103-131 137-165	4	<i>Streptomyces coelicolor</i>
NF00543851	ChiC (Chitinase C) (Putative secreted protein) (Fragment)	415-443	1	<i>Streptomyces coelicolor</i>
NF00549644	Potassium-transporting ATPase B chain	567-595	1	<i>Streptomyces coelicolor</i>
NF00548721	SCO5464 protein	6-34	1	<i>Streptomyces coelicolor</i>
NF00543475	Hypothetical protein SCO4859	126-154	1	<i>Streptomyces coelicolor</i>
NF00928474	Hypothetical protein spyM18_1868	523-551	1	<i>Streptococcus pyogenes</i>
NF01861114	Putative Fe ³⁺ -siderophore transport protein	525-553	1	<i>Streptococcus pyogenes</i>
NF01223601	Putative PBP 5 synthesis repressor	63-91	1	<i>Streptococcus pyogenes</i>
NF00442096	Penicillin-binding protein 3	167-195	1	<i>Streptococcus pneumoniae</i>
NF01965281	Hypothetical protein	196-224	1	<i>Symbiobacterium thermophilum</i>
NF02172090	Transaldolase	358-386	1	<i>Synechococcus elongatus</i>
NF02171539	Carboxyl-terminal protease	152-180	1	<i>Synechococcus elongatus</i>
P72797	Transaldolase	333-361 356-384	2	<i>Synechocystis sp.</i>
NF00423899	Slr0366 protein	171-199 277-305 383-411 489-517	4	<i>Synechocystis sp. PCC 6803</i>
NF00424738	Cation or drug efflux system protein	87-115	1	<i>Synechocystis sp. PCC 6803</i>
NF00425666	Sulfate adenylyltransferase	99-127	1	<i>Synechocystis sp. PCC 6803</i>
NF00968263	Septum formation inhibitor-activating ATPase	28-56	1	<i>Thermoanaerobacter tengcongensis</i>
NF02549883	calcium-binding EF-hand	61-89 97-125	2	<i>Thermobifida fusca</i>
NF01023077	transaldolase	356-384	1	<i>Thermosynechococcus elongatus</i>
NF00578072	50S Ribosomal Protein L4	1-28	1	<i>Thermus thermophilus</i>

Accession ID	Protein Name	Sequence Range	EF-hand number	Organism
NF01703216	Hypothetical conserved protein	240-268	1	<i>Thermus thermophilus</i>
NF00723911	Hemolysin-related protein	54-82	1	<i>Vibrio cholerae</i>
NF00722357	Hypothetical protein VCA0849	3164-3192	1	<i>Vibrio cholerae</i>
NF01185821	FlgA	125-153	1	<i>Vibrio fischeri</i>
NF02302098	Iron-regulated protein FrpC	344-372	1	<i>Vibrio fischeri</i>
NF01230959	Chemotaxis protein CheV	133-161	1	<i>Vibrio parahaemolyticus</i>
NF01234459	Putative outer membrane protein	42-70	1	<i>Vibrio parahaemolyticus</i>
NF01147763	Autotransporter adhesin	1196-1224 2841-2869	2	<i>Vibrio vulnificus</i>
NF01149533	Multidrug resistance efflux pump	128-156	1	<i>Vibrio vulnificus</i>
NF01147763	Autotransporter adhesin	887-915 990-1018 1093-1121 1196-1224 2841-2869	5	<i>Vibrio vulnificus</i>
NF01147228	FOG: Ankyrin repeat	310-338	1	<i>Vibrio vulnificus</i>
NF01150679	ATPase component of various ABC-type transport system	353-381	1	<i>Vibrio vulnificus</i>
NF01149533	Multidrug resistance efflux pump	128-156	1	<i>Vibrio vulnificus</i>
NF01411038	Hypothetical protein	214-242		<i>Wolinella succinogenes</i>
NF01840495	exopolysaccharide synthesis protein ExoD-related protein	16-44	1	<i>Wolbachia sp. wMel</i>
NF01840821	Exopolysaccharide synthesis protein ExoD-related protein	16-44	1	<i>Wolbachia sp. wMel</i>
NF01189636	Metallopeptidase	56-84	1	<i>Xylella fastidiosa</i>
NF00981304	Calcium-binding protein	38-66 74-102 112-140 143-171	4	<i>Xanthomonas axonopodis</i>
NF01010603	NodB-like protein	662-690	1	<i>Xanthomonas campestris</i>
NF00974448	Hypothetical protein XCC0163	30-58	1	<i>Xanthomonas campestris</i>
NF00977416	Polysaccharide deacetylase	714-742	1	<i>Xanthomonas campestris</i>
NF00974241	Hypothetical protein XCC1206	172-200	1	<i>Xanthomonas campestris</i>
NF02292841	EF hand domain protein	617-645	1	<i>Xanthomonas oryzae</i>
NF02292473	Phage-related protein	342-370	1	<i>Xanthomonas oryzae</i>
NF02292388	Polysaccharide deacetylase	662-690	1	<i>Xanthomonas oryzae</i>
NF02294838	Hypothetical protein	88-116	1	<i>Xanthomonas oryzae</i>
NF02293741	Hypothetical protein	505-533	1	<i>Xanthomonas oryzae</i>
NF01189016	Hypothetical protein	36-64	1	<i>Xylella fastidiosa</i>
NF02140272	DnaK molecular chaperone	173-201	1	<i>Zymomonas mobilis</i>
NF02139509	Hypothetical protein	51-79	1	<i>Zymomonas mobilis</i>
NF02139489	Hypothetical protein	58-86	1	<i>Zymomonas mobilis</i>
NF02139845	Putative RTX family exoprotein	2479-2507	1	<i>Zymomonas mobilis</i>

3.7. Evolutionary perspectives on EF-hand proteins

The classification and evolution of EF-hand proteins was first analyzed by Kretsinger *et al* (11,136-138). A dendrogram of the EF-hand proteins has been published previously. Since then, more EF-hand subfamilies, especially pseudo-EF-hand proteins, have been revealed. To analyze the potential evolutionary scenario of the EF-hands, particularly pseudo EF-hands, phylogenetic analysis was carried out with updated pseudo EF-hand members and the canonical EF-hand proteins in this study on the basis of sequence alignments (Fig. 3.5).

The pseudo EF-hand N-J tree revealed three major groups assuming that they are evolved from a common ancestral protein. The largest group consists of two closely-related subgroups, one with S100A2, S100A3, S100A4, S100A5, and S100A6 and the other with S100A1, S100P, S100B, S100Z, and S100A10. It is interesting to note that S100A10, separating early from other members in its subgroup, loses the capacity to chelate Ca^{2+} ion with mutations and deletions at the Ca^{2+} liganding positions in both canonical and pseudo-EF-hand motifs. The small phylogenetic distance between S100A2, S100A3, S100A4, S100A5, and S100A6 is consistent with the clustered organization of these genes (139). Additionally, S100A2, S100A3, S100A5, and S100A6 have been proposed to coordinate Zn^{2+} with varying affinity (140-143). The second major group is comprised of S100A8, S100A9, S100A12, trichohyalin, and MRP-126 from chicken. All of these proteins (except for trichohyalin) are excreted to the extracellular space, where Ca^{2+} concentration is at the millimolar level (144). Their common targets are the cytoskeletal or cell membrane proteins. In addition,

the proteins in this group are associated with pro-inflammatory functions by inducing chemotaxis or secretion of pro-inflammatory mediators. Interestingly, members in this group possess the Zn^{2+} -chelating motif His-x-x-x-His at the C termini, with possible involvement of an upstream glutamate (145). The third major group consists of S100A7, S100A11, S100A11P, S100A15, S100H, and repetin. Repetin contains an N-terminal S100-like domain and central tandem repeats of glutamine-rich sequence (94). It is involved in epidermal differentiation. Repetin is separated early from other members in the group. The other members (S100A13 and S100A14, S100A16 and S100A17, and S100G and hornerin) form three minor groups. They are rather diverse and no valuable evolutionary clues can be inferred at present. Repetin, trichohyalin, and hornerin belong to the “fused gene” family. A proposed evolutionary pathway for hornerin involves the fusion of an S100-like Ca^{2+} -binding protein with an ancestral epidermal structural gene containing tandem repeats that reside in the same chromosomal locus 1q21 (94,146).

The canonical EF-hands are ubiquitously distributed across eukaryote, bacteria, and archaea. The gene replication could also be tentatively used to explain the appearance of penta- (calpain subfamily) and hexa-EF-hands (calretinin and calbindin D28k), both of which have distinct phylogenetic pathways (Fig. 3.5) (147,148). The abundant single-handed EF-hand-like motifs found in the genomes of bacteria could provide clues for the origin of the prototypical EF-hand (4,25,149,150). The evolutionary mobile single Ca^{2+} binding loop first present in the ancestral protein could be “transplanted” to several

locations of the host protein or to several host proteins (5,151).

In contrast to the canonical EF-hands, our study shows that the pseudo EF-hands are exclusively found in vertebrates with tissue- and cell-specific expression profiles and are not predicted in the bacteria genomes. The lowest organism containing pseudo EF-hand reported thus far is the spiny dogfish (*Squalus acanthias*) with a pseudo EF-hand protein that is closely related to S100A1 (13). Thus, it is reasonable to postulate that pseudo EF-hands are phylogenetically younger and have a shorter history than canonical EF-hands. Although more evidence is required to confirm this postulation, the current observation in natural proteins that pseudo EF-hands always pair with canonical EF-hands but canonical EF-hands do not necessarily pair with pseudo EF-hands also indicate that pseudo EF-hands appear later than the canonical EF-hands. Genomic study on human and rat S100 proteins has also indicated the recent origin of the S100 subfamily (13,139). It has been hypothesized that the evolution of pseudo EF-hands might be achieved by domain swapping through gene duplication or exon recombination from a CaM-type protein with subsequent loss of two of the four EF-hands (13). Then, evolutionary divergence of EF-hands follows, thereby creating the sequence diversity of pseudo EF-hands. During this process, pseudo EF-hands become distant relatives of canonical EF-hand and a number of pseudo EF-hands (S100B, S100A2, S100A3, S100A5, S100A6, S100A7, S100A12) acquire the ability to bind other metal ions such as Zn^{2+} or Cu^{2+} to further adapt to the tissue-specific temporal-spatial requirements(100). They evolved largely varied Ca^{2+} affinity from nM to mM to meet the versatile

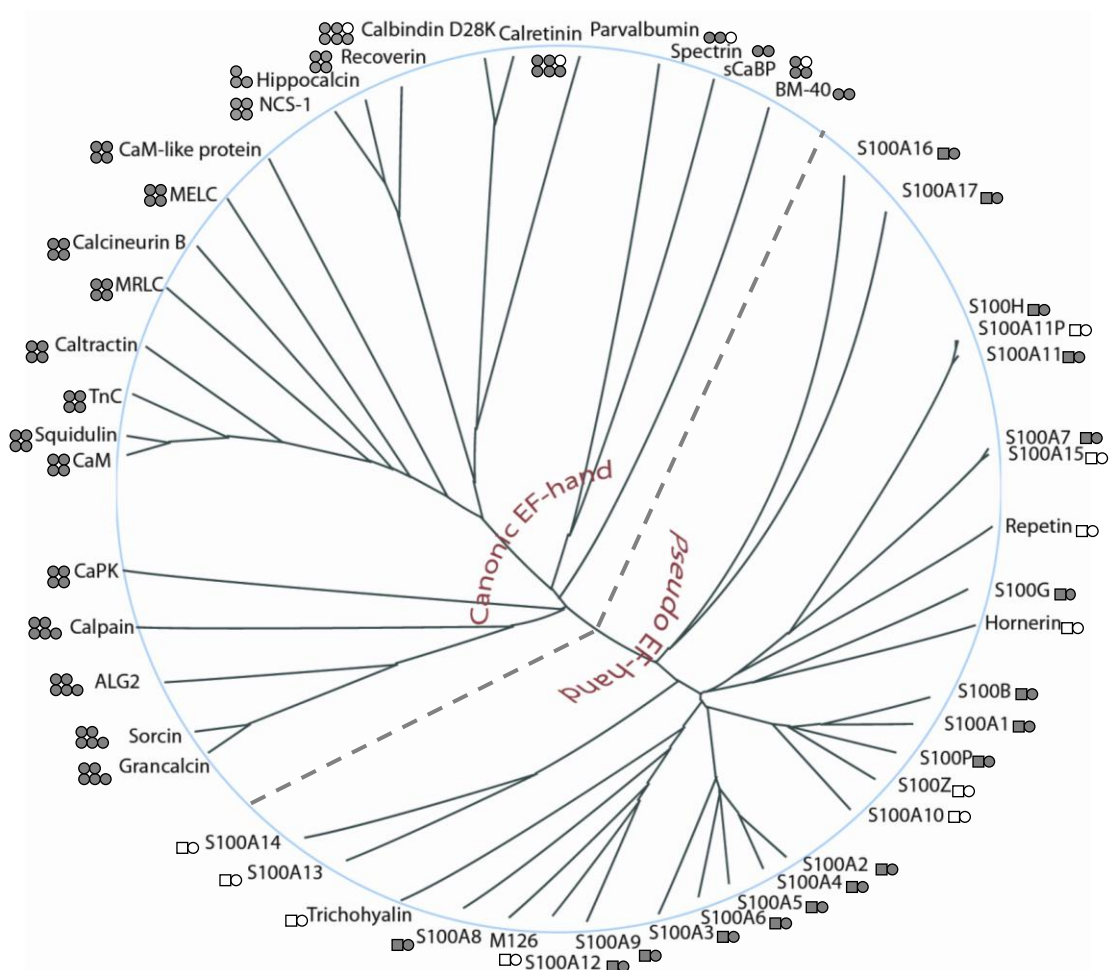


Figure 3.5. Phylogenetic analysis of the EF-hand protein family. The unrooted N-J tree was generated on the basis of multiple sequence alignments of 27 typical proteins containing pseudo EF-hand motifs and 22 proteins with canonical EF hand motifs. (Circle: canonical EF-hand; Square: pseudo EF-hand; Solid: bind Ca²⁺; Open: do not bind Ca²⁺ or Ca²⁺ binding capability is unknown).

requirements in various cellular compartments (140,141,143,152,153).

3.8. Summary

With more and more EF-hand Ca^{2+} binding proteins being discovered and characterized in bacteria, archaea, and eukaryotes, structural and functional knowledge of the EF-hand proteins has expanded steadily in recent years. The EF-hand-like proteins contain Ca^{2+} -binding sequences that closely resemble the canonical EF-hand motif yet with diversified flanking structural elements. An easy and straightforward searching method to identify both canonical and pseudo EF-hands has been established based on our modified patterns. In addition to being supplementary to the signatures PS00018 and PS00303, the newly developed patterns convey information on the flanking structural contents with higher accuracy and sensitivity. Screening of the prokaryotic genome information revealed 397 entries of putative EF-hand proteins (467 motifs) implicated in a variety of cellular activities. The results enable us to envision possible scenarios regarding the evolutionary history of EF-hands. The pseudo EF-hands are likely to be phylogenetically younger than canonical EF-hand motifs. The prediction of Ca^{2+} binding motifs in bacteria genomes is helpful for the exploration of the role of Ca^{2+} and Ca^{2+} binding proteins in bacteria. Moving toward our goal of advancing calciomics, we have also carried out the prediction studies on other genomes, such as viruses (see **Chapter 4**), with our prediction method. This will further enable us to better understand the role of Ca^{2+} in diverse biological systems.

4. Ca^{2+} , Ca^{2+} -binding proteins and virus infection

Ca^{2+} , “a signal for life and death”, is one of the most important second messengers in transducing cellular signals. It plays a central role in the regulation of a number of cellular processes, such as cell division, differentiation and apoptosis (3). The versatile speed, amplitude and spatial-temporal patterning of Ca^{2+} in eukaryotic cells controls vital biological processes by finely modulating the activity of a repertoire of signaling components including cellular receptors, ion channels, pumps, exchangers, Ca^{2+} buffers, Ca^{2+} effectors, Ca^{2+} -sensitive enzymes and transcription factors in different cellular compartments (Fig. 4.1) (154).

Being extremely adept at hijacking the host cellular machinery, viruses have been extensively reported to interfere with the Ca^{2+} signaling pathways or Ca^{2+} -dependent processes, and thereby, achieve their optimal infectivity to produce progenies. The interplays between viruses and Ca^{2+} , in general, fall into three major categories: 1) viral particles or viral proteins directly or indirectly disturb the Ca^{2+} homeostasis by altering membrane permeability and/or manipulating key components of the Ca^{2+} -signaling repertoire; 2) a number of important structural or nonstructural viral proteins directly bind to Ca^{2+} for structural integrity or optimal functions; 3) virus-host interactions that require cellular Ca^{2+} -regulated proteins or processes. This chapter, which aims to stimulate more future studies on the role of Ca^{2+} in virus infection, is focused on

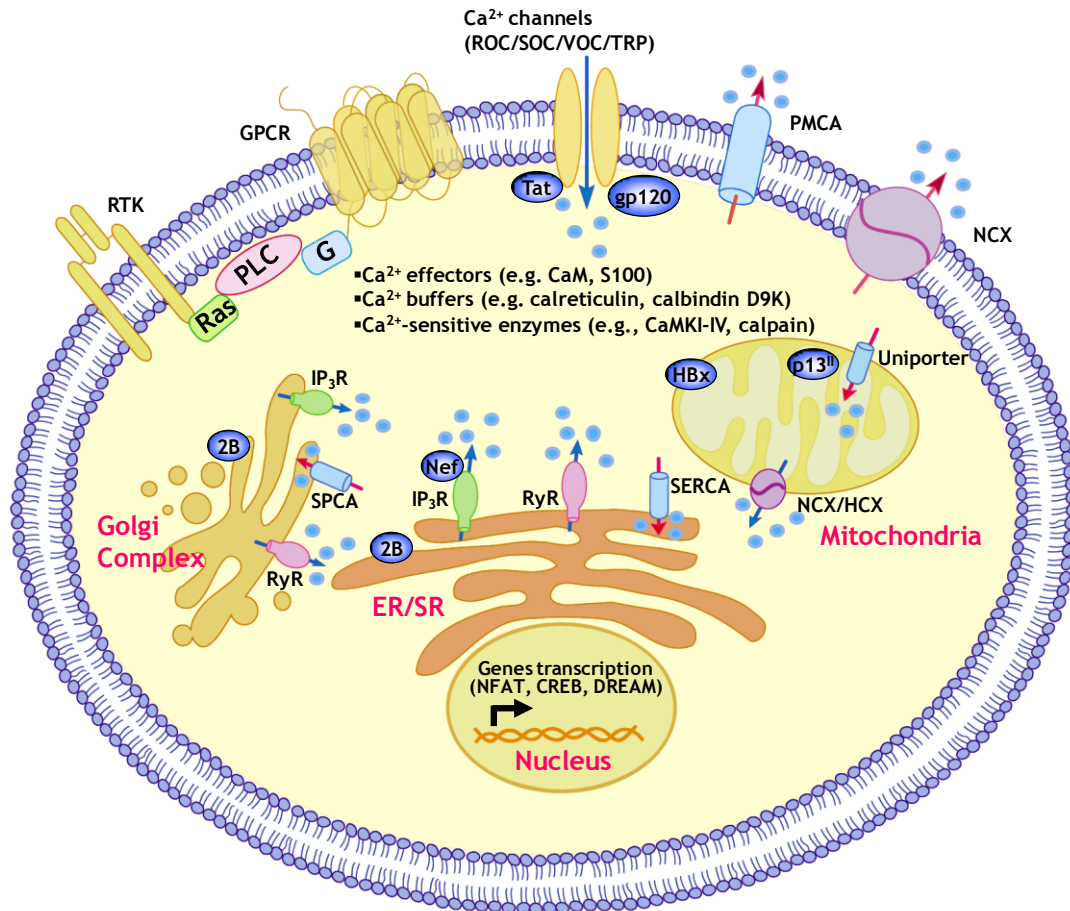


Figure 4.1. Examples of virus-induced perturbations on Ca²⁺ homeostasis. The versatile speed, amplitude and spatial-temporal patterning of Ca²⁺ in eukaryotic cells controls vital biological processes by precisely yet flexibly modulating the activity of a repertoire of signaling components including cellular receptors, ion channels, pumps, exchangers, Ca²⁺ buffers, Ca²⁺ effectors, Ca²⁺-sensitive enzymes and transcription factors in different cellular compartments. A variety of viral proteins (e.g., Tat, gp120 and Nef of HIV-1, p13^{II} of HTLV-1, HBx of HBV, 2B protein of coxsachievirus,) can disturb the sophisticated intracellular Ca²⁺ homeostasis by interacting with these key Ca²⁺ signaling components.

such interplays that have been observed in almost every step of a typical viral life cycle by featuring recent discoveries in the roles of Ca^{2+} in virus infection.

4.1. Effects of viral infections on Ca^{2+} homeostasis

The intracellular Ca^{2+} homeostasis is controlled by an array of components in the Ca^{2+} -signaling toolkit (154). The Ca^{2+} signal is derived from two sources: the extracellular medium or the internal stores (Fig. 4.1). The entry of Ca^{2+} across the plasma membrane, usually triggered by stimuli that include membrane depolarization, mechanical stretch, external agonists, depletion of internal stores and intracellular messengers, is mediated by some particular cellular receptors and Ca^{2+} channels. The mobilization of Ca^{2+} from the internal store in response to Ca^{2+} itself or intracellular messengers, is primarily mediated by the IP_3 receptors (IP_3R) and the ryanodine receptors (RyR). Five distinct pumping mechanisms (the plasma membrane Ca^{2+} -ATPase, the sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase (SERCA), the secretory pathway Ca^{2+} -ATPase (SPCA), the $\text{Na}^+/\text{Ca}^{2+}$ exchanger (NCX), and the mitochondrial uniporter) are responsible for sequestering Ca^{2+} from the cytosol by transport Ca^{2+} either to an external medium or into different cellular compartments (154).

The Ca^{2+} signaling system requires the exquisite choreography of the Ca^{2+} -signaling toolkit (Ca^{2+} “signalsomes”) and undergoes constant remodeling to meet the specific spatio-temporal requirements. This flexibility, on the one hand, provides sufficient opportunities for the host cells to adjust to the virus

infection. On the other hand, viruses are adept at utilizing the universal Ca^{2+} signal to create a tailored cellular environment that meets their own demands. The development of a variety of cell-permeable Ca^{2+} indicators (e.g., the acetoxymethyl ester forms of Fura-2, Fluo-4 and Indo-1) and genetically targetable fluorescent indicators (e.g., pericam and aequorin) provide solid foundation for the quantitative measurements of cellular Ca^{2+} signal changes in response to virus infection or the treatment of viral proteins. A detailed summary of virus-induced cellular Ca^{2+} -signaling alterations are listed in Table 4.1. Instead of exhaustively describing every virus that disturbs the Ca^{2+} homeostasis, this review will highlight several important viruses that are well-documented to alter the cellular Ca^{2+} -signaling system.

Table 4.1. Virus-induced alterations on cellular Ca^{2+} homeostasis.

Virus	$[\text{Ca}^{2+}]$ Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca^{2+} probe)	References
(+) ssRNA viruses						
Coronaviridae						
SARS-CoV	$\uparrow[\text{Ca}^{2+}]_{\text{CYT}}$	Spike protein	$\uparrow\text{Ca}^{2+}$ influx due to ER stress	$\uparrow\text{COX-2}$ expression via PLC pathway	HEK293T (Fura-2 AM)	(155)
Flaviviridae						
HCV	$\uparrow[\text{Ca}^{2+}]_{\text{MIT}}$	Core protein	Mitochondrial Ca^{2+} uniporter	\uparrow mitochondrial ROS production	Isolated liver mitochondria (Rhod-2 AM) Huh7 (Pericam)	(156)
	$\downarrow[\text{Ca}^{2+}]_{\text{ER}}$		ER stress, \uparrow Calreticulin, impaired SERCA activity	Induces apoptosis	Huh7 (erAEQ)	(157)
	$\uparrow[\text{Ca}^{2+}]_{\text{CYT}}$		ER Ca^{2+} leakage	Activate NFAT pathway	Jurkat T (Fura-2 AM)	(158)
	$\uparrow[\text{Ca}^{2+}]_{\text{CYT}}$	p7	Functions as Ca^{2+} channel	Channel activity inhibited by amantadine	Artificial membrane	(159)
	$\uparrow[\text{Ca}^{2+}]_{\text{CYT}}$	NS5A	\uparrow ER Ca^{2+} efflux	Activates NF- κ B and STAT-3, \uparrow ROS	Huh7	(160)
	$\downarrow[\text{Ca}^{2+}]_{\text{ER}}$	Virus infection	\uparrow ER Ca^{2+} release	Upregulates PP2A and inhibits IFN- α pathway	UHCV 57.3 (Fura-2 AM)	(161)
DV	$\uparrow[\text{Ca}^{2+}]_{\text{CYT}}$	virus infection	$\uparrow\text{Ca}^{2+}$ influx induced by CF	Enhances the cytotoxic activity	Mouse spleen (^{45}Ca)	(162,163)

Virus	[Ca²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca²⁺ probe)	References
JEV	↑[Ca ²⁺] _{CYT}	virus infection	↑Ca ²⁺ influx induced by MDF	Stimulate neutrophil activation	Mouse	(164)
Picornaviridae						
Poliovirus	↑[Ca ²⁺] _{CYT}	Virus infection	Voltage- sensitive Ca ²⁺ channels		HeLa (Fura-2 AM)	(165)
		2BC	↑Ca ²⁺ influx		HeLa (Fluo-3 AM)	(166)
Coxsachie B3 virus	↑[Ca ²⁺] _{CYT} , ↓[Ca ²⁺] _{ER} , ↓[Ca ²⁺] _{MIT} , ↓[Ca ²⁺] _{GOL}	2B	↑Ca ²⁺ influx; Pore formation on ER and Glogi membrane	Suppresses apoptotic host cell response	HeLa (Fura-2 AM, organelle- targeted AEQ)	(167,168)
	↑[Ca ²⁺] _{CYT} , ↓[Ca ²⁺] _{ER} ,	Virus infection and 2B protein	↑Ca ²⁺ influx; ER Ca ²⁺ leakage	Facilitates vRNA replication and virus release	HeLa (Fura-2 AM)	(169)
HRV2	↑[Ca ²⁺] _{CYT}	Virus infection	↑Ca ²⁺ influx	Channel blockers inhibit virus replication and release	HeLa T (Fluo-4 AM)	(170)
Togaviridae						
SFV	↑[Ca ²⁺] _{CYT} , ↑[Ca ²⁺] _{MIT} (<5 h), ↓[Ca ²⁺] _{MIT} (>5 h)	Virus infection	↑Ca ²⁺ influx; Impaired mithochonrial permeability	Impaired electron transport	CEC and isolated mithochondira (⁴⁵ Ca)	(171)

Virus	[Ca ²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca ²⁺ probe)	References
Sindbis virus	↓[Ca ²⁺] _{CYT} (5 h)	Virus infection		virus-induced apoptosis is dependent on viral replication	N18 (Fura-2 AM)	(172)
(-) ssRNA viruses						
Filoviridae						
Ebola virus and Marburg virus	↑[Ca ²⁺] _{CYT}	Virus infection		Activate TREM-1	Neutrophils (Fluo-4 AM, Fura-2 AM)	(173)
Orthomyxoviridae						
Influenza A virus	↑[Ca ²⁺] _{CYT} (>6 d)	Virus infection	↑ER Ca ²⁺ efflux via IP ₃ R	Reduces voltage- dependent Ca ²⁺ - currents	Hippocampal neurons (Fluo-4 AM)	(174)
	↑[Ca ²⁺] _{CYT}	Virus infection	↑IP ₃ (PLC pathway)	Activates neutrophils	Neutrophils (Fura-2 AM, ⁴⁵ Ca)	(175,176)
Paramyxoviridae						
Measles virus	↑[Ca ²⁺] _{CYT}	Virus infection		Viral yield is inhibited by verapamil	MA104, BGM (⁴⁵ Ca)	(177)
Sendai virus	↑[Ca ²⁺] _{CYT}	Virus infection		Maximizes viral fusion	Human erythrocyte ghosts (obelin)	(178)
	↑[Ca ²⁺] _{CYT}	Virus infection	↑Ca ²⁺ influx	Inducing rounding and fusion of infected cells	Chicken erythrocyte (⁴⁵ Ca)	(179)
Mumps virus	↓[Ca ²⁺] _{CYT}	Virus infection	↓Ca ²⁺ influx	Reduces voltage- dependent Ca ²⁺ - currents	Hippocampal neurons	(180,181)

Virus	[Ca ²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca ²⁺ probe)	References
(+) ssRNA viruses with DNA intermediates						
Retroviridae						
HIV-1	↑[Ca ²⁺] _{CYT}	Nef	Interacts with IP ₃ R	Promotes T cell activation via NFAT pathway	Jurkat T (Fluo-3, Fura-red AM, AEQ)	(182)
				Decreases store-operated Ca ²⁺ influx; Involves SH3-mediated interaction	Differentiated HL60 (Fura-2 AM)	(183)
	↑[Ca ²⁺] _{CYT}	Tat	↑Ca ²⁺ influx via DHP receptor ↑Ca ²⁺ influx via NMDA receptor	Increases TNF-α production	PBMC (Fluo-3 AM)	(184)
				Induces neurotoxicity	Rat hippocampal slices (Calcium-orange AM)	(185)
				Increases IP ₃ production; Induces neurotoxicity	Fetal neurons and astrocytes (Fura-2 AM)	(186)
				Stimulates cytokine and chemokine production	Microglia cells (Indo-1)	(187)
			↑Ca ²⁺ influx via glutamate receptor; ↑ ER Ca ²⁺ release via IP ₃ R ↑Ca ²⁺ influx	Mutation C31W in subtype C Tat reduces Ca ²⁺ flux	PBMC (Fluo-4 AM)	(188)

Virus	[Ca ²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca ²⁺ probe)	References
HTLV-1	↓[Ca ²⁺] _{CYT}	gp120	Blocks the L-type Ca ²⁺ channels ↑Ca ²⁺ influx via L-type Ca ²⁺ channels ↑Ca ²⁺ influx	Reduces glutamate and ATP-induced [Ca ²⁺] _{CYT} increase	Glial cells (Fura-2 AM)	(189)
	Inhibits NK cell function			NK cells (Fura-2 AM)	(190)	
	Induces neurotoxicity			Fetal neurons and astrocytes (Fura2-AM)	(191)	
	Activates CD4 ⁺ T cell signaling			CHO-745, T cells (Fluo-4 AM)	(192)	
	↑Ca ²⁺ influx			Embryonic neurons (Indo-1)	(193)	
	↑[Ca ²⁺] _i	gp120/gp160	↑Ca ²⁺ influx	Potentiates NMDA-induced Ca ²⁺ response	Synaptosomes (Fura2-AM)	(194)
	Induces cytotoxicity			HT-29-D4 (Fura2-AM)	(195)	
	↑[Ca ²⁺] _{CYT}	virus infection	↑ ER Ca ²⁺ release via IP ₃ R; ↑Ca ²⁺ influx		H9	(196)
		virus infection		Induces T-cell hyporesponsiveness	PMBC, MT4, Jurkat T	(197)
	↑[Ca ²⁺] _{CYT}	P12 ¹		Activates T cells via NFAT pathway	Jurkat T (Indo-1, Fura-2 AM)	(198)
				Enhances p300 transcription		(199)

Virus	[Ca ²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca ²⁺ probe)	References
	↓[Ca ²⁺] _{MIT}	P13 ^{II}	↑Mitochondrial membrane permeability	Induces apoptosis	Isolated mitochondria (Calcium- Green)	(200-202)
Reoviridae						
Rotavirus	↑[Ca ²⁺] _{CYT}	Virus infection	↑PM permeability; ↓ER Ca ²⁺ pool	Tunicamycin and brefeldin A inhibits ↑[Ca ²⁺] _{CYT}	HT29, MA104 (⁴⁵ Ca, Fura-2 AM)	(203)
			↑PM permeability; ↑ER Ca ²⁺ release via PLC pathway	Facilitates virus entry and maturation	Caco-2 (Quin-2 AM)	(204)
			↑Ca ²⁺ influx via L-type Ca ²⁺ channels		HT29, MA104 (Fura-2 AM)	(205)
		NSP4	↑ ER Ca ²⁺ release via IP ₃ R; ↑Ca ²⁺ influx		HEK293, sf9 (Fura-2 AM)	(206)
CAV	↑[Ca ²⁺] _{CYT}	Virus infection			Lymphocytes	(207)
ssDNA viruses						
Parvoviridae						
B19	↑[Ca ²⁺] _{CYT}	VP1	Activation of I _{CRAC}	Exhibits PLA2-like activity	HMEC-1 (Fura-2)	(208)
dsDNA viruses						
Herpesviridae						

Virus	[Ca²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca²⁺ probe)	References
CMV	↑[Ca ²⁺] _{CYT}	UL37x1	↑ ER Ca ²⁺ release	Blocks apoptosis and induces mitochondrial fission	Fibroblast (Calcium- Orange, Fluo- 3)	(209)
		Virus infection	↑Ca ²⁺ influx	Inhibited by nifedipine and verapamil	Fibroblast	(210)
HHV-1/2	↑[Ca ²⁺] _{CYT}	Virus infection	↑ ER Ca ²⁺ release via IP ₃ R; ↑Ca ²⁺ influx via voltage-gated channels	Triggers FAK phosphorylation	CaSki, Vero (Fura-2 AM)	(211)
	↑[Ca ²⁺] _{CYT} (H ₂ O ₂ - induced)	Virus infection		H ₂ O ₂ enhances virus release	FI cells (Fura-2 AM)	(212)
KSHV	↑[Ca ²⁺] _{CYT}	K1	Interacts with SH2- containing proteins	Activates T cells	293T(Indo-1)	(213)
		vMIP I/II	↑Ca ²⁺ influx	Induces signal transduction and chemotaxis	THP-1, K562 (Indo-1 AM)	(214)
		K7	↑ ER Ca ²⁺ release	Associates with CAML to inhibit apoptosis	BJAB (Indo-1 AM)	(215)
EBV	↑[Ca ²⁺] _{CYT}	Virus infection	↑Ca ²⁺ influx Ca ²⁺ channels	Activates B cells; Inhibited by verapamil or diltiazem	B95-8 (Quin-2)	(216)

Virus	[Ca ²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca ²⁺ probe)	References
Polyomaviridae SV40	↓[Ca ²⁺] _{CYT}	LMP2A	Blocks BCR	Blocks B cell signal transduction	BJAB (Indo-1)	(217)
	↑[Ca ²⁺] _{CYT}		↑Ca ²⁺ influx via capacitative channels		SV3T3 (AEQ)	(218)
	↑[Ca ²⁺] _{CYT}	Virus infection	Promotes progression of tumorigenicity		Keratinocytes	(219)
HPV-18	↑[Ca ²⁺] _{CYT}	Virus infection				
Poxviridae Vaccinia virus	↑[Ca ²⁺] _{CYT}	Virus infection		Viral yield is inhibited by verapamil	MA104, BGM (⁴⁵ Ca)	(177)
		A38L	↑Ca ²⁺ influx	Increases PM permeability; decreases production of infectious viral particles	BS-C-1 (⁴⁵ Ca)	(220)
	↓[Ca ²⁺] _{CYT} (agonist-induced)	Virus infection	↓PLC production or activity	Alter agonist-induced Ca ²⁺ homeostasis	BS-C-1 (Fura-2 AM)	(221)
dsDNA viruses with RNA intermediates						
Hepadnaviridae HBV	↑[Ca ²⁺] _{CYT}	HBx	Intracellular Ca ²⁺ store release	Transactivates JNK and MAPK pathways	CHL-X (Fura-2 AM)	(222)

Virus	[Ca ²⁺] Alteration	Caused by	Mechanisms	Cellular effects	Model System (Ca ²⁺ probe)	References
			↓Mitochondrial Ca ²⁺ uptake; Impaired PMCA activity	Induces apoptosis	HepG2, HeLa (organelle- targeted AEQ)	(223)
			Mitochondrial Ca ²⁺ leakage	Activated PyK2	HepG2	(224)

Abbreviations: [Ca²⁺]_{CYT}, cytosolic Ca²⁺ concentration; [Ca²⁺]_{ER}, endoplasmic reticulum Ca²⁺ concentration; [Ca²⁺]_{MIT}, mitochondrial Ca²⁺ concentration; [Ca²⁺]_{GOL}, Golgi complex Ca²⁺ concentration; [Ca²⁺]_i, intrasynaptosomal Ca²⁺ concentration; SARS-CoV, severe acute respiratory syndrome-associated coronavirus; COX-2, cyclooxygenase-2; HCV, hepatitis C virus; SERCA, sarcoplasmic/endoplasmic reticulum Ca²⁺-ATPase; PP2A, protein phosphatase 2A; IFN, interferon; erAEQ, ER-targeted aequorin; NFAT, nuclear factor of activated T cells; DV, dengue virus; CF, cytotoxic factor; JEV, Japanese encephalitis virus; MDF, macrophage derived neutrophil chemotactic factor; HRV2: human rhinovirus 2; SFV, Semliki Forest virus; TREM, triggering receptors expressed in myeloid cells; CEC: primary chicken embryo cells; PLC, phospholipase C; IP₃R, inositol 1,4,5-triphosphate receptor; HIV, human immunodeficiency virus; Nef, negative factor; Tat, transactivator of transcription; DHP, dihydropyridine; PBMC, peripheral blood mononuclear cell; TNF, tumor necrosis factor; NK, natural killer cell; NMDA, N-methyl-D-aspartic acid, HTLV, human T-lymphotropic virus; PM, plasma membrane; CAV, chicken anemia virus; CRAC, Ca²⁺ release activated Ca²⁺ channel; PLA2, phospholipase A2; CMV: cytomegalovirus; HHV, human herpesvirus; FAK, focal adhesion kinase; KSHV, Kaposi's sarcoma-associated herpesvirus; CAML, Ca²⁺-modulating cyclophilin ligand; EBV, Epstein-Barr virus; LMP2A, latent membrane protein 2A; HPV, human papillomavirus; HBV, hepatitis B virus; JNK, ; MAPK, ; PMCA, plasma membrane Ca²⁺-ATPase; PyK2, proline-rich tyrosine kinase-2;

Hepatitis C virus. Hepatitis C virus (HCV), one of the major causative agents of hepatitis with a single-stranded positive-sense RNA genome, belongs to the *Flaviviridae* family. HCV infection has been shown to decrease the ER Ca^{2+} concentration and activate ER stress by upregulating the protein phosphatase 2A (PP2A), leading to the inhibition of interferon alpha signaling. Three viral proteins, the core protein, NS5A and p7, have been found to interfere with the cellular Ca^{2+} signaling.

HCV core protein, which induces apoptosis in Huh7 cells, has been recently reported to induce ER Ca^{2+} depletion by impairing the function of SERCA2 (157). The core protein also increases mitochondrial ROS production and Ca^{2+} uptake partially by stimulating the activity of mitochondrial Ca^{2+} uniporter (156). It can also lead to the elevation of cytosolic Ca^{2+} concentration due to ER Ca^{2+} leakage, and thereby activates the NFAT pathway in Jurkat T cells (158). All these findings indicate that the core protein potently modulates the apoptosis of HCV-infected cells and is possibly responsible for chronic liver disease and tumor transformation.

The nonstructural protein NS5A, which is embedded on the ER membrane, has been shown to alter the intracellular Ca^{2+} levels with increased ER Ca^{2+} efflux. The released Ca^{2+} can be readily taken by mitochondria and increase the production of mitochondrial oxidants. As a consequence, the NS5A functions as a transcriptional transactivator to induce the activation of NF- κ B and STAT3 (160).

Another viral protein p7, a small hydrophobic uncharacterized protein on the ER membrane, has been found to form hexamers and function reminiscent of a Ca^{2+} ion channel on artificial lipid membrane (159). The reconstituted channel can be blocked by amantadine or long-alkyl-chain iminosugar derivatives. This finding raises the possibility that p7 might be responsible for the flow of Ca^{2+} from ER to the cytoplasm.

Enterovirus. Poliovirus and coxsachievirus are two typical enteroviruses in the *picornaviridae* family that contains a positive-sense ssRNA. The single RNA genome encodes four structural proteins (VP1-4) and ten nonstructural proteins (2A^{pro} , 2B, 2C, 3A, 3B, 3C^{pro} , 3D^{pol} , 2BC, 3AB, and 3CD^{pro}). Among these, the 2BC protein of poliovirus has been shown to induce an increase in cytosolic Ca^{2+} concentration, possibly via the voltage-sensitive Ca^{2+} -channels 2 or 3 h postinfection. This time point is coincident with a period during which the viral proteins are actively synthesized (165). The 2B protein of coxsachievirus leads to the decrease of Ca^{2+} concentrations in subcellular compartments, such as mitochondria, Golgi complex and ER, by forming pores on these membranes and subsequently increasing Ca^{2+} efflux from these organelles. The substantial decrease of Ca^{2+} in ER and Golgi complex has been correlated with the inhibition of protein trafficking pathways, and thus down-regulating host anti-viral immune response (225). More interestingly, the perturbation of Ca^{2+} homeostasis by the 2B protein induces the inhibition of

apoptotic cell response triggered in early infection and serves to provide the virus a suitable time window for viral RNA replication and protein synthesis (166-169).

Human immunodeficiency virus type 1 (HIV-1). HIV, the etiological agent of acquired immune-deficiency syndrome (AIDS), contains a 9.2 Kb single-stranded RNA genome that encodes for polyproteins processed into matrix protein, capsid protein, nucleocapsid (NC), gp 120, gp41, protease, reverse transcriptase, integrase and p6, as well as six accessory proteins (Vif, Vpr, Nef, Tat, Rev and Vpu). Three HIV-1 proteins, namely gp120, Nef, and Tat, have been individually shown to alter the Ca^{2+} homeostasis in HIV-infected cells.

Gp120, one of the envelope glycoproteins that forms the HIV-1 surface spikes, is essential for the attachment of HIV to the membrane immunoglobulin CD4 and viral entry into host cells. The processing of the precursor protein gp160 into gp120 and gp41 has been shown to be a Ca^{2+} -dependent process (226). Furthermore, gp120 is capable of inducing neurotoxicity by increasing Ca^{2+} influx into the cytoplasm of fetal neurons and astrocytes via Na^+/H^+ exchangers, L-type voltage-sensitive Ca^{2+} channels and NMDA receptors (191,193). Gp120 can also induce the elevation of Ca^{2+} in presynaptic terminals of rat cortical synaptosomes (194) and in the cytoplasm of human intestinal epithelial cells HT-29-D4 (195). These findings suggest that drugs that block gp120-induced intracellular Ca^{2+} changes in the human brain might be potential candidates for the treatment of HIV-1 dementia and enteropathy.

Nef, a highly conserved accessory protein among primate lentiviruses, plays a major role in the down-regulation of immunologically relevant cell surface protein CD4 and MHCI (227). In addition, Nef has been demonstrated to induce cytosolic Ca^{2+} increase through its interaction with IP_3R residing on ER membrane, and subsequently promotes the T cell receptor-independent activation of the NFAT pathway in Jurkat T cells (182). The activated NFAT, a transcriptional factor responding to the low-amplitude intracellular Ca^{2+} oscillations, can further promote the viral gene transcription and replication (228). The increase in cytosolic Ca^{2+} is also demonstrated in differentiated myelomonocytic HL60 cells (183). The elevation in the intracellular Ca^{2+} storage seems to decrease store-operated Ca^{2+} influx and might be associated with Nef-related pathophysiology. In addition, it is interesting to note that a Src-like protein-tyrosine kinase (PTK) coimmunoprecipitates with both Nef and the IP_3R . Given this, it is hypothesized that Nef might modulate the activity of the IP_3R activity via its interaction with the Src-like PTK (183).

Tat, the HIV transcriptional transactivator, is mainly responsible for the regulation of the viral gene expression and replication. Tat has been reported to either increase or decrease cytosolic Ca^{2+} in mainly two types of cells: the immune cells and neuronal cells. Mediated by the L-type Ca^{2+} channel DHP receptors, Tat has been shown to promote extracellular Ca^{2+} influx to the cytoplasm and stimulate the production of $\text{TNF}\alpha$ in primary human monocytes (184). In comparison, the treatment of NK cells with HIV Tat blocks the phenylalkylamine-sensitive L-type Ca^{2+} channels and impairs the cytotoxic

function mediated by NK cells, thus exacerbating progressive immunosuppression during HIV infection (190). The increase in cytosolic Ca^{2+} , due to either increased Ca^{2+} influx via NMDA type glutamate receptor or the release of Ca^{2+} from IP_3 -sensitive internal stores via IP_3R , is observed in both fetal neurons, astrocytes and microglia cells (185-187). The Tat-induced dysregulation of intracellular Ca^{2+} leads to neurotoxicity and contributes to HIV-related dementia.

Taken together, these three HIV-1 proteins function in concert to alter intracellular Ca^{2+} signaling and further regulate T cell activation, apoptosis and cell proliferation.

Human T-lymphotropic virus type 1 (HTLV-1). HTV -1, also called adult T-cell lymphoma virus type 1, is a human retrovirus that causes adult T-cell leukemia/lymphoma, HTLV-I-associated myelopathy, and Strongyloides stercoralis hyper-infection. The 8.5 Kb, single-stranded RNA genome encodes the gag, pol, pro and env genes, as well as five accessory proteins (p12^{I} , p13^{II} , p21^{Rex} , p27^{I} , p30^{II} and) and two regulatory proteins (Tax and Rex). HTLV-1 dysregulates the Ca^{2+} signaling mainly through two accessory proteins: p12^{I} and p13^{II} .

The accessory protein p12^{I} , residing on the membrane of ER and Golgi complex, increase cytosolic Ca^{2+} by triggering Ca^{2+} release from ER possibly via IP_3R , and meanwhile, enhances Ca^{2+} influx through Ca^{2+} release-activated Ca^{2+} channels (198). The elevated basal intracellular Ca^{2+} leads to the activation of

nuclear factor activated T-cell (NFAT) mediated transcription, and thus promoting lymphocyte activation to support early viral infection. p12^I has also been shown to interact with Ca²⁺-responsive protein, p300, to mediate various transcription factors (199) and further promote lymphocyte survival.

The other accessory protein p13^{II} is an 87-amino-acid mitochondrial protein that can cause swelling and depolarization of mitochondria by increasing inner membrane permeability to cations, such as Ca²⁺, Na⁺ and K⁺ (200). Such changes are believed to be responsible for promoting ceramide- or Fas ligand-induced apoptosis in T lymphocytes expressing p13^{II} (201).

Thus, the two viral accessory proteins that are targeted to different cellular compartments function independently to modulate apoptosis and proliferation of infected cells.

Rotavirus. Rotavirus, a member of the *reoviridae* family that contains double-stranded RNA genomes, is the major etiological agent of viral gastroenteritis in young children. Ca²⁺ plays a key role in the morphogenesis, replication and pathogenesis of rotavirus (229). Rotavirus-infected MA104 and HT29 cells have been shown to exhibit a progressive elevation of cytosolic Ca²⁺ due to increased membrane permeability (205). The entry of Ca²⁺ is partially blocked by L-type voltage-sensitive Ca²⁺ channel blocker metoxyverapamil (D600). More importantly, the permeability pathway is selectively permeable for the passage of divalent ions that include Ca²⁺, Ba²⁺, Sr²⁺, Mn²⁺ and Co²⁺, but is impermeable to the trivalent cations La³⁺ and Cr³⁺. Such metal ion selectivity is

observed commonly in divalent cation channels in both excitable and nonexcitable cells (230). Other mechanisms that involve both phospholipase C (PLC)-dependent and PLC-independent pathways have also been shown to account for intracellular Ca^{2+} increase (206,231). The exogenous NSP4, which can induce diarrhea in rodent pups, increases cytosolic Ca^{2+} concentration via the activation of PLC and the resultant ER Ca^{2+} depletion through IP_3R (206). In contrast, endogenously expressed NSP4 in Sf9 cells and HEK293 alters the ER membrane permeability and causes a sustained increase of cytosolic Ca^{2+} concentration that is independent of the PLC pathway (206,231). In this regard, further investigation is needed to elucidate the mechanisms underlying the alteration of intracellular Ca^{2+} homeostasis in rotavirus-infected cells.

Herpesviruses. An increase in intracellular Ca^{2+} has been observed in epithelial and T cells infected with HHV-1, HHV-2 (211), HHV-4 (Epstein-Barr virus, EBV) (216,217), HHV-5 (cytomegalovirus, CMV) (210), and HHV-8 (Kaposi's sarcoma-associated herpesvirus, KSHV) (213-215). The alteration in Ca^{2+} signaling leads to a variety of cellular responses in cells transfected with different HSV.

Exposure of epithelial cells (CaSki and Vero) with HHV-1 and HHV-2 has been shown to result in a rapid and transient increase in intracellular Ca^{2+} by triggering ER Ca^{2+} release via IP_3R , but not Ca^{2+} influx through plasma membrane. Such change in turn triggers FAK phosphorylation 5-10 min following virus exposure (211). Notably, FAK phosphorylation has been demonstrated to

promote actin cytoskeleton reorganization and virus trafficking. In this way, altered Ca^{2+} signaling is directly linked to the phosphorylation event and plays a central role in facilitating early events in HHV-1/HHV-2 invasion.

Epstein-Barr virus (EBV), or HHV-4, is the etiological reagent for infectious mononucleosis and a number of malignant diseases, such as Burkitt's lymphoma and nasopharyngeal carcinoma. It has been shown that EBV infection can induce a rise in intracellular Ca^{2+} due to increased Ca^{2+} influx from the extracellular space (216). The rise in intracellular Ca^{2+} , which is associated with protein kinase C translocation from the cytosol to membrane bound compartments and the activation of B lymphocytes, can be blocked by verapamil, a L-type Ca^{2+} -channel blocker (232). The EBV latency-associated gene product latent membrane protein 2A (LMP2A) has been shown to function as a negative modulator of the activation of B lymphocytes by binding to the Src family kinases Lyn and Syk (233). Such interaction is responsible for blocking the Ca^{2+} mobilization in B lymphocytes and inhibiting the induction of lytic infection of EBV, and thus evading the host immune clearance of EBV infection (217).

Infection of fibroblast cells by cytomegalovirus (CMV), or HHV-5, has been shown to induce more than 2-fold increase in cytosolic Ca^{2+} (210). The rise can be further blocked by nifedipine and verapamil, indicating that L-type Ca^{2+} channel might contribute to the intracellular Ca^{2+} increase. Moreover, a CMV-encoded protein pUL37x1, which is capable of trafficking from ER to mitochondria, can induce Ca^{2+} release from the ER store with concomitant morphological changes, such as cell rounding, cell swelling, actin cytoskeleton

reorganization and mitochondrial fission (209). The altered Ca^{2+} -signaling is further linked to the antiapoptotic activity of pUL37x1.

Kaposi's sarcoma-associated herpesvirus (KSHV), also called HHV-8, is the etiological agent of Kaposi's sarcoma and closely associated with primary effusion lymphoma and multicentric Castleman's disease (234,235). Alteration in Ca^{2+} -signaling has been most often reported in cells expressing K1, K7 and viral macrophage inflammatory protein (vMIP) -I and -II, all of which are abundantly expressed during lytic phase of the viral life cycle. The KSHV K1 protein is a lymphocyte receptor-like protein that can transduce external stimuli to elicit downstream signaling pathways and induce cytokine production through its cytosolic immunoreceptor tyrosine-based activation motif (ITAM). By interacting with cellular Src homology 2-containing proteins including Lyn, Syk, p85, PLC γ 2, and RasGAP, the expression of K1 protein leads to the intracellular Ca^{2+} immobilization and the subsequent activation of NFAT and AP-1 pathways (213). The K7 protein is a mitochondria-targeted protein expressed during the lytic replication of KSHV. K7 has been shown to enhance ER Ca^{2+} release through its interaction with the cellular Ca^{2+} -modulating cyclophilin ligand and further inhibit apoptosis (215). Thus, the lytic K7 protein protects KSHV-infected cells from apoptosis and allows the completion of the viral lytic replication. vMIP-1 and vMIP-II are chemokine homologs involved in the signal transduction and chemotaxis of monocytes and T lymphocytes. Both proteins are capable of inducing transient Ca^{2+} mobilization by activating the CC chemokine receptor 5 in

K562 cells. In addition, vMIP-1 can induce Ca^{2+} signal through its interaction with CC chemokine receptor 8(214).

HBV. HBV is a para-retrovirus that infects 3-6% of the world's population. The multifunctional hepatitis B virus X protein (HBx) has been shown to alter the Ca^{2+} signaling to benefit viral replication and core assembly. Upon HBx expression, the intracellular Ca^{2+} increases due to Ca^{2+} release from ER and mitochondria. The rise in cytosolic Ca^{2+} triggers the activation of a Ca^{2+} -dependent proline-rich tyrosine kinase-2, which in turn activates the Src kinase and up-regulates reverse transcription (224). More strikingly, such effects exerted by HBx can be substituted with simple increase in cytosolic Ca^{2+} , suggesting a fundamental requirement of Ca^{2+} for HBV replication and infection. More than that, the increased Ca^{2+} by HBx has also been shown to enhance the HBV core assembly (40). Treatment of transfected HepG2 cells with the cell-permeable Ca^{2+} -chelator BAPTA-AM and cyclosporine A results in reduced HBV capsid assembly, whereas the application of thapsigargin, a strong inhibitor of the sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase (SERCA), increased HBV core assembly. Overexpression of HBx has been further demonstrated to stimulate caspase-3-dependent cleavage of the plasma membrane Ca^{2+} -ATPase (PMCA), a pump that extrudes Ca^{2+} from cytoplasm to extracellular environment. HBx is also linked to apoptosis by inducing morphological changes and reducing Ca^{2+} uptake in mitochondria (223).

4.2. Viral Ca^{2+} -binding proteins

Ca^{2+} plays important roles in the replication cycle of some of viruses through its direct interactions with viral proteins. For some viruses, such as polyomaviruses (33,34), rotavirus (36,38,39,236), hepatitis B virus (40), turnip crinkle virus (41,42) and tobacco mosaic virus (33,43), Ca^{2+} is required to maintain the structural integrity and/or the proper assembly of virions. Ca^{2+} is required for cell fusion mediated by human immunodeficiency virus type 1 envelope glycoprotein gp 160 (44,45). For influenza B virus, the binding of Ca^{2+} has been reported to enhance the stability of the virion associated neuraminidase and to promote its enzymatic activity (46,237). To date, the majority of the reported viral CaBPs are structural proteins, including both capsid proteins and envelope proteins.

According to the location of Ca^{2+} -coordinating residues, Ca^{2+} -binding sites can be primarily divided into two types: discontinuous and continuous Ca^{2+} -binding sites. Both types of Ca^{2+} -binding sites have been found in different families of viruses (Table 4.2).

Discontinuous viral Ca^{2+} -binding sites

In discontinuous Ca^{2+} -binding sites, Ca^{2+} -coordinating ligand residues that are distantly separated in the primary sequence or originate from different polypeptides approximate in space to form Ca^{2+} -binding pockets. By searching through the Protein Data Bank (PDB), such Ca^{2+} -binding sites have been most frequently found in the coat protein of both RNA viruses (e.g., enteroviruses,

rhinovirus, sobemoviruses and rotavirus) and DNA viruses (e.g., parvoviruses and polyomaviruses), as well as in the NSP4 of rotavirus and the neuraminidase of Influenza viruses (Table 4.2).

Coat protein

The binding of viral coat protein to Ca^{2+} ions have been seen in some plant viruses (such as Sobemoviruses, Tobamoviruses and some viruses in the family of *Tombusviridae*), as well as a number of animal viruses including the picornaviruses, rotavirus, parvoviruses and polyomavirus (Table 4.2). In these viruses, Ca^{2+} ions are required for the efficient assembly and/or disassembly of viral particles. In the extracellular milieu, the Ca^{2+} concentration is maintained at millimolar level, whereas the Ca^{2+} concentration drops ~100-1000 fold in the cytoplasm. The viruses can make use of the much lower cytosolic Ca^{2+} concentration to initiate the uncoating event (238,239). Upon entry into the host cells, the viral particles undergo swelling or “solubilization” due to the repulsive forces between negatively charged residues that are otherwise neutralized by the bound Ca^{2+} . The incorporated Ca^{2+} ions may situate between the interacting interfaces of capsid subunits (Fig. 4.2A) or sit on the symmetric threefold or fivefold axis (Fig. 4.2B). Moreover, the numbers of bound Ca^{2+} and coordinating geometry of the Ca^{2+} -binding sites also differ for different viruses. Representative examples can be found in both plant and animal viruses that have icosahedral or helical (e.g., tobacco mosaic virus) capsids (Table 4.2).

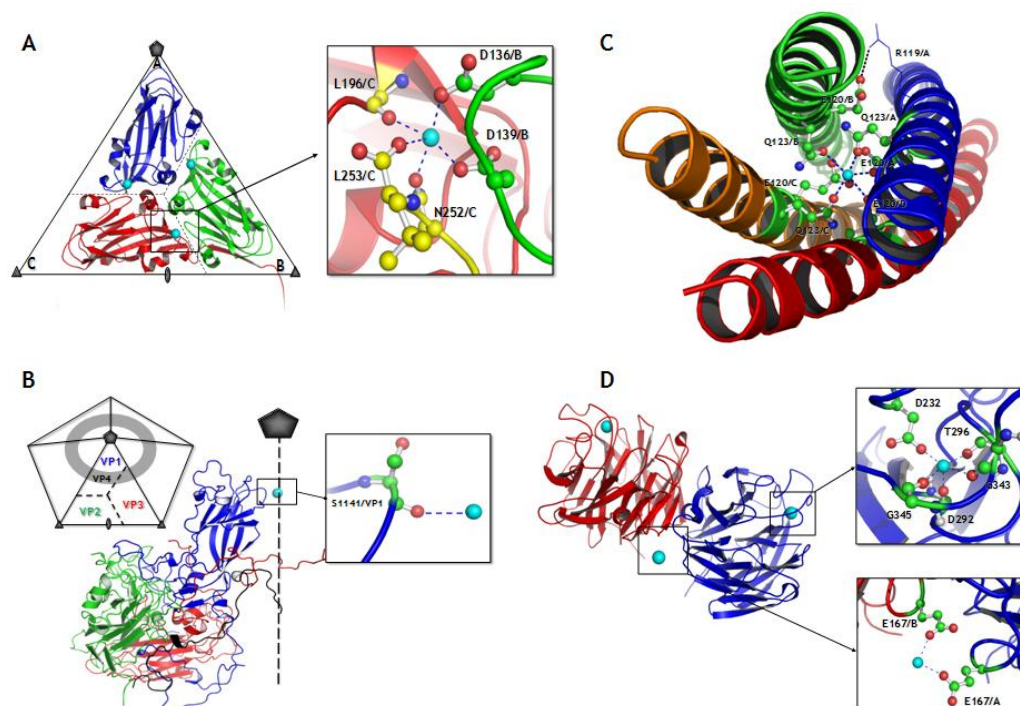


Figure 4.2. Examples of viral Ca^{2+} -binding proteins with determined 3D structures. (A), 3D representation of the icosahedral asymmetric unit of the cocksfoot mottle virus capsid and the location of the incorporated Ca^{2+} ions (PDB code: 1ng0). The assembling unit is formed by three subunits, A (blue), B (green) and C (red) that are chemically identical but slightly different in conformational arrangement. Ca^{2+} situates between the interfaces of neighboring subunits (A-B, A-C or B-C). The solid pentagon, triangle and oval represent five-, three-, and two-fold axes of the icosahedron. (B), Ca^{2+} ion located on the fivefold axis of the capsid of human rhinovirus 3 (HRV3) (PDB code: 1rhi). The icosahedral capsid of HRV3 is composed of 60 copies of each of the four capsid proteins VP1 (blue), VP2 (green), VP3 (red) and VP4 (black). VP1, VP2 and VP3 are exposed to the external surface of the viral particle, whereas VP4 lines in the internal surface. A Ca^{2+} ion is found situated on the fivefold axis of the capsid and is coordinated by 5 oxygen atoms from the main chain carbonyl group of S¹¹⁴¹ on VP1 (enlarged area). (C), The core Ca^{2+} -binding pocket in the oligomerization domain of NSP4 from rotavirus (PDB code: 2o1j). The domain self-assembles into a paralleled tetrameric coiled-coil. Chains A, B, C and D are shown in blue, green, orange and red, respectively. The Ca^{2+} ion is coordinated by six oxygen atoms from the side chains of Q¹²³ on chains A to D, as well as the side chains of E¹²⁰ on chains A and C. In addition, the E¹²⁰ on chains B or D form salt bridges with the residue R¹¹⁹ on the neighboring chains A or C. (D), The 3D structure of neuraminidase of influenza B virus (PDB code: 1nsb). The cartoon only represents half of the tetrameric form of this enzyme. Three Ca^{2+} -binding sites are found in two identical subunits, A (blue) and B (red). Each subunit contains one octahedral Ca^{2+} -binding site (upper panel). Another site (lower panel), coordinated by the fourfold symmetry-related E¹⁶⁷.

Table 4.2. Roles of Ca^{2+} in the life cycle of virus and viral Ca^{2+} -binding proteins (CaBPs).

Virus	Roles of Ca^{2+}	Viral CaBPs (proposed Ca^{2+} -binding site)	Methods (K_d , μM)	Reference
(+) ssRNA viruses				
Dengue Virus	Virus release			(240)
Human rhinovirus	Virus stability; acid-induced uncoating	Capsid proteins [PDB code: 1rhi; 1v9u] (D1137[HRV1A]; S1141[HRV3]; Site 1: N1141; Site 2: E3200[HRV14])	X-ray	(170,241,242)
Coxsackievirus A21	Capsid stability	VP1 [PDB code: 1z7s] (S1021/S1024/T1022/N1067)	Cryo-EM	(243)
Coxsackievirus B3	Capsid stability	VP1 (D1133); VP3 (D3203)	X-ray	(244)
HAV	Virus attachment			(245,246)
STNV	Capsid stability	Coat protein [PDB code: 2buk] (Site 1: E25/S61/Q64/D194; Site 2: D55; Site 3: T138)	X-ray; Fluorescence (Eu^{3+} : 0.0011)	(247-249)
Rubella virus	↑viral protease stability	NSP (D1206/S1208/D1210/T1012/D1217)	Fluorescence; NMR (La^{3+} : 14; Tb^{3+} : 47; Ca^{2+} : 247)	(250)
TBSV	↑virus swelling	Coat protein [PDB code: 2TBV] (aa 149-155; 181-186)	X-ray; Cryo-EM	(251-253)
TCV	Virus movement	Coat protein (D155/D157/K160/E127/D199)	X-ray	(41,42,238)
SeMV	Capsid stability and assembly	Coat protein [1smv] (D145/D149/Y205/N267/N268)	X-ray	(254)
CfMV	Capsid stability and assembly	Coat protein [1ng0] (D136/D139/L196/N252/L253)	X-ray	(255)
RYMV	Capsid stability and assembly	Coat protein [1f2n] (D126/D129/V182/N237/T238)	X-ray	(142)
SCPMV	Capsid stability and assembly	Coat protein [4sbv] (D138/D141/F199/N259/L260)	X-ray	(256,257)

Virus	Roles of Ca ²⁺	Viral CaBPs (proposed Ca ²⁺ -binding site)	Methods (K _d , μM)	Reference
TMV	Capsid stability and assembly	Coat protein	X-ray; Electrode (Ca ²⁺ : ~100)	(33,43)
RCNMV	Capsid stability	Coat protein (aa 93-99; 125-130)	Cryo-EM	(258)
TNV	Capsid stability and assembly	Coat protein [1c8n] (Site 1: D160/D163/T219/N275; Site 2: N93/S94/S173/N175)	X-ray	(259)
(-) ssRNA viruses				
Ebola virus	Triggers conformational change; ↑fusion with PM	GP/fusogenic domain	CD	(260)
Hantavirus	↑N-N interactions	N protein (aa 401–412)		(261)
Influenza A virus	↑enzymatic activity and virus fusion	Neuraminidase [PDB code: 3nn9] (D293/G297/N347/D324)	X-ray	(262-266)
Influenza B virus	↑thermostability and enzymatic activity	Neuraminidase [PDB code: 1nsb] (Site 1: E167; Site 2: D292/T296/D323/G343/G345)	X-ray	(46,237)
Parainfluenza viruses	Enzymatic activity	HN protein [PDB code: 1z4v; 1v2i; 1e8t] (D250/S253/A255/A285)	X-ray	(267-269)
Sendai virus	Virus morphogenesis; envelope protein transportation;			(270-272)
	Within the core of tetrameric coiled coil	Phosphoprotein ₃₂₀₋₄₃₃ [PDB code: 1ezj] (N389)	X-ray	(273)
RSV	↑fusion protein synthesis and syncytium formation			(274)
(+) ssRNA viruses require DNA intermediates				

Virus	Roles of Ca ²⁺	Viral CaBPs (proposed Ca ²⁺ -binding site)	Methods (K _d , μM)	Reference
HIV-1	Triggers conformational change (α-helix→β sheet); ↑virus fusion to PM Virus fusion to PM	gp41 N-terminal fusion peptide	IR	(275)
	gp160 processing virus attachment and cell fusion	gp41 ₆₂₈₋₆₈₃ (aa 628-648)	Fluorescence (Ca ²⁺ : 40)	(44,276) (226,277) (45,278,279)
ASV	Alters enzymatic activity	Integrase [PDB code: 1svi] (D64/D121)	X-ray	(280,281)
BLV	↑gp51 and p24 synthesis			(282)
dsRNA viruses				
IBDV	Capsid stability and assembly	VP2 [PDB code: 2gsy; 2df7] (D31/D174)	X-ray	(283,284)
Rotavirus	Virus maturation and stability Capsid stability; assembly	VP7 (aa 270-282) VP6 [PDB Code: 1qhd] (Site 1: Y248/N250; Site 2: D340/161/E134/D337; Site 3: D62/Q47; Site 4: D286/N266; Site 5: E379/D380/Q383/Q268)	Cryo-EM	(38,39,229) (285)
	NSP4 oligomerization	NSP4 [PDB code: 1g1i; 2o1k] (E120/Q123)	X-ray	(286)
ssDNA viruses				
CPV/FPV	capsid stability and assembly	VP2 [PDB code: 1c8d; 1c8g] (Site 1: D373/D375/G362/G363; Site 2: D237/D239; Site 3: D237/D240/D405)	X-ray	(287,288)
dsDNA viruses				
Adenovirus 7	Virus internalization			(289)

Virus	Roles of Ca ²⁺	Viral CaBPs (proposed Ca ²⁺ -binding site)	Methods (K _d , μM)	Reference
BKV	Capsid stability	VP1 (Site 1: E331; Site 2: D346)	Cryo-EM	(290)
MPyV	Capsid assembly	VP1 (aa 266-277)	⁴⁵ Ca binding	(31)
SV40	Assembly, cell entry and nuclear entry	VP1 (Site 1: E46/E48/S213/E216/E330; Site 2: E157/E160/K214/E216/D345)	X-ray	(34,291,292)
BFDV	Capsid assembly	VP1 (aa 237-248)	⁴⁵ Ca binding	(293)
<i>dsDNA viruses with RNA intermediates</i>				
HBV	Core assembly and virus replication		TEM	(40)
<i>Others</i>				
Bacteriophage φX174	Trigger conformational change; virus assembly	F protein (Site 1: G1321/D1421/M1424/S1426/Q1004 Site 2: S1001/N1002); G spike protein (Site 3: D2117)	X-ray	(294,295)
Bacteriophage G4	Virus assembly	F protein [PDB code: 1gff] (A1321/N1322/D1421/M1424/S1426)	X-ray	(296)
Bacillus subtilis phage 41c	Adsorption; penetration; Intracellular phage development			(297)
fd and M13 virus	Lateral aggregation and bundle formation		EM	(298,299)

Abbreviations: K_d , dissociation constant; HAV, hepatitis A virus; STNV, Satellite tobacco necrosis virus; NSP, nonstructural protein; TBSV, tomato bushy stunt virus; EM, electron microscope; TCV, turnip crinkle carmovirus; SeMV, Sesbania mosaic virus; CMV, cocksfoot mottle virus; RYMV, rice yellow mottle virus; TMV, tobacco mosaic virus; RCNMV, Red clover necrotic mosaic virus; PM, plasma membrane; GP, envelope glycoprotein; CD, circular dichroism spectroscopy; HN: Hemagglutinin-neuraminidase; N protein, nucleocapsid protein; PDB, protein data bank; RSV, respiratory syncytial virus; HIV, human immunodeficiency virus; IR, infrared spectroscopy; BLV, bovine leukemia virus; IBDV, infectious bursal disease virus; CPV, canine parvovirus; FPV, feline panleukopenia virus; MPyV, Murine polyomavirus; SV40, simian virus 40; BFDV, budgerigar fledgling disease virus; HBV, hepatitis B virus; TEM, transmission electron microscope.

The virion of cocksfoot mottle virus (CfMV), which belongs to the genus *Sobemovirus* with a single-stranded, positive-sense RNA genome, has an icosahedral capsid composed of 180 copies of coat protein monomers assembled in $T=3$ quasi-equivalent symmetry (255). These monomers, all of which contain 253 amino acids, have a jelly-roll β -sandwich topology and exhibit three slightly different conformations, denoted as quasi-equivalent positions A, B and C (Fig. 4.2A). The subunits A, B and C are assembled as asymmetric units that eventually form the icosahedral capsid. Three Ca^{2+} ions, each incorporated between the interacting surfaces of subunits (A-B, B-C and A-C), function as reusable “glue” to stick adjacent subunits together and stabilize the capsid. Each Ca^{2+} ion is coordinated by O δ 1 of residues Asp¹³⁶ and Asp¹³⁹ from one subunit and the main chain carbonyl oxygen of Leu¹⁹⁶, O δ 1 of Asn²⁵², and the C-terminal carboxyl oxygen from the other interacting subunit (Fig. 4.2A). Such Ca^{2+} -binding sites seem to adopt an octahedral geometry with only five ligands and an average Ca-O distance of 2.4 Å (142,254,255).

The incorporation of Ca^{2+} ions into the capsid is also observed in the animal viruses with either RNA or DNA genomes. In human rhinovirus 1A (HRV1A), HRV3 and HRV14, Ca^{2+} ions are located at the five-fold axes of the icosahedral capsids (Fig. 4.2B). The metal ions are coordinated with the oxygen atoms from the main chain carbonyl group of Asn¹¹⁴¹ and S¹¹⁴¹ in five copies of the symmetry-related VP1 of HRV14 (241) and HRV3 (242), respectively. In HRV1A, the Ca^{2+} coordinates with the O δ 1 of Asp¹¹³⁷ of fivefold-related VP1 (300). With two additional oxygen atoms from water molecules above or below the metal ion

as coordinating ligands, the Ca^{2+} -binding pocket forms a pentagonal bipyramidal geometry with Ca-O distance ranging from 2.3 Å to 3.1 Å. In HRV14, another Ca^{2+} ion with similar property coordinates with the Oε1 of Glu³²⁰⁰ and situates itself on the threefold axis of the capsid. Similar to HRV14, each asymmetric unit of the icosahedral capsid of Coxsachievirus B3, which belongs to the genus enterovirus in the *picornaviridae* family, contains two Ca^{2+} ions, with one interacts with Oδ1 of Asp³²⁰³ from VP3 on the threefold axes and the other involves water-mediated interaction with the fivefold-related residues Asp¹¹³³ on VP1 (244). Three Ca^{2+} ions are incorporated in the ssDNA virus bacteriophage φ174. Site 1 is located approximately 10 Å from the threefold axis and is coordinated with six ligands from the main chain of Glu¹³²¹, Asp¹⁴²¹, Met¹⁴²⁴, S¹⁴²⁶ and the side chain of Gln¹⁰⁰⁴ and Asp¹⁴²¹ in the F protein. Site 2 in the F protein and Site 3 in the G spike protein situate on the threefold and fivefold axes, respectively. The Ca^{2+} ion at Site 1 is functionally important for the injection of viral DNA into the host cells, whereas the binding of Ca^{2+} ions to Site 2 and Site 3 on the axes stabilize the capsid (294,295).

The ssDNA parvoviruses, feline panleukopenia virus (FPV) and canine parvovirus (CPV), bind three and two Ca^{2+} ions, respectively (287). CPV was discovered in 1978 as a host range variant of FPV. Both FPV and CPV contain a pH-dependent dual Ca^{2+} -binding site that is separated by 4.6 Å. The largest structural difference between these two viruses is confined to a flexible surface loop (aa 359-375) in the capsid protein VP2 that could possibly interact with host cellular surface molecules. This loop forms a third Ca^{2+} -binding site in FPV, but

not in CPV. The binding of Ca^{2+} to this particular site causes conformational changes in VP2 and seems to influence the host range of different parvoviruses (287).

The binding of Ca^{2+} to the capsid of dsDNA viruses is best exemplified in simian virus 40, a small DNA virus that belongs to the *polyomaviridae* family. Mutations at the Ca^{2+} binding ligands lead to defects in capsid assembly at nonpermissive temperature (301). Even after the formation of virus-like particles (VLPs), the Ca^{2+} mutants are much less infectious than the wild type virus. More interestingly, some of the mutants fail to enter the cell (e.g., E330K in Site 1) or enter the nucleus (e.g., E157A-E160A and E216K in Site 2) due to its incapability to bind Ca^{2+} and subsequent premature VLP dissociation. Thus, the binding of two Ca^{2+} ions to the dual Ca^{2+} binding site on VP1 are essential for the assembly of the icosahedral capsid, as well as the cellular and nuclear entry (34,292).

NSP4 of rotavirus

The NSP4 of rotavirus, a transmembrane glycoprotein primarily embedded in the membrane of endoplasmic reticulum (ER) of rotavirus-infected cells, is required for the budding of immature viral particle into the ER lumen and thus plays a central role in the morphogenesis of rotaviruses (229). The NSP4 also functions as an enterotoxin and triggers the ER Ca^{2+} release, eventually resulting in gastrointestinal symptoms (36,231). The NSP4 itself has been revealed to contain a core Ca^{2+} -binding site when it oligomerizes into a functional homotetramer (286). The Ca^{2+} ion is coordinated by the side chains of four Gln¹²³ and

two Glu¹²⁰ residues from four neighboring identical polypeptides (Fig. 4.2C). This Ca²⁺-binding site appears to stabilize the tetrameric state of NSP4 in the extracellular space and the cytoplasm, where the Ca²⁺ concentration could be increased by 5 folds after rotavirus infection (206).

Neuraminidase of Influenza viruses

The neuraminidase of influenza viruses catalyses the cleavage of the glycosidic linkages between the terminal silica acid residues and the adjacent carbohydrate moieties on the surface of infected cells. This activity is required for the release of virus from the cell surface. Two distinct Ca²⁺-binding sites are found in the neuraminidase (237,264). One high-affinity site (Fig. 4.2D) is close to the active site and adopts an octahedral geometry with six coordinating ligands from the carboxyl group of the side chains of Asp³²³, the carbonyl groups of the main chains of Gly³⁴³, Gly³⁴⁵, Thr²⁹⁶, Asp²⁹² and one molecule of water. This Ca²⁺-binding site is conserved among influenza A virus (262,264), influenza B viruses (237) and the parainfluenza virus (268). The other relative low-affinity site is located on the fourfold axis of the tetrameric neuraminidase. The high-affinity Ca²⁺-binding site is needed for the thermostability of the enzyme and optimal enzymatic activity; whereas the low-affinity site has been postulated to hold the tetramer together (46).

Continuous viral Ca²⁺-binding sites

Continuous Ca²⁺ binding sites are best characterized by the 29-residue the EF-hand motifs, where the Ca²⁺ ions are coordinated by seven oxygen atoms from the sidechain carboxyl or hydroxyl groups (loop sequence positions 1, 3, 5, 12), the main chain carbonyls (position 7), and a bridged water (position 9) by adopting a pentagonal bipyramidal geometry (302). The helix-loop-helix EF-hand moiety is one of the most widely used motifs in proteins and is ubiquitously distributed across bacteria, archaea, and eukaryotes (6,138). Upon binding to Ca²⁺, this motif may undergo conformational changes that enable Ca²⁺-regulated functions, as seen in Ca²⁺ effectors such as calmodulin (CaM) and troponin C (TnC) or Ca²⁺ buffers such as calreticulin and calbindin D9k. While the majority of the known EF-hand CaBPs contain paired EF-hand motifs, single nuclear EF hands have also been discovered in bacteria and eukaryotes (5,6). The recently reported ER Ca²⁺ sensor, stromal interaction molecule 1, has been confirmed to function as a monomer in the presence of Ca²⁺ (303).

Though EF-hands have been found abundantly in eukaryotes and bacteria, literatures on EF-hand or EF-hand like Ca²⁺-binding motifs in virus proteins are, nevertheless, scarce, possibly due to lack of accurate prediction methods and robust validating methodologies. A thorough search in PubMed with the key words “EF-hand and virus” only results in 4 examples of viral EF-hand or EF-hand like motifs: the NSP protease domain (DASPDGTGDPLD) of rubella virus (250), the VP1 (DENGVGPLCKGE) of polyomavirus (31), the VP7 outer capsid protein (DITADPTTAPQTE) of rotavirus (39), and the transmembrane

protein gp 41 of HIV-1 (44). The binding of Ca^{2+} to these sequences either enhances the protein stability or promotes the enzyme activity.

Our laboratory has developed a pattern search method to predict continuous Ca^{2+} -binding sites from genomic information (<http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm>). This bioinformatic tool has been successfully applied to predict and analyze potential bacterial EF-hand and EF-hand like Ca^{2+} -binding motifs at genome-wide level (Chapter 3) (6). Given the diversity of viral genomes and its close association with host cells that are abundant with the EF hand motif, it would be surprising to find only 4 cases. Driven by that, we initiated a comprehensive search for potential viral EF-hand motifs by screening all viral genomic information that is available on the protein database Swiss-Prot/TrEMBL. With our developed method, along with the pattern PS00018 (<http://ca.expasy.org/prosite/PDOC00018>) from Expert Protein Analysis System (ExPASy) proteomic server, we have detected a number of additional potential EF-hand motifs (summarized in Table 4.3), though the Ca^{2+} -binding capability of these sequences remain to be experimentally verified. The 93 putative EF-hand or EF-hand-like motifs are found in the genomes of almost 80 different viruses, spreading throughout the majority of virus families. Almost all of these matches are found to be single EF-hand motifs except for two EF-hand or EF-hand-like motifs detected simultaneously in the envelope protein of HIV-1 and the immediate-early protein RSP40 of pseudorabies virus. These putative viral EF-hand-containing proteins are involved in a wide range of viral or cellular events, such as viral adsorption and fusion (neuraminidase of Influenza A virus,

Sendai virus and human parainfluenza virus 1; envelope glycoprotein of HIV-1; spike protein of rat coronavirus, murine hepatitis virus and bovine ephemeral fever virus; glycoprotein B of feline herpesvirus 1), virion assembly and disassembly (coat protein of beet yellow stunt virus, papaya ringspot virus and African horse sickness virus), viral precursor protein processing (nonstructural protease of rubella virus), viral nucleic acid modification and replication (mRNA-capping enzyme of alphavirus; RNA-directed RNA polymerase of tobamovirus, respiratory syncytial virus, and influenza A virus; DNA methylase of *Sulfolobus* virus; DNA polymerase of nucleopolyherpesvirus and human herpesvirus 2) and transcriptional regulation of viral genes (ICP0 of bovine herpesvirus 1; IE63 of human herpesvirus 3; ICP4 of equine herpesvirus 1). In addition, the functions of almost 20% of these matched proteins remain uncharacterized. We hope that our prediction would serve as a prelude to more extensive searching for additional viral Ca^{2+} -binding proteins that are closely associated with virus-host interacting events.

Table 4.3. Putative EF-hand and EF-hand-like Ca²⁺-binding motifs predicted in virus genomes.

Virus	Accession ID	Protein name	Match Ranges	Sequences
EF-hand motifs				1 2 3 4 5 6 7 8 9 10 11 12
(+) ssRNA viruses				
MHV	Q5ICX2/Q83331	Spike protein	727-738	D N S T E Q S V D A C D
Rat coronavirus	Q9IKD1	Spike protein	726-737	D N S T E Q S V D A C D
Rubella virus	P13889	NSP (Protease)	1206-1217	D A S P D G T G D P L D
VEEV	P36327	NSP (mRNA-capping enzyme nsP1)	401-412	D Q E D E R P L G L R D
WEEV	P13896	NSP (mRNA-capping enzyme nsP1)	399-410	D L D D E K E L G V R E
Sindbis virus	P27283	NSP (mRNA-capping enzyme nsP1)	402-413	D L D N E K M L G T R E
SFV	P08411	NSP (mRNA-capping enzyme nsP1)	402-413	D L D D E K P L G V R E
RRV	P13887	NSP (mRNA-capping enzyme nsP1)	402-413	D L D N E K M L G T R E
SqMV	P36341	Movement protein	218-229	N N S G D N E V E F S E
SBLV	Q8JW06	3a protein	52-63	N L S S D N R L N F I D
Nora virus	Q27YG9	Replication polyprotein	462-473	N K S P D K S V T I E D
Tobamovirus (RMV/TMV/TVCV/YoMV)	Q1L1D7/Q88604/Q88920/Q66220	RNA polymerase	1358-1369/ 1362-1373	D L D S T Q A M E I L E
TTV	A1XIP9/A6NA62	Polyprotein/ RNA polymerase	1535-1546/ 196-207	D V S S S K M L D L S E
BSBMV	Q9IF43	29K protein	222-233	D D D G D G V V G D D D
BYSV	Q65858	Coat protein	11-22	D S S A S Q T M T A K D
PRV	Q98WK3	Coat protein	50-61	D K D N D G A C D G N D
SMoV	Q8UYV9	Polyprotein	1388-1399	D K D G D R W V A K D E
(-) ssRNA viruses				

Virus	Accession ID	Protein name	Match Ranges	Sequences
HRSV-A	P28887/O36635/ Q82021/Q82027	RNA Polymerase (subunit L)	1557-1568	D M N T S D L L C V L E
HRSV-B	P24567/O42062	Phosphoprotein	229-240	D N D S D N D L S L D D
BRSV	P33454/Q77KZ1	Phosphoprotein	229-240	D E S S D N D L S L E D
PPRV	Q4PIR9/Q91QS4	Phosphoprotein	74-85	D M S P E D N L G F R E
BEFV	P32595/Q56118	Spike protein	400-411	D K N E D G Y I D I Q E
Puumala virus	Q6QTA9	RNA polymerase	1975-1986	D S D E D D D V S Q L D
NDV	A0SZV5	RNA polymerase	1529-1540	N H D G S H Q L A D T D
Influenza A virus	Q110S8	Neuraminidase	68-79	N I S N T K I V N V Q D
	Q0P290	Neuraminidase	68-79	N I N N T N V V A G K D
(+) ssRNA viruses require DNA intermediates				
HIV-1	Q2PNG6	Envelope protein	29-41	D N S T E G T V G G G E
	Q2PN97		29-41	N T S T G G T V G G G E
	Q2PN98		29-41	N N S T G G T V G G G E
	A1Z0A9		132-143	D L N T T N T I N S S D
	Q6EJM9		28-39	N T N N S R I M E G G E
	Q8AF19		139-150	N S S S E N K M E I G E
	P19549/Q03807		141-152	N N S S G G T V E K E E
	Q1G4R9		151-162	N N S S G Q I M E K G E
	Q1HSZ5		137-148	N S S S G K M M E E G E
	Q2NN71		137-148	N S S S E R T M E K G E
	Q1HSY3		143-154	N S S R G K M M E K G E
	Q1HT34		155-166	N N S G G G T V E E R E
	Q1HT21		149-160	N N S S G R T M E E R E
	Q1HT20		147-158	N N S S G R T I E E R E
	Q2PN97		147-158	N T S S T G G T V G G E
	Q6QBL0		32-43	N N N S E G R I E R G E
	Q6QUE1		34-45	N T S S E G M V E K G E
	Q71129		24-35	N N S S G K L I E L G E
	Q71313		17-28	N S S D E G K I E K G E

Virus	Accession ID	Protein name	Match Ranges	Sequences
CYMV	Q91B26	Putative protein	129-140	N R D H E Q L C E V V E
<i>dsRNA viruses</i>				
PsV-F	Q4G3H1	Capsid protein	294-305	D E D R S D L L T T L D
AHSV-6	Q64913	VP6 protein	144-155	N D D A T R N V G S S E
<i>ssDNA viruses</i>				
FPV	P04864/P24840	Coat protein VP1	658-669	S A N M S R I V T Y S D
MEV	P27437	Coat protein VP1	653-664	S A N M S R I V T Y S D
<i>dsDNA viruses</i>				
Megalocytivirus (RBIV/DGIV/ISKNV/OSGIV/R SIV)	Q8B4N1/Q65858/ Q8QUT8/Q1X6P8/ Q71G61	Putative Phosphatase	80-91	D M S Q D G F V N Y Q D
Ranavirus (ATSV/FV-3)	Q6YH54/Q6GZT5	Orf2-like protein	327-338	D P D A S H V M Q T D E
LCDV	Q677V5	Purative protein	31-42	D I N N S Q I I T V S D
EHV-1	P28925	ICP4	733-744	D S D P T H R L G S D E
FeHV-1	Q86665/Q90050	Glycoprotein B	871- 882/ 876-887	D F D E E K L M Q A R E
WSSV	Q8JNB7/ Q8JNC2	ORF2735/ ORF4162	312- 323/ 281-292	D D D D D D D C E G M D
CPV	Q80DU9/Q8QMS8	A27L	383- 394/ 57-68	D D D D D D V I D D D D
Squirrelpox virus	Q1HTR0	C3R	22-33	D D D T D G E L E S K D
Ectromelia virus	Q9PXN5	C11R	32-43	N R N S T H K I Q E N E
MOCV-1	Q98173	MC002L	251-262	N T N E N G A M S S A E
Nucleopolyhedrovirus (AgMNPV/Cfdef/CfNMPV)	Q06KP7/Q6VTX1/ Q7TLX0	Putative protein	63- 74/61- 72/ 61-72	D D N N G R C C N V V D / D D S A N R C C S V V D

Virus	Accession ID	Protein name	Match Ranges	Sequences
Granulovirus		Alk-exo	350-361	N L N E S Q T L S L S E
PBCV-1	Q84528	A208R	214-225	D D D S D D D I N D N D
ESV	Q9J3U5	EsV-1-56	333-344	D V S G N D L L N V S D
Sulfolobus virus	Q684G2	Adenine-specific DNA methylase	151-162	D D D A E G Y I G T W E
Archaeal BJ1 virus	A0ZYT1	Putative protein	18-29	D I S D T N V L A A A E
AFV-2	Q573C5	Putative protein	85-96	D S S S S Q E V T I P D
HZV-1	Q8JKJ1	Microtubule-associated-like protein	423-434	D T D G S N D L A K L D
Others				
ABV	A4ZUC5	ORF188	169-180	D P N Q N Q T I S E S E
EF-hand-like motifs				
(+) ssRNA viruses				
EEEV	Q306W6	NSP	2401-2414	D D D Q D G D R R R A L Y D
CSFV	P19712	Genome polyprotein	929-942	D C N R D G V V I S T E G E
(-) ssRNA viruses				
Influenza A virus	Q6XV43	Neuraminidase	385-398	D K D S N G V Q D I I D N D
	P21427/Q0A2K9/O89750/Q0A2Q7	RNA polymerase subunit P2	294-306	D P S H E G E G I P L Y D
HPIV-1	P16017	Hemagglutinin-neuraminidase	281-293	D Y S S E G I E D L V F D
Sendai virus	Q783Y1/P19758/P03425/Q88261	Hemagglutinin-neuraminidase	281-293	D Y S S D G I E D L V L D
VHSV				
dsRNA viruses				
AHSV-4	Q64929	VP4 core protein	263-276	D V S A D G L K G T I E W E

Virus	Accession ID	Protein name	Match Ranges	Sequences
(+) ssRNA viruses require DNA intermediates				
HIV-2	P24103	Negative factor Nef	181-194	D T S Q E G E D T E T D T E
MPMV	P07567	Gag polyprotein	133-146	D S D D E G A K S S S L Q D
dsDNA viruses				
BoHV-1	P29128	Trans-acting transcriptional protein ICP0	287-300	D S D S E G S E D D S W S E
HHV-1	P10211/P06347	Glycoprotein B precursor	824-837	D A S G E G E E G G D F D E
HHV-2	P07918	DNA polymerase	675-688	D D D E D G D E D G D E R E
HHV-3	P09255	Transcriptional transactivator IE63	164-177	D S D D D G S T P S D V I E
HHV-6	Q01350	Protein U3	284-297	D N D E D G R P R F V A E
HHV-7	P52520	Protein U3	299-311	D N D P D G N L T F I A E
CeHV-9	Q04548	Glycoprotein E precursor	40-53	D M D E D G V Y G D D I Q D
PRV	P24827	Immediate-early protein RSP40	224-237	D E D E E G E E E E D E E E
	P24381	Serine/threonine-protein kinase	335-348 77-90	D E D E D G E G E E G G K D D G D G D G D S S G D E D D
EBV	P12977	Nuclear antigen 3	711-724	D E S G E G S D T S E P C E
	P03210	Putative protein BRRF2	492-505	D E D E D G S E D G E F S D
BmNPV	P41712	DNA polymerase	955-967	D D D D D G C D S S D S E
SINPV	O57030	DNA-directed RNA polymerase beta chain	422-435	D E D E N G S G G D D D D D

Abbreviations: MHV, murine hepatitis virus; VEEV, Venezuelan equine encephalitis virus; WEEV, western equine encephalitis virus; SFV, Semliki forest virus; RRV, Ross river virus; SqMV, squash mosaic virus; SBLV, spring beauty latent virus; RMV, ribgrass mosaic virus; TMV, tobacco mosaic virus; TVCV, turnip vein-clearing virus; YoMV, youcai mosaic virus; TTV, tomato torrado virus; BSBMV, beet soil-borne mosaic virus; BYSV, beet yellow stunt virus; PYV, papaya ringspot virus; SMoV, strawberry mottle virus; HRSV, human respiratory syncytial virus; BRSV, bovine respiratory syncytial virus; PPRV, peste-des-petits-ruminants virus; BEFV, bovine ephemeral fever virus; NDV, Newcastle disease virus; HIV, human immunodeficiency virus; BSV, banana streak virus; DBV, dioscorea bacilliform virus; CYMV, citrus yellow mosaic virus; GVBV, gooseberry vein banding virus; KTSV, kalanchoe top-spotting virus; PsV-F, penicillium stoloniferum virus; AHSV, African horse sickness virus; FPV, feline panleukopenia virus; MEV, mink enteritis virus; RBIV, rock bream iridovirus; DGIV, dwarf gourami iridovirus; ISKNV, infectious spleen and kidney necrosis virus; OSGIV, orange-spotted grouper iridovirus; RSIV, red sea bream iridovirus; ATSV, ambystoma tigrinum stebbensi virus; FV, frog virus; LCDV, lymphocystis disease virus; FeHV, feline herpesvirus; EHV, equine herpesvirus; WSSV, white spot syndrome virus; CPV, cowpox virus; MOCV, molluscum contagiosum virus; AgMNPV, anticarsia gemmatilis nuclear polyhedrosis virus; Cfdef, choristoneura fumiferana defective polyhedrosis virus; CfNMPV, choristoneura fumiferana nuclear polyhedrosis virus; PBCV, paramecium bursaria chlorella virus; ESV, ectocarpus siliculosus virus; AFV, acidianus filamentous virus; HZV, heliothis zea virus; ABV, acidianus bottle-shaped virus; EEEV, Eastern equine encephalomyelitis virus; CSFV, classical swine fever virus; HPIV, human parainfluenza 1 virus; VHSV, viral hemorrhagic septicemia virus; MPMV, Mason-Pfizer monkey virus; BoHV, bovine herpesvirus; HHV, human herpesvirus; CeHV, Cercopithecine herpesvirus; PRV, pseudorabies virus; EBV, Epstein-Barr virus ; BmNPV, Bombyx mori nuclear polyherosis virus; SINPV, spodoptera littoralis nuclear polyherosis virus.

4.3. Ca^{2+} dependent virus-host interactions

Compared to the scarcity of reported viral Ca^{2+} -binding proteins, the host cells are abundant with cellular Ca^{2+} -binding proteins. Accordingly, viral proteins might utilize a number of important cellular Ca^{2+} -binding proteins as alternative strategies in its response to the versatile and universal Ca^{2+} signals (listed in Table 4.4). These proteins are widely distributed in the extracellular matrix (fibulin-1), cytoplasm (annexin, calmodulin, and S100) and endoplasmic reticulum (ERC-55, calreticulin and calnexin).

Rubella virus. Rubella virus (RUB), a positive-sense single-stranded RNA virus of the *Togaviridae* family, is the causative agent of German measles. The RUB genome consists of two open reading frames: one for nonstructural proteins and the other for structural proteins (capsid, glycoprotein E1 and E2). Both glycoproteins E1 and E2 have been demonstrated to interact with two Ca^{2+} -binding ER chaperones, e.g., calreticulin and calnexin (304). The 46-KDa calreticulin consists of three domains, a globular N domain, a P domain rich in proline residues, and an acidic C-terminal tail domain. The N and P domains are responsible for the chaperone function and the highly acidic C-terminal tail is involved in buffering Ca^{2+} storage with high capacity (20-30 mol Ca^{2+} /mol protein) (305). Calreticulin mainly functions as a molecular chaperone to facilitate the folding of a number of proteins. Its activity is influenced by changes in the ER, such as the concentration of Ca^{2+} and ATP. Calnexin is a membrane-bound Ca^{2+} -binding ER chaperone that retains incorrectly or incompletely folded proteins (306). During posttranslational modification, virtually all glycoproteins associate

with these two chaperones. Both proteins can promote folding, oligomerization and suppress degradation. They work in concert as the “quality control” for the secretory proteins. By differentially interacting with E1 and E2, calreticulin and calnexin have been established to control the transport of these two glycoproteins. Since glycoproteins associate with calreticulin and calnexin with a prolonged duration (>60 min), the maturation of RUB glycoproteins has been proposed to be the rate limiting step of the transport of glycoproteins from ER to Golgi complex (304).

In addition, the 3' end of the RUB genomic RNA stem-loop (SL) structure, which is essential for the initiation of negative-strand RNA synthesis, has also been found to interact with the N domain of the autophosphorylated calreticulin (307). Several possibilities have been postulated to interpret the necessity of such interaction: 1) the interaction might promote viral RNA replication or translation, thus ensuring a productive infection of RUB, and 2) calreticulin-bound viral RNA might be compartmentalized to escape the surveillance of the host immune system. However, the exact biological consequence of the RNA SL-calreticulin interaction is not clear (308).

HIV-1. HIV-1 and several HIV-1 encoded gene products have been demonstrated to interact with cellular Ca^{2+} -binding proteins including annexin II, calmodulin, calreticulin and calnexin.

Annexin II, traditionally thought of as a Ca^{2+} -dependent phospholipid-binding protein expressed in monocytes, microglia and macrophages, is

implicated in membrane trafficking, endosomal formation and vesicles aggregation (309). It has been shown that the HIV polyprotein precursor p55^{GAG} interacts with annexin II in macrophages. However, the Gag-annexin II interaction is only seen in productively infected macrophages but not in quiescently infected cells. Such interaction localizes at the limiting membranes of late endosomes, the site where HIV assembly and budding takes place. Depletion of annexin II is further shown to be responsible for the destabilization of lipid rafts and subsequent abortion of HIV assembly. In addition, annexin II is reported to promote HIV entry into macrophage through its interaction with phosphatidylserine in the viral particles (82).

CaM is a small (148 amino acids; MW: 16.7 KDa) and acidic (pI: ~4.0) EF-hand Ca^{2+} -binding protein that is ubiquitously expressed in eukaryotic cells. It consists of two globular and autonomous domains, each of which contains two helix-loop-helix EF-hand motifs. Through its binding to Ca^{2+} and the concomitant conformational changes, CaM is capable of transducing the intracellular Ca^{2+} signal changes into divergent cellular events by targeting to an array of cellular proteins (14,310). Two HIV proteins, Nef and gp160 have been demonstrated to interact with calmodulin in a Ca^{2+} -dependent fashion. Nef is a myristoylated protein expressed in early infection of HIV. Nef has been shown to downregulate the cell surface receptors CD4 and MHCI, and alter T lymphocyte signaling pathways. The latter effect is partially associated with its ability to strongly interact with CaM with an apparent dissociation constant of 13.7 nM (311). The CaM-targeting sequence in Nef is further narrowed down to a 20-amino-acid N-

terminal basic domain that shares high similarity with the myristoyl group of a neuron-specific protein kinase C substrate, NAP-22 (311,312). Besides Nef, the viral envelope protein gp160/gp41 has been shown to interact with CaM (313,314). Two C-terminal CaM-targeting sequences in gp41 are shown to bind to CaM with dissociation constants of 31-41 nM (315). Such interaction is speculated to disrupt the anti-apoptotic CaM signaling pathway by either reducing the amount of free cytosolic CaM or changing its subcellular localization (315). The similar CaM-targeting sequence is also detected in simian immunodeficiency virus gp41 (316). Adding more complexity to the scenario, the Gag gene products of HIV have been shown to interact with CaM (317). The N terminus of p17 contains two contiguous CaM-binding sites, each of which binds CaM with affinities of $\sim 10^{-9}$ M. In view of the diverse roles of Ca^{2+} /CaM-dependent signaling pathways, the interaction between all these HIV proteins with CaM are expected to play multiple roles to fit the HIV life cycle in response to altered Ca^{2+} signals.

The newly synthesized HIV envelope glycoprotein gp160 has been shown to interact with both calreticulin and calnexin to form a transient ternary structure, and thereby facilitates the folding and maturation of HIV glycoproteins (318).

HTLV-1. The Tax protein of HTLV-1 is primarily located in the nucleus and functions as a transcriptional transactivator. Nevertheless, the Tax protein can be also exported to cytoplasm to interact with a number of host transcription factors including NF- κ B. This nuclear transport process possibly involves the interaction

of Tax with the Ca^{2+} -binding protein calreticulin (319). Interestingly, the calreticulin-mediated nuclear transport is dependent on Ca^{2+} (320). It has been reported that Ca^{2+} -loaded calreticulin does not support nuclear export of leucine-rich NES-containing proteins. However, calreticulin restores the nuclear export capability after treatment with EGTA. It still remains to be tested whether this property has effects on the subcellular distribution of Tax. In addition to Tax, the viral protein P12^l has been shown to interact with the ER-resident Ca^{2+} -binding chaperones, calreticulin and calnexin (321). This interaction is expected to facilitate the folding of p12^l and to modulate the level of Ca^{2+} storage. Another possible role of such interaction is to retain calreticulin-MHCI complexes in the ER or cis-Golgi, thereby blocking its association with β_2 -microglobulin and the trafficking of this protein complex (322).

HPV. The HPV E6 protein, a small polypeptide of approximately 150 amino acids, plays multiple roles in HPV infection. E6 has been found to possess oncogenic activity by stimulating immortalization of human keratinocytes and transforming established fibroblasts (323). By targeting p53, Bak and Myc for degradation, E6 protein is capable of inhibiting apoptosis of infected cells. Moreover, yeast two-hybrid screening studies reveal that the oncogenic HPV E6 protein interacts with two cellular Ca^{2+} -binding proteins, fibulin-1 and ERC-55.

Fibulin-1 is a Ca^{2+} -binding extracellular matrix protein that has been implicated in cellular motility modulation, cellular transformation and tumor genesis. The E6 protein has been demonstrated to interact with fibulin-1 in COS-

7 cells transiently transfected with E6 from cancer-related HPV 16 or the transforming bovine papillomavirus type 1 (BPV-1) (324). Interestingly, overexpression of fibulin-1 inhibits E6-mediated cellular transformation, suggesting that the fibulin-1 serves as a tumor suppressor. The two-hybrid assay further localizes the interacting region in fibulin-1 to amino acids 343-483, which covers an important region (amino acids 356 to 440) that is involved in the self-association, Ca^{2+} -binding and fibronectin-binding events. Given that, it has been postulated that the interaction between the oncogenic E6 protein and fibulin-1 disrupts these functions and further inactivate the inhibitory effect of fibulin-1 on cell migration and invasion.

The cancer-related HPV and BPV-1 E6 proteins also bind a hexa-EF-hand Ca^{2+} -binding protein, ERC-55, that resides in the endoplasmic reticulum (325). Biochemical and biophysical studies further map the interacting region in ERC-55 to a stretch of 25 amino acids in the fourth EF-hand motif that is capable of binding Ca^{2+} and folds into the typical helix-loop-helix conformation (326). The exiting helix of the EF-hand motif (residues L²¹, F²⁴, L²⁵ and D²⁷), but not the Ca^{2+} -binding loop and the entering helix, is further found to be fully responsible for this interaction. The biological role of this interaction still remains uncertain.

HBV. The HBV Pol protein plays multiple roles in virus replication by signaling the encapsidation and degradation of pgRNA, priming reverse transcription and catalyzing both RNA- and DNA-dependent DNA synthesis (327). Besides its association with host cellular factors such as Hsp60, Hsp70, Hsp90

and p23, Pol interacts with a Ca^{2+} -modulated protein, p11 or S100A10 (328). The p11 protein is capable of recruiting HBV Pol to PML nuclear bodies, a region that is of vital importance for cell proliferation, differentiation and antiviral responses. This interaction is affected by intracellular Ca^{2+} concentration. Treatment of cells with valinomycin, a drug that promotes Ca^{2+} influx, leads to a decrease of these proteins in the nuclear bodies; whereas the blocking of Ca^{2+} influx by EGTA results in a marked increase of Pol-p11 complex in the nuclear bodies. Thus, it is expected that the virus would favor a higher cytosolic Ca^{2+} environment to avoid host antiviral activities. Indeed, as mentioned above, the HBx protein induces an increase in the cytosolic Ca^{2+} concentration by reducing mitochondrial Ca^{2+} uptake and impairs the activity of PMCA (223) .

The small hepatitis B surface antigen (HBsAg) has been demonstrated to specifically bind to human liver annexin V, a Ca^{2+} -dependent phospholipid binding protein present on the plasma membrane of human hepatocytes (329). The binding process, with a dissociation constant of 1.7 nM, has proven to be dependent on Ca^{2+} .

Other viruses. Tobacco mosaic virus (TMV) is a positive-sense, single-stranded RNA virus that infects plants, especially tobacco and other members of the plant family *Solanaceae*. The TMV cell-to-cell spread is mediated by the movement protein (MP), which is responsible for the transport of the viral genomic RNA through plasmodesmata. The Ca^{2+} -sequestering protein calreticulin has been implicated in this process through its interaction with MP

(330). Overexpression of calreticulin interferes with the normal targeting of MP and further delays the cell-to-cell viral movement.

Sendai virus is a negative-stranded RNA virus that belongs to the paramyxoviridae family. Two envelope glycoproteins, the fusion protein (F) and the hemagglutinin-neuraminidase (HN), mediate the virus entry into host cells. Both proteins undergo maturation by interacting with different chaperones during their transport through the ER and the Golgi complex (331). These two proteins have been reported to interact with ER chaperones with different kinetics due to difference in their oligomeric state and posttranslational processing. The F precursor protein, which is a type I integral membrane protein that can be cleaved into two active subunits linked by disulfide-bonds, only shows transient interaction with calnexin ($t_{1/2} = 8$ min). The HN protein, a type II integral membrane protein that functions as a tetramer, exhibits a slower kinetics by sequentially interacting with Bip, calnexin and calreticulin to ensure the folding and assembly quality (331).

Bluetongue virus, a double-stranded RNA arbovirus of the genus *Orbivirus* in the reoviridae family, causes catarrhal fever in ruminants that are of economical importance. The nonstructural protein NS3 has been shown to interact with a host cellular protein S100A10 or p11, which harbors an N-terminal pseudo EF-hand and a C-terminal canonical EF-hand Ca^{2+} -binding motif. Two p11 as the light chain components, along with two heavy chains composed of annexin II, form the tetrameric calpactin complex. The calpactin complex is closely associated with Ca^{2+} -dependent exocytosis and secretory pathways (332).

Since NS3 also interacts with outermost protein (VP2) of assembled virions, the BTV NS3 is proposed to bring the virions together with the calpactin complex via p11, and further directs the virus to the cellular exocytic machinery. Such interaction provides a new mechanism for the virus to make use of this cellular exocytosis pathway for nonlytic release of progeny virions assembled within the cell.

Direct interaction between annexin II and viral glycoprotein is also observed in CMV. The binding of annexin II to the CMV glycoprotein B protein (gpUL55) appears to be a Ca^{2+} -dependent process (333). Similar to its interaction with HIV, annexin II also enhances the binding and fusion of CMV to phospholipid membranes, thereby maximizing its infectivity (334).

4.4. Summary

We have summarized here the majority of the known mechanisms underlying virus- or viral proteins-mediated alterations in intracellular Ca^{2+} signaling events. The remodeled Ca^{2+} signaling network postinfection is closely associated with virus fusion and entry, viral gene replication and expression, virus maturation and release, and cellular transformation. Our predictions also result in a number of EF-hand-containing viral proteins that are worthy of further studies with respect to their Ca^{2+} binding capability and physiological relevance. The intensive association of viral proteins with cellular Ca^{2+} -binding proteins adds even more complexity to the Ca^{2+} -dependent virus-host interaction. A major

challenge in the future will be to integrate all aspects of the Ca^{2+} -virus interplay and capture the entire picture of viral calciomics.

Table 4.4. Interactions between cellular Ca^{2+} -binding proteins and viruses.

Cellular CaBP	Viral molecular identity	Virus	Consequences of interaction	Reference
Annexin II	p55 ^{GAG}	HIV-1	Facilitates virus entry and fusion in macrophages	(335)
	glycoprotein B	CMV	Enhances binding and fusion to membranes	(333,336)
Annexin V	Small HBsAg	HBV	Participates in initial steps of HBV infection	(336)
Calmodulin	Nef	HIV-1	Alters T lymphocyte signaling pathway	(311,312)
	gp 160/gp41	HIV-1	Disrupts CaM signaling pathway	(315)
	gp 41	SIV		(316,337)
	p17 ^{GAG}	HIV-1		(317)
Calreticulin/ calnexin	E1 and E2	RUB	Regulates viral glycoprotein maturation	(304)
	Viral RNA	RUB		(307)
	MP	TMV	Regulates cell-to-cell virus movement	(330)
	F, HN	SeV	Mediates maturation of glycoproteins	(331)
	gp160	HIV-1	Facilitates protein maturation	(318)
	P12 ^I	HTLV-1		(198)
	Tax	HTLV-1	Facilitates viral protein folding; possibly mediates the interaction with MHCI	(319)
ERC-55	E6	HPV		(325)
Fibulin-1	E6	HPV	Regulates cell migration and invasion	(324)
S100A10 (p11)	pol	HBV	Inhibits viral replication	(328)
	NS3	BTV	Mediates nonlytic virus release	(338)

Abbreviations: HIV, human immunodeficiency virus; CMV, cytomegalovirus; HBV, hepatitis B virus; SIV, simian immunodeficiency virus; RUB, rubella virus; SeV, Sendai virus; TMV, tobacco mosaic virus; HTLV, human T-cell lymphotropic virus; HPV, human papillomavirus; EBV, Epstein-Barr virus; BTV, bluetongue virus.

5. Probing the metal-binding properties of rubella virus nonstructural protease

5.1. Molecular biology of rubella virus (RUB)

Rubella virus (RUB), the only member of the genus rubivirus in the Togaviridae family, is the causative agent of a disease called as rubella or German measles. RUB is an enveloped, single-stranded positive-polarity RNA virus with a genome size of 9,762 nucleotides (Fig.5.1). The RUB genome contains two long open reading frames (ORFs): a 5'-proximal ORF or NSP-ORF (41 to 6,389 nt; 2,116 amino acids; MW: 240 kDa) that encodes the nonstructural protein (NSP) involved in viral RNA replication, and a 3'-proximal ORF or SP-ORF (6,512 to 9,700 nt; 1,063 aa; MW: 110 kDa) that encodes for one capsid protein (C) and two envelope glycoproteins (E1 and E2) (339,340). Multiple sequence alignments of the RUB NSP ORF against other positive-stranded RNA virus genomes revealed a number of conserved domains, namely, methyltransferase (M), protease (P), helicase (H), replicase (R), and a proline-rich region (G) as well as an X domain (X) showing high homology to the Appr-1-p processing enzyme (339,341). Upon uncoating and release of viral mRNA after entry into host cells, the NSP-ORF is immediately translated into a polypeptide precursor (P200), which is subsequently self-cleaved by its protease domain (termed the nonstructural or NS protease) at a cleavage site mapped between residues G¹³⁰¹ and G¹³⁰² into two products (P150 and P90) involved in viral

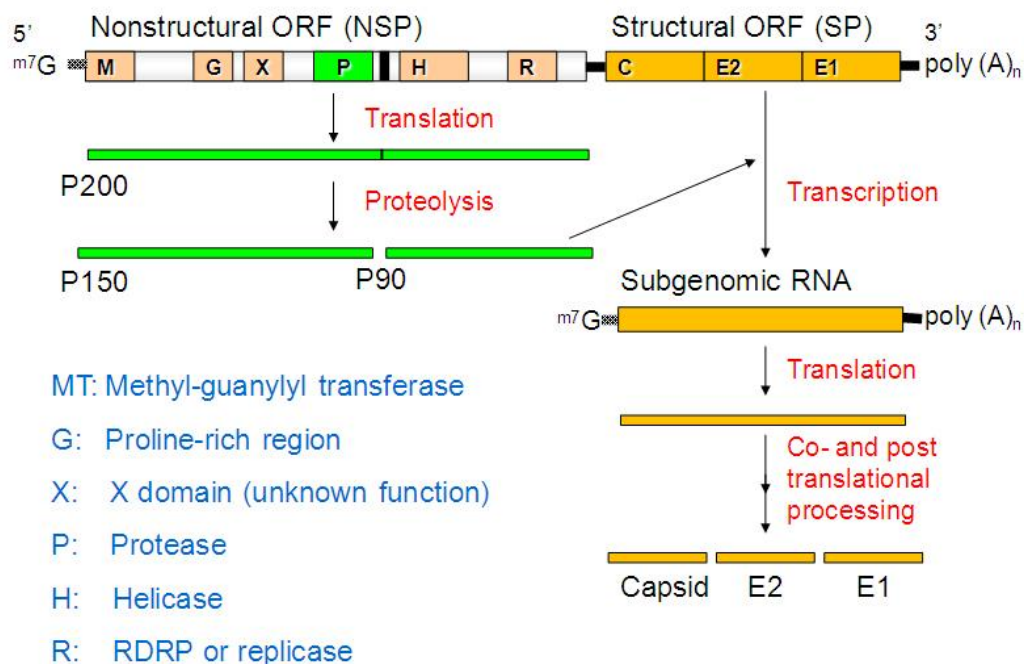


Figure 5.1. Genomic organization of RUB and viral protein processing. The RUB genome consists of two ORFs, one encoding nonstructural proteins (NSP) and the other encoding structural proteins (SP). Upon uncoating and release of viral mRNA after entry into host cells, the NSP-ORF is immediately translated into a polypeptide precursor (P200), which is subsequently self-cleaved by its protease domain (termed the nonstructural or NS protease) at a cleavage site mapped between residues G¹³⁰¹ and G¹³⁰² into two mature products (P150 and P90) involved in viral RNA replication. The SP-ORF polypeptide translated from the 24S subgenomic mRNA is eventually cleaved into three proteins, a capsid phosphoprotein and two envelope glycoproteins that form the spike complexes on the surface of the virion.

RNA replication in association with late endosomal or lysosomal membranes. The NS protease has been reported to be a Zn^{2+} -dependent papain-like cysteine protease with C¹¹⁵² and H¹²⁷³ constituting the catalytic dyad (342-346).

Metalloproteins constitute almost one third of all known proteins and viral metalloproteins are both common and the target of antiviral drug development (347-349). In the present study, we report our prediction of a single EF-hand Ca^{2+} -binding motif within the NS protease domain of RUB (section 5.2). To confirm this prediction, the binding loop was grafted into a scaffolding protein, CD2 domain 1, and the minimal metal binding domain of the NS protease was bacterially expressed and both were found to bind Ca^{2+} or its trivalent analogs Tb^{3+} and La^{3+} . We also found that mutagenesis of critical Ca^{2+} binding residues negatively impacted virus infectivity and rendered the activity of the NS protease temperature sensitive. This is the first demonstration of an EF-hand Ca^{2+} -binding motif in a virus-encoded protein. Moreover, with the minimal metal binding domain that also contains a putative Zn^{2+} -binding cysteine-rich motif, we have further carried out conformational and functional studies of this polypeptide and its cysteine mutants (section 5.3).

5.2. Identification of a Ca^{2+} -binding domain in RUB NS protease

5.2.1. Prediction of the EF-hand Ca^{2+} -binding motif and homology modeling of the RUB NS protease

By taking advantage of the sequence alignment of currently available EF-hand proteins and considering the structural context of the Ca^{2+} -binding loop, we

generated a series of patterns for the prediction of EF-hand proteins (6). Using this prediction method (<http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm>), we detected one putative EF-hand Ca^{2+} -binding motif (aa 1197-1225) within RUB NS protease, intertwined with the partially characterized cysteine-rich Zn^{2+} -binding motif (Fig. 5.2A) (346). This putative Ca^{2+} -binding motif shared high homology with well-known EF-hand proteins such as calmodulin and calcyphosine (Fig. 5.2B). In canonical EF-hand Ca^{2+} -binding motifs, the bidentate ligand residue at the loop position 12 is either a Glu (95%) or an Asp (5%). Considering its longer side chain and closer distance to the metal ion, Glu at this position (as seen in calmodulin) ensured a stronger binding affinity for Ca^{2+} than Asp. This predicted motif contained an Asp at loop position 12, similar to the first EF-hand in calcyphosine, regulatory light chain of myosin (350,351), and sarcoplasmic Ca^{2+} -binding protein (352) that exhibited relatively weaker Ca^{2+} -binding affinities. Prediction of the secondary structure elements revealed that the Ca^{2+} -binding loop was flanked by two helices, as observed in most EF-hand CaBPs (16). The predicted Ca^{2+} -binding motif is highly conserved among the eight genotypes of RUB for which sequence of this region is available, including all of the Ca^{2+} -binding coordination ligands (Fig. 5.2B).

Since the three dimensional structure of a viral papain-like cysteine protease, the leader protease of foot-and-mouth-disease virus (FMDV), has been determined (353,354), it provides an excellent template for homology modeling of the NS protease of RUB. Pairwise sequence alignment of the leader protease of

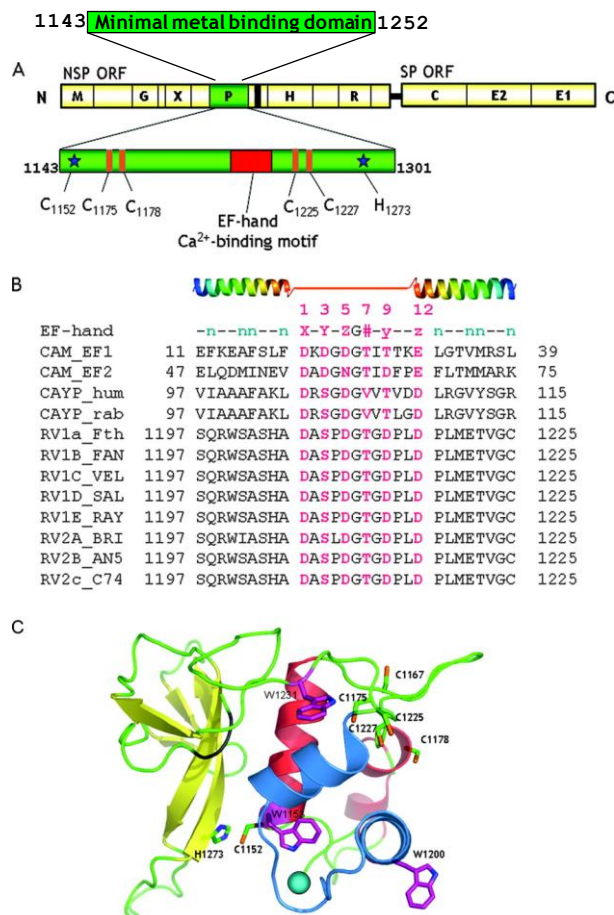


Figure 5.2. A putative EF-hand Ca^{2+} -binding motif in RUB NS protease. (A) Schematic representation of the RUB genome organization and domain locations. The protease contains at least one Zn^{2+} -binding site as well as a putative EF-hand Ca^{2+} -binding motif. It has been proposed that the residues C1152 and H1273 constitute the catalytic dyad (stars). The minimal metal-binding domain RUBCa (aa 1143 to 1252) is bacterially expressed and used in this study. (B) Sequence alignment results of putative EF-hand Ca^{2+} -binding motif RUBCa in the protease domain with calmodulin and calcyphosine. The motif remained conserved in both clades I and II of RUB. Boldface residues represent the potential Ca^{2+} -coordinating residues. NCBI or GenBank accession numbers are: CaM_EF1 and CaM_EF2 (calmodulin, P62158); CAYP_hum and CAYP_rab (calcyphosine from human, Q13938; rabbit, P41150). (C) Homology modeling of RUB NS protease. The leader protease of FMDV (pdb code 1qmy, chain A), a papain-like cysteine protease with a high-resolution structure available, is chosen as the template for homology modeling. The active site consists of C1152 and H1273 (shown as sticks). The predicted EF-hand Ca^{2+} -binding motif, located on the opposite side, is highlighted in blue with the calcium ion shown as a cyan sphere. The cysteine (green) and tryptophan (magenta) residues in the protease domain are shown as sticks.

FMDV (pdb code: 1qmy) with the RUB NS protease revealed 12.1% sequence identity and highly similar secondary structure arrangement. Based on the sequence alignment results, a homology model of the RUB NS protease has been generated using the program SWISS-MODEL (Fig. 5.2C). With the program GG (85), we further detected a potential Ca^{2+} -binding pocket around the predicted EF-hand Ca^{2+} -binding motif in this model structure, indicating the perfect agreement of the prediction made on the basis of both primary sequence and the predicted 3D structure of RUB NS protease.

5.2.2. Probing the metal binding properties of RUB NS protease by grafting

To dissect the Ca^{2+} binding capability of the putative Ca^{2+} -binding motif and meanwhile, to circumvent the difficulty encountered in purifying the cysteine-rich protease domain and the interference from Zn^{2+} binding, we grafted the predicted 12-residue Ca^{2+} -binding loop (aa 1206-1217) into a non- Ca^{2+} -binding scaffold protein CD2.D1. As seen in Fig.5.3A, CD2 is a cellular adhesion molecule composed of nine β -strands with Ig-like fold. We inserted the Ca^{2+} -binding loop from the RUB NS protease into CD2.D1 (named as CD2.RUBCa) between S⁵² and G⁵³ within the loop of strands C' and D. Three glycines at each side of the loop serve as linkers to provide sufficient conformational freedom for the grafted Ca^{2+} -binding loop and to minimize perturbation on the host protein while allowing the inserted loop to retain its capability of chelating Ca^{2+} (60,68-70).

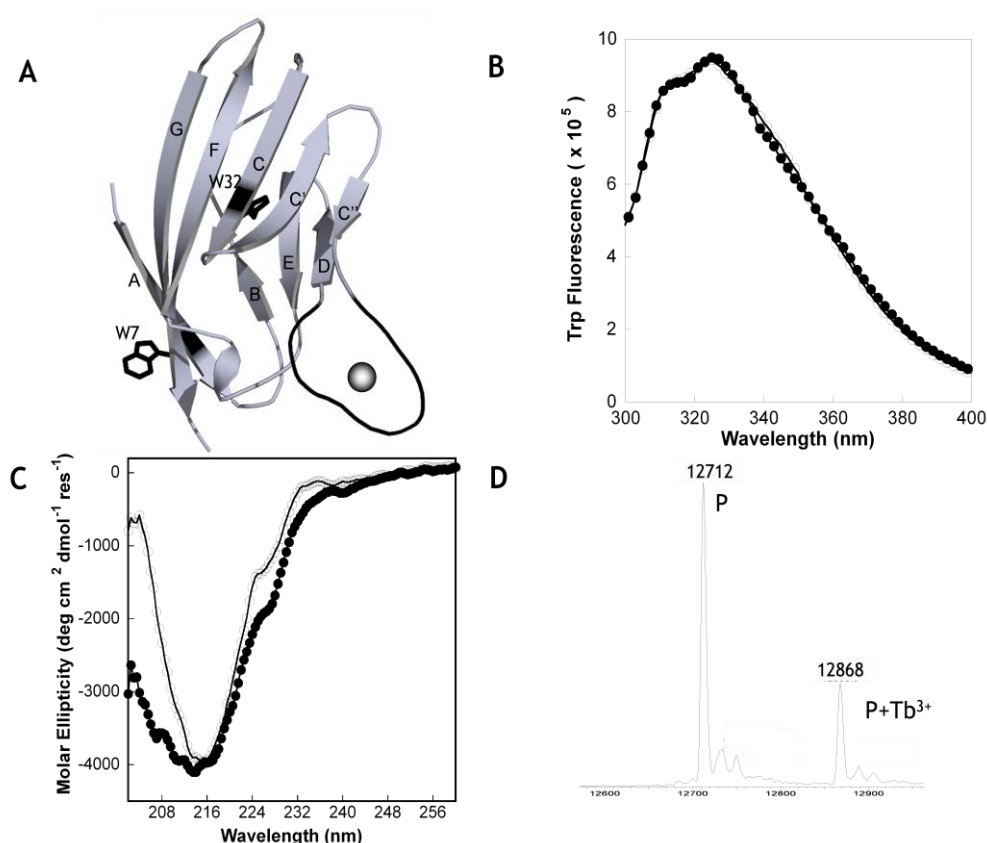


Figure 5.3. Grafting the predicted EF-hand Ca²⁺-binding loop into CD2.D1 and formation of metal-protein complex. (A), Model structure of the engineered protein CD2.RUBCa. The Ca²⁺-binding loop (black) from the RUB NS protease (a.a. 1206-1217), with two glycines on the left and three glycines on the right rendering flexibility, is grafted to the loop which connects strands C'' and D. The model structure is built using the automated comparative protein modeling server SWISS-MODEL. Aromatic residues Trp (black) are shown as sticks. The Ca²⁺ ion is shown as a sphere. Intrinsic Trp fluorescence emission spectra (B) and far UV CD spectra (C) spectra of CD2 (open circle) and engineered protein CD2.RUBCa (closed circle) are compared to examine the perturbation of insertion on the scaffold protein CD2.D1. Buffers consists of 10 mM Tris, 10 mM KCl, pH 7.4. (D), Electrospray mass spectra of CD2 with grafted EF-loop from RUB NS protease (P) in the presence of a 5-fold molar excess of TbCl₃ (+156).

A variety of spectroscopic methods confirmed that the structure of the CD2.D1 had not been perturbed by the inserted loop either in the absence or presence of metal ions. The Trp intrinsic fluorescence spectrum of CD2 overlapped with that of CD2.RUBCa, both of which exhibited the emission maximum at 325 nm and a shoulder peak at 313 nm (Fig. 5.3B). The far-UV CD spectra of CD2.RUBCa remained similar to that of CD2.D1 alone with a deep trough at 216 nm, characteristic of typical β -sheet elements (Fig. 5.3C). The differences observed could be primarily from the inserted 12-residue loop and flanking triple-glycine linkers. The two dimensional ^1H NMR further showed that the majority of the dispersed signals in the fingerprint regions of the TOCSY and NOESY spectra of CD2.RUBCa were located at the same positions as that of CD2, suggesting that the integrity and packing of the host protein frame are maintained after the insertion of the Ca^{2+} -binding loop from RUB NS protease. In general, the insertion of the foreign acidic sequence did not cause major conformational changes of the host protein, and the overall fold of CD2.RUBCa remained similar to that of CD2.D1. The ability of the grafted protein (calculated MW: 12712.9 Da) to form 1:1 metal/protein complexes was revealed by ESI-MS. As shown in Fig. 5.3, the molecular mass of the engineered protein CD2.RUBCa agreed well with theoretical mass. The presence of 5-fold excess of Tb^{3+} led to the emergence of a new peak with additional masses of 156. Similar results were also obtained for La^{3+} . This provides the foundation for measuring the intrinsic metal-binding affinity with a minimized contribution from protein conformational

change and minimal influence of the host protein environment on the grafted EF-hand loop (69).

5.2.3. Obtaining the metal binding affinities of grafted Ca^{2+} -binding loop by FRET and NMR

The spectroscopic silence of Ca^{2+} makes it extremely challenging to investigate the Ca^{2+} binding properties directly. Therefore, lanthanides are frequently used to probe Ca^{2+} -binding sites due to their similar ionic radii (Ca^{2+} : 1.00 Å; Tb^{3+} : 0.92 Å; La^{3+} : 1.03 Å) and metal coordination chemistry. In CD2.RUBCa, the four aromatic residues in the host protein and their distances (<20 Å) to the inserted binding pocket made it possible to monitor the binding of Tb^{3+} using FRET (Fig. 5.3A). As shown in Fig. 5.4A, the addition of increasing amounts of CD2.RUBCa to 10 μM Tb^{3+} resulted in an increase of Tb^{3+} emission fluorescence intensity at 545 nm when excited at 282 nm, suggesting the formation of a metal:protein complex. However, the addition of both CD2.D1 and the engineered protein with two mutated ligands ($\text{D}^5\text{A}/\text{D}^{12}\text{A}$) had substantially weaker enhancement than that of CD2.RUBCa, suggesting that the mutant lost its specific Tb^{3+} binding capability. This study further confirmed that Tb^{3+} bound to the inserted binding loop involving ligands D^5 and D^{12} . By monitoring the changes of Tb^{3+} emission enhancement of fixed amount of proteins as a function of Tb^{3+} concentrations (Fig. 5.4A) and by assuming a 1:1 (Tb^{3+} to protein) binding model, a dissociation constant (K_d) of $47 \pm 4 \mu\text{M}$ was obtained for CD2.RUBCa. In addition, the dissociation constant of Ca^{2+} was indirectly estimated by metal

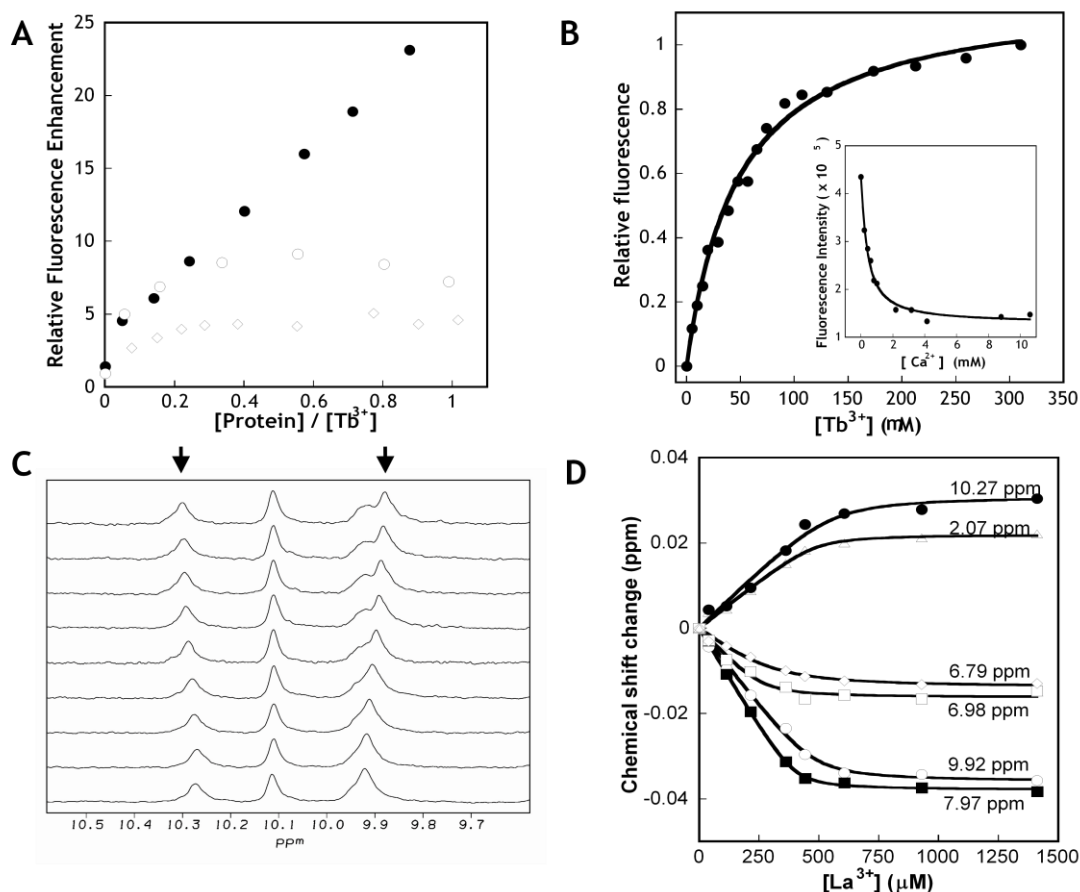


Figure 5.4. Obtaining metal-binding affinity using aromatic residue sensitized Tb^{3+} -FRET and 1D ^1H NMR. (A), Tb^{3+} fluorescence enhancement at 545 nm due to resonance energy transfer (excited at 282 nm) as a function of protein concentrations of CD2 (open diamond), CD2.RUBCa (closed circle) and its mutant $\text{D}^5\text{A}/\text{D}^{12}\text{A}$ (open circle). (B), Tb^{3+} titration of CD2.RUBCa. Normalized fluorescence intensity was plotted as a function of the Tb^{3+} concentration. The inset showed the Ca^{2+} competition titration curve of CD2.RUBCa (1.5 μM) preincubated with 40 μM Tb^{3+} in 20 mM PIPES, 10 mM KCl, pH 6.8. (C), La^{3+} titration of CD2.RUBCa monitored by 1D ^1H NMR. Some resonances (arrows) at amide regions of 1D ^1H NMR spectra of CD2.RUBCa (0.2 mM) shifts with the increased concentrations of La^{3+} (from bottom to top: 0, 39.2, 113.2, 214.3, 442.2, 360.7, 605.1, 929.1, and 1411.2 μM , respectively) in 20 mM PIPES, 10 mM KCl at pH 7.4. (D), the chemical shift change at different resonant regions as a function of the concentration of La^{3+} .

ion competition assay, in which 1.5 μM of CD2.RUBCa with 40 μM Tb^{3+} were titrated with increasing amounts of Ca^{2+} in excess. As shown in the inset of Figure 3B, the Tb^{3+} fluorescence emission maximum at 545 nm decreased due to the competition between Tb^{3+} and Ca^{2+} for the grafted metal binding site, which gave an apparent dissociation constant of $395 \pm 135 \mu\text{M}$ and a real dissociation constant of $214 \pm 73 \mu\text{M}$ according to Eq. 3.

1D ^1H NMR was conducted to monitor the La^{3+} -induced chemical shift changes of CD2.RUBCa as a function of La^{3+} concentration. Several resonances such as those at 7.97, 6.98, 6.79, and 2.67 ppm (Fig. 5.4C) gradually shifted or sharpened as the La^{3+} concentration increases, indicating a single binding process. The La^{3+} -binding affinity, with an average of $14 \pm 7 \mu\text{M}$, was obtained by plotting the chemical shift of these peaks as a function of La^{3+} concentration (Fig. 5.4D).

5.2.4. Metal selectivity of the isolated EF-hand Ca^{2+} -binding motifs from RUB NS protease

Competition assays based on Tb^{3+} -FRET have been used as a convenient method to probe the binding capacity of physiologically competing metal ions (e.g. 100 mM K^+ and 10 mM Mg^{2+}) to EF-hand Ca^{2+} -binding proteins, such as calmodulin (355) and galactose-binding protein (356). As seen in Fig. 5.5A, a significant decrease in Tb^{3+} -FRET elicited from Tb^{3+} -loaded CD2.RUBca was only observed with following the addition of 1 mM Ca^{2+} or 0.1 mM La^{3+} (another Ca^{2+} analog, as discussed above). However, the addition of 100 mM K^+ or 10

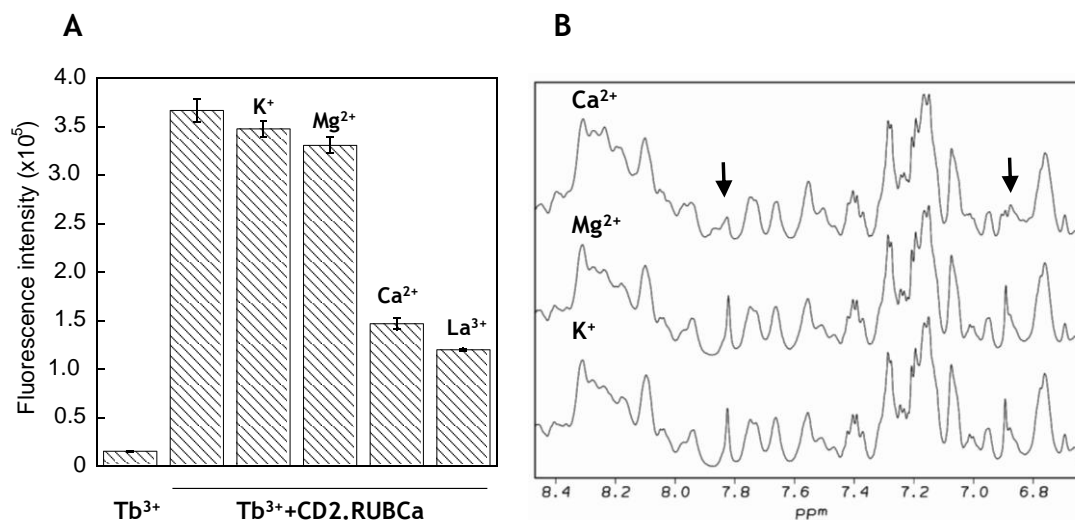


Figure 5.5. Metal selectivity of the engineered protein CD2.RUBCa. (A), Metal competition assay. The addition of 1.5 μM CD2.RUBCa to free Tb^{3+} (40 μM) solution resulted in the increase of fluorescence intensity at 545 nm by over 20 folds due to the binding of Tb^{3+} to the protein and the resultant FRET. 100 mM K^+ , 10 mM Mg^{2+} , 1 mM Ca^{2+} , and 0.1 mM La^{3+} , were subsequently added to individually prepared solutions containing 40 μM Tb^{3+} and 1.5 μM CD2.RUBCa. (B), Amide region of 1D 1H NMR spectrum of CD2.RUBCa with sequential addition of 100 mM K^+ , 10 mM Mg^{2+} , and 1 mM Ca^{2+} . Resonances that exhibited changes are indicated by arrows.

mM Mg^{2+} led to only a slight decrease (<10%) in the intensity of Tb^{3+} fluorescence. To further clarify whether the engineered protein CD2.RUBCa exhibited selectivity for Ca^{2+} , we monitored the 1D ^1H NMR spectrum with sequential additions of K^+ , Mg^{2+} and Ca^{2+} . As shown in Fig. 5.5B, while excess K^+ and Mg^{2+} did not result in any changes in the chemical shifts, Ca^{2+} was able to specifically induce changes of chemical shifts in the main chain amide proton region even in the presence of 100 mM K^+ and 10 mM Mg^{2+} . These results clearly demonstrate that CD2.RUBCa is capable of binding Ca^{2+} with selectivity over Mg^{2+} or K^+ , similar to other known EF-hand proteins such as calmodulin (357,358).

5.2.5. Oligomeric states of the grafted EF-hand motif

The EF-hand motifs tend to occur in pairs in most well-characterized EF-hand-containing Ca^{2+} -binding proteins (359). To clarify whether the predicted EF-hand motif is capable of undergoing dimerization in solution, we further grafted the 29-residue helix-loop-helix EF-hand motif to CD2.D1 (denoted as CD2.RUBCa.EF). Pulsed-field gradient NMR (PFG-NMR), a valuable technique for the study of molecular motions and measurement of the dimensions of molecules in solution (360), was then applied to determine the oligomeric state of the grafted EF-hand motif from the RUB NS protease. As shown in Fig 5.6, the 1D ^1H NMR signals corresponding to internal standards (dioxane and PIPES) and proteins decayed as the field strength increased. Small molecules (dioxane [88 Da] and PIPES [335 Da]) had a significantly faster decay than the protein

samples (12 KDa). The integrated areas of resonances corresponding to dioxane, PIPES and the Ca^{2+} -free or Ca^{2+} -loaded engineered protein CD2.RUBCa.EF were fitted using equation 6 with $R^2 > 0.999$. The obtained diffusion constants for dioxane, PIPES, apo-CD2.RUBCa.EF and Ca^{2+} -CD2.RUBCa.EF were $98.9 \times 10^7 \text{ cm}^2/\text{s}$, $61.9 \times 10^7 \text{ cm}^2/\text{s}$, $13.5 \times 10^7 \text{ cm}^2/\text{s}$ and $13.4 \times 10^7 \text{ cm}^2/\text{s}$ respectively. Using the reported hydrodynamic radius of dioxane (2.12 Å) as a reference of size, the effective hydrodynamic radii of CD2.RBUCa.EF in the absence and presence of Ca^{2+} are 15.53 and 15.41 Å, respectively. Similar to our previous finding that the 12-residue EF-hand loop from EF hand site III of calmodulin remain as a monomer (70), this piece of data suggests that the isolated EF-hand motif remains as a monomer even after metal binding in solution. However, we still cannot rule out the possibility that other motifs in the protease might lead to oligomerization. These findings are reminiscent of the ER-residing protein stromal interaction molecule 1 (STIM1) that contains a single EF-hand. STIM1 is a 90-KDa type 1 single-pass transmembrane protein that is believed to be one of the key components of store operated Ca^{2+} entry (SOCE). The ER-luminal portion of STIM1 protein has a single EF-hand motif and a sterile-alpha motif (SAM),. Biophysical studies on the luminal EF-SAM domain of STIM1 reveal that ER Ca^{2+} -depletion results in the oligomerization of EF-SAM domain mediated by the SAM motif. In contradiction to the EF-hand pairing paradigm, the protein remains as a monomer at high Ca^{2+} concentration (303).

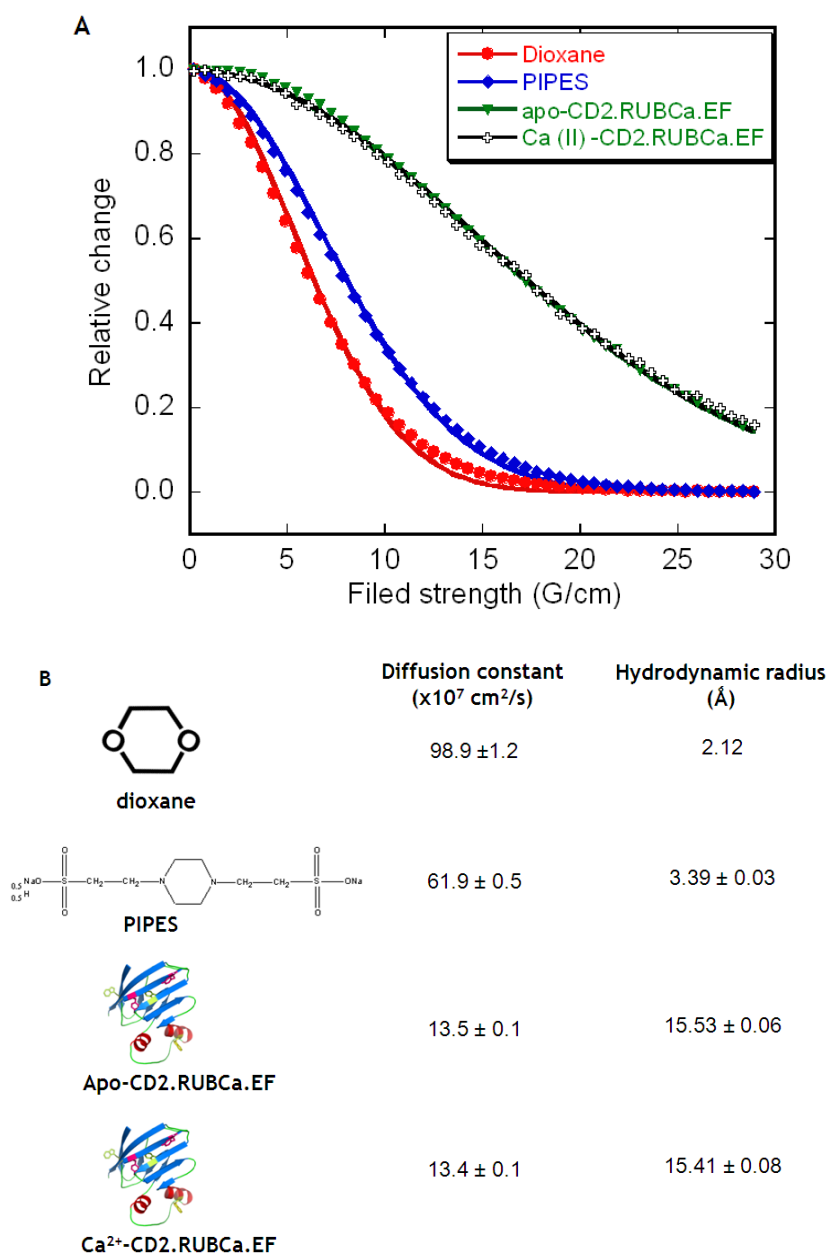


Figure 5.6. Determination of oligomeric states of the grafted EF-hand motif by PFG NMR. (A), 1D ^1H NMR signal decay (integrated area) as a function of pulsed field gradient strength. (B), Diffusion constants and hydrodynamic radii of samples in 20 mM PIPES-10 mM KCl, pH 6.8 at 25 °C.

5.2.6. Using the Minimal Metal Binding Domain RUBCa for Metal-Binding and Conformational Studies.

While our attempts to express the entire NS protease (aa 1000-1300 of P150) were fruitless, we successfully expressed as a GST-fusion protein and purified the minimal metal binding domain (RUBCa, aa 1143-1252, see Fig. 5.2A), which contains the Ca^{2+} binding loop, both catalytic residues, and the ligands proposed to be involved in Zn^{2+} binding. Purified RUBCa was used to characterize the predicted EF-hand Ca^{2+} -binding motif in its native protein environment. When purified RUBCa was reconstituted in the presence of Ca^{2+} and subjected to ICP-MS to measure its metal content, it was found to bind Ca^{2+} with a molar ratio ($\text{Ca}^{2+}/\text{protein}$) of 0.7 ± 0.2 to 1 ($n=2$).

RUBCa contains three Trp residues and according to the model of the structure of the RUB NS protease (Figure. 5.2C), one of these aromatic residues, W1153, is in close proximity to the predicted Ca^{2+} -binding loop (3.5 Å). This feature enabled us to probe the metal-binding properties of RUBCa using Tb^{3+} -FRET. As shown in Figure 5A, a significant enhancement of Tb^{3+} fluorescence intensity was observed when Tb^{3+} was added to RUBCa, indicating Tb^{3+} binding to the predicted loop. A K_d for Tb^{3+} of 3 ± 1 μM was obtained (inset, Fig. 5.7A).

In comparison with the 355 nm for free Trp, the emission maximum of the purified RUBCa blue-shifted to 339 nm, suggesting that at least some Trp in RUBCa was partly shielded from the solvent though not fully buried inside the hydrophobic core (Fig. 5.7B). The shoulder at 355 nm suggests that some Trp residues were exposed to the solvent. No Ca^{2+} -induced emission peak position

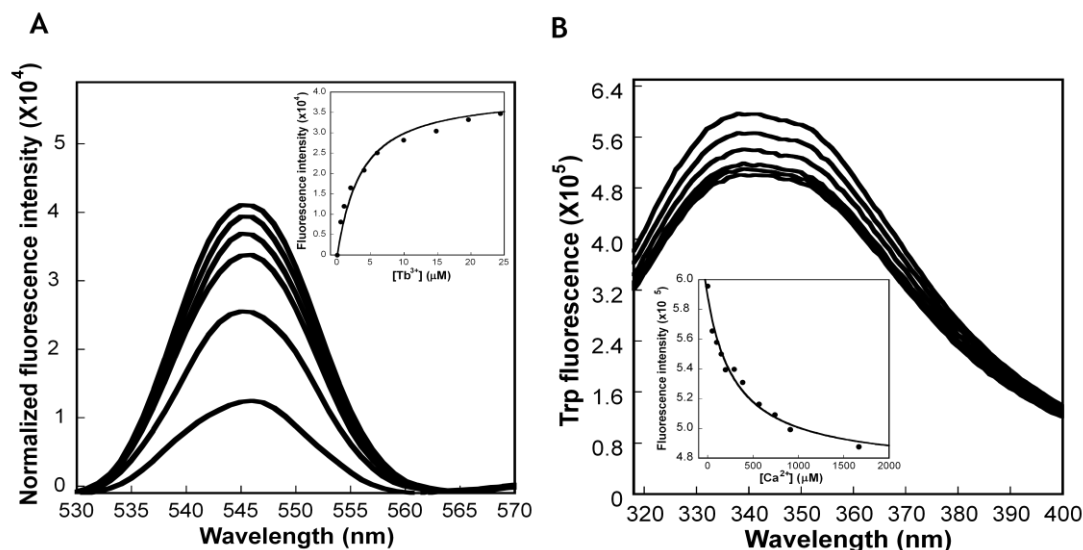


Figure 5.7. Metal ions titration of the minimal metal-binding domain RUBCa monitored by aromatic residue sensitized Tb^{3+} fluorescence (A) and intrinsic Trp fluorescence (B). (A), Normalized Tb^{3+} fluorescence spectra of RUBCa with increasing concentration of Tb^{3+} (from bottom to top: 0, 1.0, 4.0, 9.9, 14.8, 19.6, 14.4 μM , respectively). The inset showed the Tb^{3+} fluorescence enhancement at 545 nm due to energy transfer as a function of the concentration of Tb^{3+} . (B), Intrinsic Trp fluorescence emission spectra of RUBCa (2.5 μM) with increasing concentration of Ca^{2+} (From top to bottom: 0, 49.8, 291.3, 566.0, 740.7 and 909.1 μM , respectively). Inset: the intrinsic Trp fluorescence intensity plotted as a function of the concentration of Ca^{2+} . An average dissociation constant of 316 μM was obtained by assuming a 1:1 binding model. The excitation wavelength was set at 282 nm. All the buffers used in metal titration consist of 20 mM PIPES, 10 mM KCl, pH 6.8.

change was observed. However, the addition of Ca^{2+} led to a decrease in the emission intensity, indicating the local changes of the chemical environment around the aromatic residues (Fig. 5.7B). By monitoring this intensity change, a dissociation constant of $316 \pm 4 \mu\text{M}$ for Ca^{2+} was obtained (Fig. 5.7B, inset), which was in agreement with the Ca^{2+} -binding affinity obtained using the grafted protein CD2.RUBCa (Table 5.1).

Far UV CD was performed to reveal any possible changes of secondary structure of RUBCa induced by Ca^{2+} . As seen in Fig. 5.8A, the spectra of both EGTA-treated and Ca^{2+} -loaded RUBCa had two troughs at 222 and 206 nm, indicating the existence of α -helical secondary structure. Using the program DICHROWEB (66), the best fit of the CD spectrum of RUBCa indicated that 17.8% was α -helix and 19.5% was β -sheet, whereas the remainder was random coil. In excess Ca^{2+} , the CD signal of RUBCa at 208 nm and 222 nm was 5% more negative than that of RUBCa in 1 mM EGTA and the DICHROWEB-predicted α -helix and β -sheet contents were 23.8% and 14.9%, respectively. Thus, the observed gain in negative ellipticity could be attributed to the formation of a higher degree of α -helical content induced by Ca^{2+} binding. The anionic amphiphile ANS was further used as a hydrophobic probe to examine the conformational properties of RUBCa. As shown in Fig. 5.8B, upon the addition of RUBCa, the emission peak of ANS fluorescence blue-shifted from 510 nm to 500 nm and the maximal emission intensity increased by 30%, suggesting that part of the hydrophobic regions of the purified RUBCa were exposed to the solvent and thus accessible to ANS. The addition of excess Ca^{2+} did not cause significant

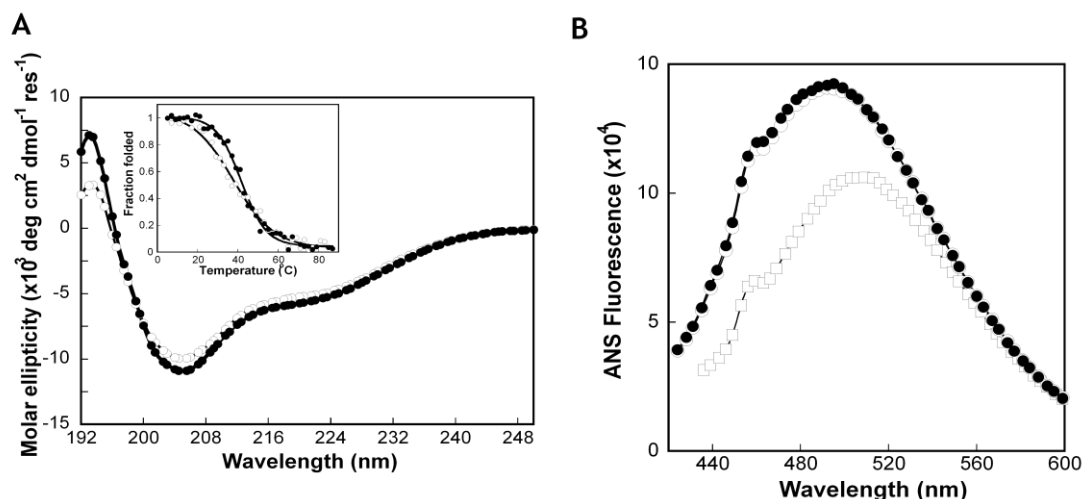


Figure 5.8. Ca^{2+} -induced conformational changes and thermal unfolding of the putative Ca^{2+} -binding domain RUBCa. (A), Far UV CD spectra of RUBCa with 1 mM EGTA (open circle) or 1 mM Ca^{2+} (closed circle) in 10 mM Tris-HCl, 10 mM KCl. Inset: Normalized CD signal at 222 nm plotted as a function of increasing temperature (5-90 $^{\circ}\text{C}$) in the presence of 1 mM EGTA (open circle) or 1 mM Ca^{2+} (closed circle). (B), Fluorescence emission spectra of 40 μM ANS (open square) and ANS:RUBCa complex with 1 mM EGTA (open circle) or 1 mM Ca^{2+} (closed circle). The excitation wavelength was set at 390 nm. The buffer consists of 10 mM Tris-HCl, 10 mM KCl (pH 7.4).

conformational changes in these hydrophobic regions considering the overlapping emission fluorescence spectra of RUBCa in the presence of 1 mM EGTA or 1 mM Ca^{2+} . Taken together, these data indicated that the binding of Ca^{2+} induced a local conformational change, whereas the secondary structure and hydrophobic surface were not significantly altered.

In order to gain more insight into the possible role of Ca^{2+} binding, thermal unfolding was carried out by monitoring the CD signal change at 222 nm as a function of temperature under Ca^{2+} -depleted or Ca^{2+} -loaded conditions. With increasing temperature from 5 to 90 °C, RUBCa gradually underwent thermal denaturation, leading to the decrease of CD signals. Compared with 1 mM EGTA, the melting temperature (T_m) of RUBCa with Ca^{2+} increased from 37.7 ± 0.8 to 41.8 ± 0.4 °C (Fig. 5.8B). The observed increase of melting temperature suggested that the binding of Ca^{2+} stabilized the overall structure of RUBCa under physiological conditions (100 mM KCl).

5.2.7. Mutating the Potential Ca^{2+} -Binding Ligands in Infectious cDNA Clones and Replicons.

To determine if the binding of Ca^{2+} by the RUB NS protease is of physiological significance, the $\text{D}^5\text{A}/\text{D}^{12}\text{A}$ double mutant shown to abrogate binding in CD2.RUBCa was introduced into the Robo502 infectious cDNA clone of RUB (the mutant was termed Robo502 AA). Following transfection, the virus titer obtained from Robo502AA was ~20-fold lower than from the Robo502 parent (Fig. 5.9A), indicating that these residues with their potential Ca^{2+} binding were

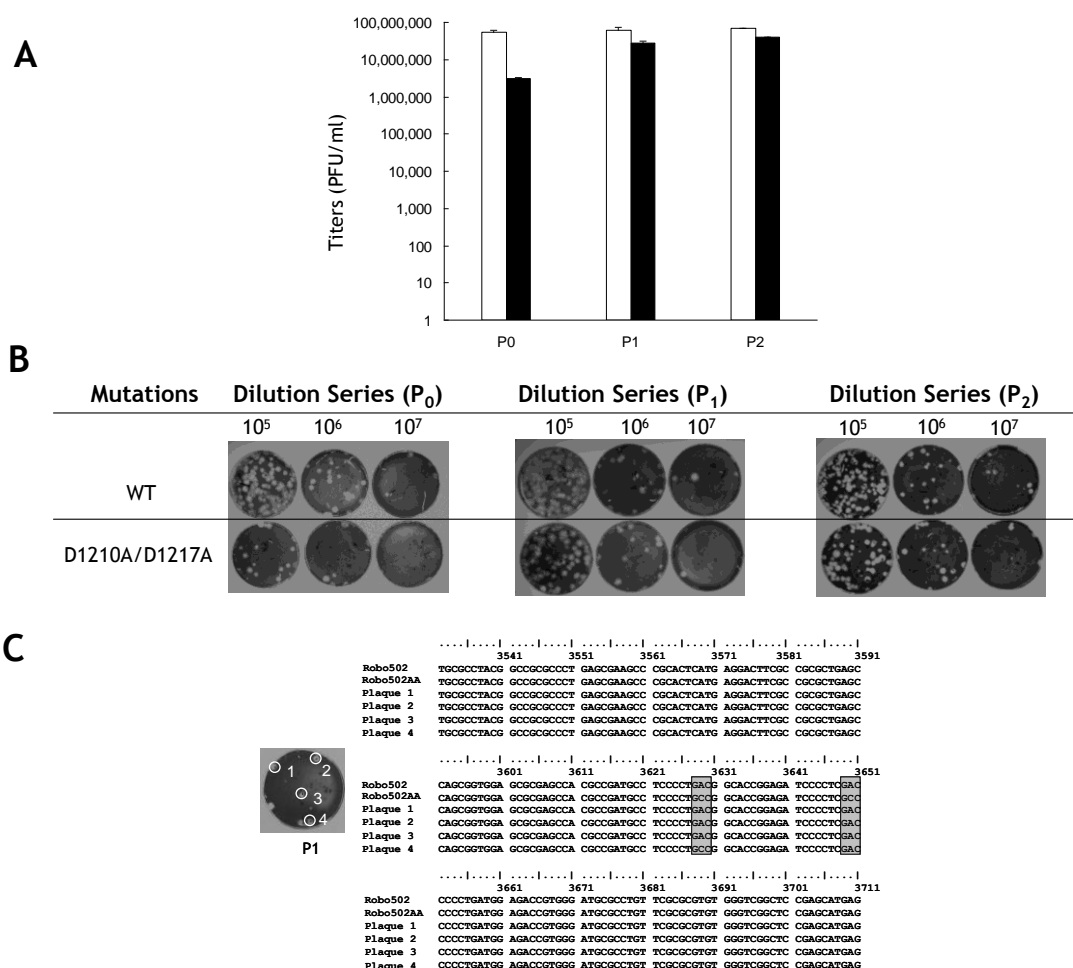


Figure 5.9. Effects of mutations of the potential Ca^{2+} -coordination ligands on RUB replication. Transcripts from the WT infectious cDNA clone, Robo502, or Robo502 AA containing the D¹²¹⁰A and D¹²¹⁷A mutations in the Ca^{2+} -binding loop, were used to transfect Vero cells. Culture fluid from the transfection plate (P0) was harvested on day 7 post-transfection and passaged twice in Vero cells (P1 and P2). The virus titer in the P0, P1, and P2 culture fluids was determined by plaque assay in triplicate (A). Open bar: Robo502; Black bar: Robo502AA. Representative plaques at each passage are shown in B. To check for the generation of revertants in the Robo502AA population, four plaques were picked from terminal plaque assay dilution plates from P1 (left, C) culture fluid and after one round of amplification in Vero cells, the sequence of the metal binding domain in the NSP was determined, as shown in Panel C in comparison to the wt sequence.

necessary for optimal RUB replication. Following two subsequent passages of the transfected culture fluids, the Robo502AA titer rose to within threefold of Robo502 (Fig. 5.9A and B). When four individual plaques from the first passage of Robo502AA were amplified and the sequence of the NS protease determined, three were found to have reverted to the wt sequence at both sites while the fourth had reverted to the wt sequence at one site (Fig. 5.9C). Thus, the increase in titer was due to reversion of the mutant sequences.

To prevent the accumulation of revertants, the AA mutation was also introduced into the replicon, RUBrep/GFP, in which the SP-ORF is replaced with the GFP reporter gene. Following transfection, RUBrep/GFP replicates and expresses GFP but does not spread from cell to cell. As shown in Fig. 5.10A, RNA synthesis by RUBrep-AA/GFP was delayed by ~1 day post-infection in comparison with the wt replicon. Concomitantly, GFP expression was similarly delayed (data not shown). We also assayed P200 NSP precursor cleavage in cells transfected with RUBrep-HA/GFP derivative, replicons expressing a P150 tagged with the HA epitope (71). These experiments were done at 6 hrs post-transfection, when translation from the input transcripts is detectable but replication has not yet started. As shown in Fig. 5.10B, at 35 °C P200 to P150 cleavage was efficient in wt RUBrep-HA transfected cells, but only ~50% efficient in RUBrepAA-HA/GFP-transfected cells. RUBrep-NS*-HA/GFP, a construct with a C¹¹⁵²S substitution at the catalytic site unable to undertake cleavage, served as an uncleaved control. Given our biophysical measurements on RUBCa that indicate that Ca²⁺ binding increased the T_m of RUBCa from 37.7 °C to 41.8 °C

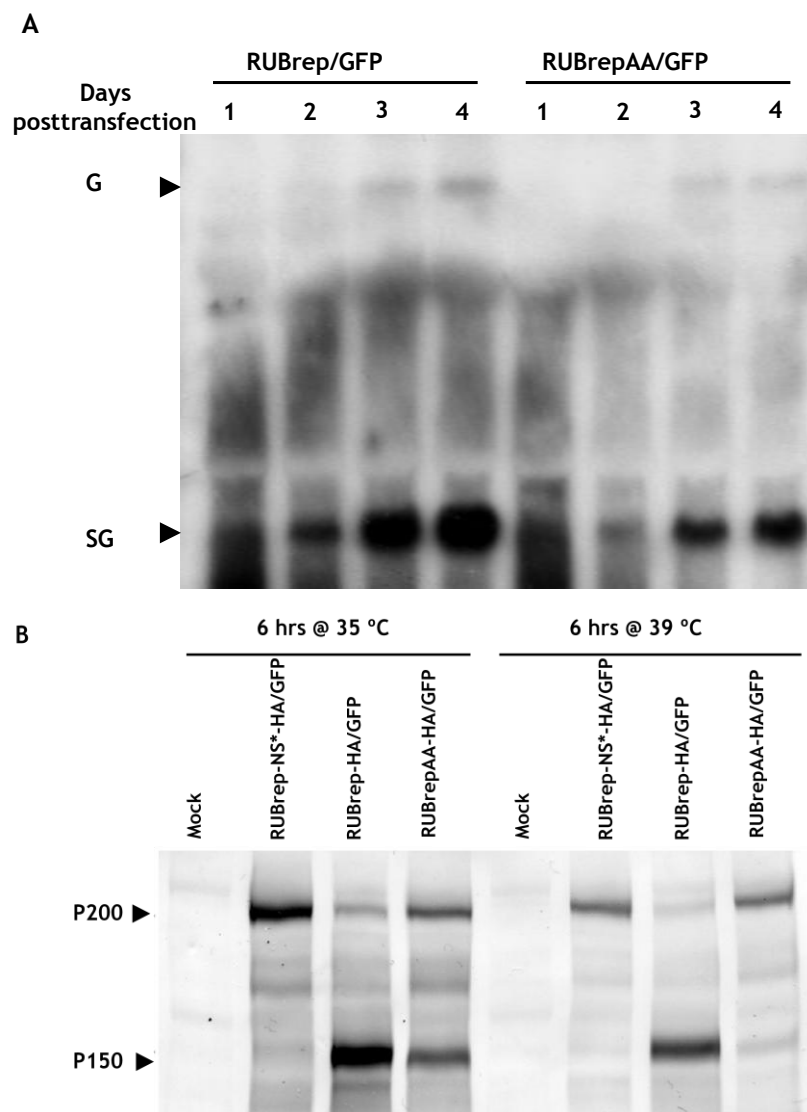


Figure 5.10. Replicon RNA synthesis and P200 cleavage in transfected Vero cells. Vero cells were transfected with transcripts from RUBrep/GFP or RUBrepAA/GFP containing mutations D⁵A and D¹²A (A) or from RUBrep-HA/GFP or RUBrepAA-HA/GFP, which express an HA-epitope-tagged P150 (B). (A), Total cell RNA was extracted 1 to 4 days posttransfection, and replicon plus-strand RNA species (G, genomic; SG, subgenomic) were resolved by Northern blotting following agarose gel electrophoresis. (B), Transfections were performed at 35°C or 39°C. Six hours posttransfection, cells were lysed and the P200 precursor and P150 product were resolved by Western blotting probed with anti-HA antibodies following SDS-PAGE (the other product, P90, does not appear because it does not contain the HA epitope). Cells transfected with a replicon containing a C¹¹⁵²S catalytic site mutation (RUBrep-NS*-HA/GFP) that cannot mediate P200 cleavage served as an uncleaved control.

(Fig. 5.10B), we also did the experiment in cells transfected at 39 °C. As shown in Fig. 5.10B, wt replicon cleavage of the P200 precursor was more efficient than at 35 °C, however, cleavage in AA mutant replicon-transfected cells could not be detected, indicating that the mutant protease is temperature sensitive. Thus, not only is Ca^{2+} binding important for NS protease activity, in the absence of the Ca^{2+} binding the protease is rendered temperature sensitive.

5.2.8. Summary

Identification of an EF-Hand Ca^{2+} -Binding Motif in a Viral Nonstructural Protein. In this report, we have presented one of the few predictions and confirmations of a Ca^{2+} -binding motif in a virus nonstructural protein. Most other viral Ca^{2+} -binding motifs occur in virion structural proteins and of these, only the rotavirus VP7 outer capsid protein (DITADPTTAPQTE) (39), the HIV-1 transmembrane protein gp 41 (44), and the polyomavirus VP1 (DENGVGPLCKGE) (31) contain EF-hand Ca^{2+} -binding motifs. In a canonical EF-hand Ca^{2+} -binding motif, the coordination geometry in the loop is formed by oxygen atoms from the sidechain carboxyl or hydroxyl groups (loop positions 1, 3, 5, 12), the main chain carbonyls (position 7), and a bridged water (position 9) and the flanking sequences around the loop preferentially form a helical secondary structure. The negatively-charged residue pairs at Z axis (loop positions 5 and 12) are essential for the binding of cations (358). Examination of the primary sequence of the motifs in the above-mentioned virion structural proteins reveals that none of these are genuine canonical EF-hand Ca^{2+} -binding

moieties due to deviations in flanking sequence structural requirements (HIV gp 41), the length of the loop (13 residues instead of 12 in rotavirus VP7), and the coordinating residues (V instead of D/E/N/Q/T/S at loop position 5 in polyomavirus VP1). Thus, the Ca^{2+} -binding motif in the RUB NS protease is the first to be identified with a canonical EF-hand structure.

Interestingly, the same bioinformatics algorithm that successfully predicted the Ca^{2+} -binding loop in the RUB NS protease also predicted an EF-hand Ca^{2+} -binding motif in nsP1 of alphaviruses (Fig. 5.11A). nsP1 is one of the four nonstructural proteins produced by alphaviruses and is involved in membrane binding and has methyl/guanylyl transferase activity. We have grafted this loop from Sindbis virus (SINV) into the CD2 scaffold and determined its K_d for Tb^{3+} as 16.4 μM (Fig. 5.11B). We are in the process of creating the corresponding mutation from the RUBV loop in the SINV loop ($\text{E}^5\text{I}/\text{E}^{12}\text{I}$). If we find that this mutation abrogates Ca^{2+} binding by the grafted SINV loop, it will be of great interest to determine if infectivity is affected after we introduce mutations into a SINV infectious cDNA clone.

Because of difficulty in expressing the entire NS protease, the metal-binding properties of this Ca^{2+} -binding motif were probed by grafting the loop into the scaffold protein CD2.D1 and by expressing the entire motif within the native minimal metal-binding domain of the NS protease. Table 1 provides a summary of the metal-binding affinities of the Ca^{2+} -binding motif in both the engineered CD2.RUBCa and the minimal metal binding domain, RUBCa. It is worth noting that numerous studies of EF-hand calcium binding motifs have been made at the

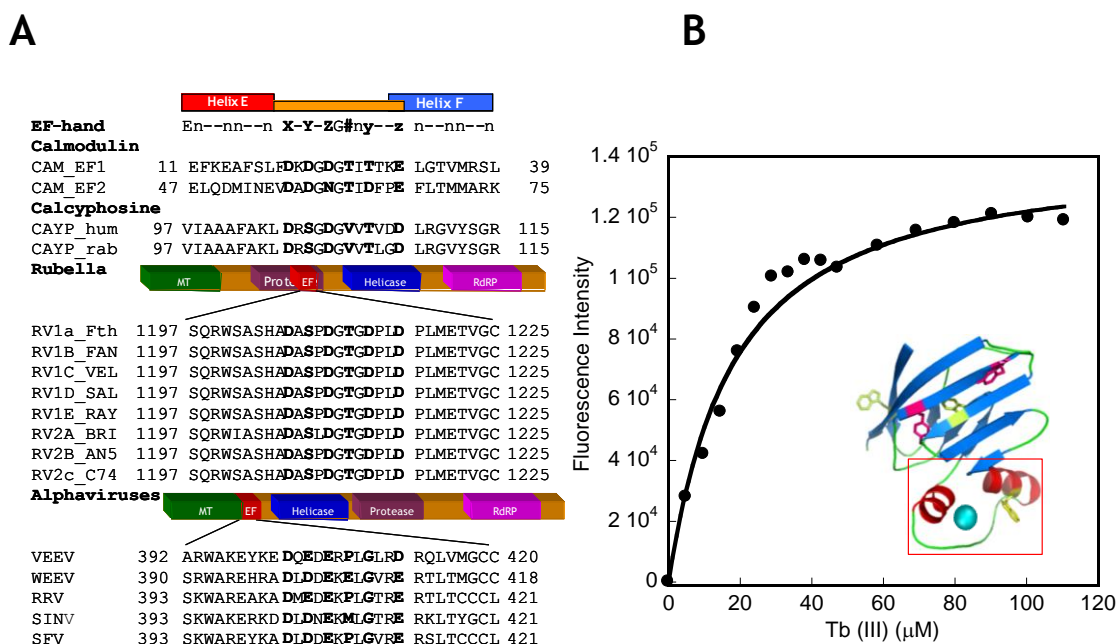


Figure 5.11. Prediction of a putative EF-hand Ca^{2+} -binding motif in the nsP1 of alphaviruses. (A), Location of predicted EF-hand Ca^{2+} -binding motifs in rubella virus and alphaviruses, both of which belong to the *Togaviridae* family. The EF-hand Ca^{2+} -binding motif in alphaviruses is immediately next to the methyltransferase domain of nsP1. VEEV, Venezuelan equine encephalitis virus; WEEV, Western equine encephalitis virus; RRV, Ross River virus; SINV, Sindbis virus; SFV, Semliki Forest virus. (B), Grafting the 29-residue helix-loop-helix EF-hand motif (red box, inset) from sindbis virus (SINV) to CD2.D1. The engineered protein binds to Tb^{3+} with a dissociation constant of 16 μM .

peptide or subdomain level (361-363). The Ca^{2+} -binding affinity of the RUBCa loop grafted into CD2 (214 μM) is weaker than the grafted EF loop I from calmodulin (34 μM) (69), likely due to the presence of an Asp at the bidentate position 12 of the loop rather than a Glu, which ensures a high Ca^{2+} -binding affinity due to its larger side chain and stronger interaction with Ca^{2+} . Substitution of Glu at position 12 in CaM with Lys or Gln reduces Ca^{2+} -binding affinity by 10-100 fold (361,364-366) while the substitution of Asp with Glu at position 12 in the regulatory light chain of myosin leads to a 15-fold increase in binding affinity (350). The Ca^{2+} -binding motif in the RUB NS protease also exhibits selectivity for Ca^{2+} and La^{3+} in the presence of excess Mg^{2+} and K^+ , similar to paired EF-hand Ca^{2+} -motifs in Ca^{2+} -binding proteins such as calmodulin and troponin C (367).

Table 5.1. The metal-binding affinities of the engineered protein CD2.RUBCa and the minimal metal-binding domain RUBCa.

Protein	Metal ions	Dissociation constants (K_d , μM)
CD2.RUBCa	^a Tb^{3+}	47 ± 4
	^b Ca^{2+}	214 ± 73
	^c La^{3+}	14 ± 7
RUBCa	^a Tb^{3+}	3 ± 1
	^d Ca^{2+}	316 ± 4
	Zn^{2+}	^d 0.2 ± 0.1 ^e 0.8 ± 0.2

Results obtained from Tb^{3+} -FRET assay ^a, Ca^{2+} competition assay ^b, 1D ^1H NMR with La^{3+} titration ^c, metal titration monitored by intrinsic Trp fluorescence ^d and ANS fluorescence ^e.

Ca²⁺-Induced Local Conformational Change in the RUB NS Protease and Its Potential Stabilizing Role. Since the expression and purification of the entire NS protease was not achieved, conformational studies were done using the RUBCa subdomain which contained both catalytic residues, the putative Zn²⁺ binding ligands and the EF-hand Ca²⁺-binding motif. Boundaries of the RUBCa subdomain were chosen near unstructured or loop regions according to our model (Fig. 5.2C) and within this model, the RUBCa subdomain had a well-folded, globular structure. We found that the RUBCa subdomain bound stoichiometric amounts of Ca²⁺ and underwent metal induced local changes in conformation. Ca²⁺ binding to RUBCa resulted in a modest decrease in its intrinsic Trp fluorescence intensity without emission peak position shift. However, the addition of Ca²⁺ did not significantly alter the secondary structure of RUBCa as revealed by far UV CD. In addition, ANS binding suggested that no hydrophobic surface was perturbed upon Ca²⁺ binding. These results indicate that Ca²⁺ binding does not lead to a global conformational change of RUBCa. In the RUB NS protease model structure, the predicted helix-loop-helix EF-hand Ca²⁺-binding site is located at the exposed surface on the opposite side of the active site C¹¹⁵² and H¹²⁷³ (Fig. 5.2C), consistent with our finding that Ca²⁺ binding to RUBCa did not induce global conformational changes. In the model, the W¹¹⁵³ residue is 3.5 Å, on average, away from the predicted EF-hand motif. This proximity could lead to perturbation in the chemical environment around W¹¹⁵³ upon metal-binding and thus the model is also consistent with our observation of significant enhancement of Tb³⁺ fluorescence intensity in RUBCa

due to Tb³⁺-FRET (Fig. 5.4A). Finally, we found that the binding of Ca²⁺ increased the T_m of RUBCA from 37.7 to 41.8 °C indicating that Ca²⁺ binding serves to stabilize the NS protease. Significantly, we found that a replicon with two mutations in the EF-hand Ca²⁺ binding loop that abrogate metal binding by the CD2.RUBCa chimera expressed a protease that was temperature sensitive at 39°C, confirming this prediction. Such a stabilizing role for Ca²⁺ has been observed in non-viral protease, such as subtilisin (368) and thermolysin (369). It will be of interest to determine if the putative EF hand Ca²⁺-binding motif in the MT domain of the alphaviruses plays a similar stabilizing role.

Role of Ca²⁺ in the Replication Cycle of Rubella Virus. When the double Ca²⁺ binding mutant was introduced into the Robo502 RUB infectious cDNA clone, titers produced were ~20-fold lower than those produced by wt Robo502, however during subsequent passaging, the mutant virus titer rose by over 10-fold and sequencing of isolated virus plaques indicated that this was due to reversion of the mutants to the wt sequence. In a replicon incapable of cell-to-cell spread, the double Ca²⁺ binding mutant exhibited delayed RNA synthesis. We finally showed that cleavage of the P200 nonstructural precursor was impaired at 35 °C and temperature sensitive at 39 °C. While minus strand RNA synthesis is catalyzed by the uncleaved precursor, the precursor must be cleaved before plus strand RNA synthesis can occur (370). Thus, an impaired cleavage is consistent with delayed plus strand RNA synthesis as observed with the replicon and lower titers as observed with the infectious cDNA clone.

However, ongoing (though delayed) RNA synthesis would lead to the generation of revertants as observed with the infectious cDNA clone.

A submillimolar K_d for Ca^{2+} binding ($\sim 200\text{-}300\ \mu\text{M}$) was derived from both CD2.RUBCa and RUBCa. In the cell, cytosolic Ca^{2+} concentrations range from $0.1\text{-}10\ \mu\text{M}$ while the Ca^{2+} concentration in endosomes or lysosomes is as high as $400\text{-}600\ \mu\text{M}$ and is maintained in part by the proton gradient across lysosomal membranes (371,372). Presumably, in the replication complex the protease resides on the cytosolic side of the lysosomal membrane, however release of Ca^{2+} from the lysosome could result in local Ca^{2+} concentrations high enough for binding. For example, in acidifying endosomes the uptake of extracellular Ca^{2+} (mM) is immediately accompanied by rapid release of Ca^{2+} to the cytoplasm and with an external $[\text{Ca}^{2+}]$ of 2 mM, the endocytosed Ca^{2+} may be sufficient to increase the total cellular Ca^{2+} concentration by $2\ \mu\text{M}$ per min (372). It is also possible that association of the viral replicase proteins with late endosomal or lysosomal membranes alters the Ca^{2+} gradient leading to a localized increase in Ca^{2+} concentration. Nevertheless, given the weak Ca^{2+} binding affinity of the Ca^{2+} binding loop, the protease is possibly inactive in the cytosol and must be associated with the late endosomal or lysosomal membrane before encountering Ca^{2+} concentration sufficient for binding. Thus, Ca^{2+} may play a role in regulating the virus replication cycle by delaying the nonstructural precursor cleavage until it is associated with late endosomal or lysosomal membranes and thus modulating the synthesis of plus strand RNA. A similar situation has been shown to be the case with rotaviruses whose NSP4 glycoprotein mediates an increase in

intracellular Ca^{2+} concentration which modulates transcriptase activity (37,229,286,373).

5.3. Conformational and functional studies on a Zn^{2+} -binding cysteine-rich domain of RUB NS protease

Viral Zn^{2+} -binding proteins include proteins that interact with nucleic acids and are important in transcription and/or replication, such as UL52 of HSV-1, a member of a helicase-primase complex (374), the nucleocapsid of retroviruses (HIV-1 and murine leukemia virus) (375,376), the V protein of Sendai virus involved in RNA editing (377), helicases of nidoviruses such as SARS coronavirus and equine arteritis virus (378,379), and the NS5A replicase component of hepatitis C virus (HCV) (380-382), and viral proteases, such as the papain-like cysteine protease 2 of SARS coronavirus (383), nsp1 papain-like cysteine protease of equine arteritis virus (a nidovirus) (384), the leader protease of encephalomyocarditis virus (385), and the NS3 protease of HCV (386). Zn^{2+} has also been to be required in the RUB NS protease (345,346); however, the binding ligands have yet to be identified.

Structural Zn^{2+} binding sites are generally coordinated by four cysteines and histidines as ligands and based on our model structure of the RUB NS protease, a cluster of cysteine residues (C^{1167} , C^{1175} , C^{1178} , C^{1225} , and C^{1227}) could potentially form a Zn^{2+} binding pocket, in which there is no histidine residue nearby (Fig. 5.12). We have expressed a subdomain of the NS protease, RUBCa (aa 1143-1252 of P150), which contains all of the putative cysteine

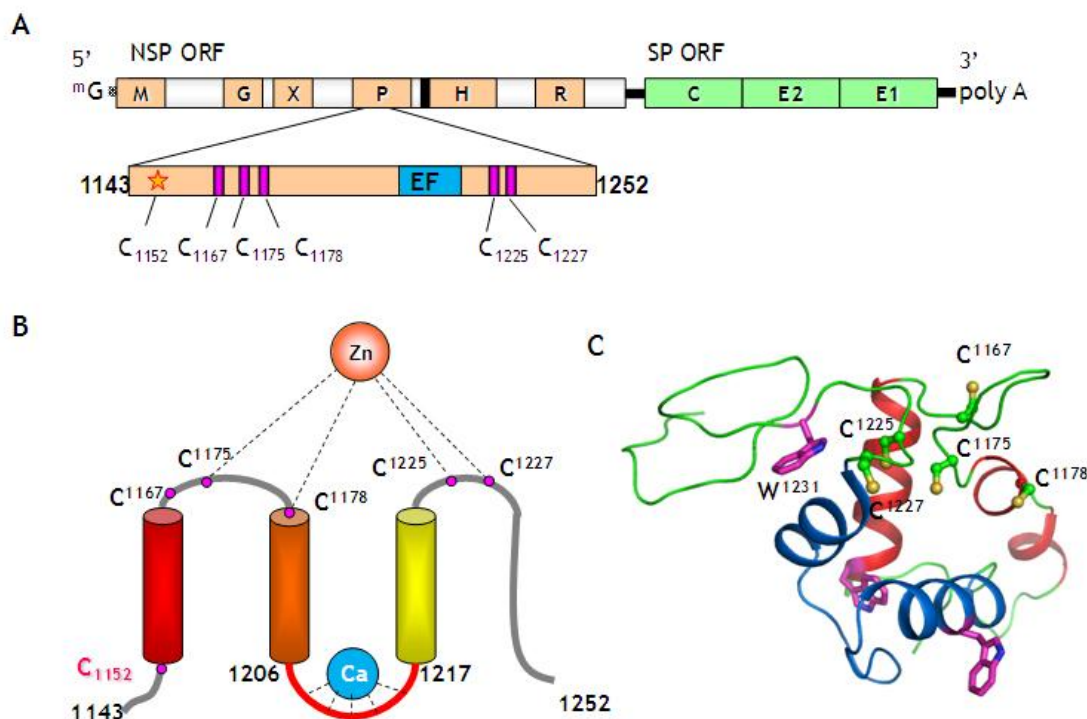


Figure 5.12. Cartoon representation of a cysteine-rich domain in RUB NS protease. (A), Domain assignment of the RUB genome. The protease domain contains at least one Zn^{2+} -binding site as well as a putative EF-hand Ca^{2+} -binding motif (blue box). It has been proposed that the residues C¹¹⁵² and H¹²⁷³ constitute the catalytic dyad. The minimal metal-binding domain RUBCa (aa 1143 to 1252), encompassing both the Zn^{2+} - and Ca^{2+} -binding motifs, is bacterially expressed and used in this study. (B) Secondary structure prediction of RUBCa and the location of Zn^{2+} - and Ca^{2+} -binding motifs. (C), Model structure of RUBCa. A cluster of cysteine residues (shown as ball-and-stick) forms a potential Zn^{2+} -binding pocket. The EF-hand Ca^{2+} -binding motif is shown in blue. The aromatic residue W¹²³¹ (magenta) situates close to the cysteine-rich region and serves as a probe for Zn^{2+} -induced conformational changes.

residues involved in Zn^{2+} binding plus the EF-hand Ca^{2+} binding motif (Section 5.2), as a GST fusion protein and found that it could be cleaved and purified, remained soluble after purification (Fig. 2.1), bound stoichiometric amounts of Ca^{2+} and Zn^{2+} , had a relatively well-folded structure, and underwent metal induced changes in conformation. Unfortunately, limited solubility of this subdomain precludes structural analysis by NMR to identify the Zn^{2+} binding ligands. Therefore, the residues involved in Zn^{2+} binding will be identified by mutating each cysteine in RUBCa individually to S (C^{1152} , the catalytic site residue, will be mutated to S as a negative control) and measuring Zn^{2+} content by a modified PAR (4-(2-pyridylazo)resorcinol) colorimetric assay (387). The P200 precursor cleavage in cells transfected with RUBrep-HA/GFP derivative, replicons expressing a P150 tagged with the HA epitope and its cysteine mutants will be similarly assayed as described in section 5.2.7.

5.3.1. Determination of Zn^{2+} contents in RUBCa and its cysteine mutants

In previous studies, C1175, C1178, C1227, H1204 and H1273 have been assigned as candidate ligands involved in zinc binding on the basis of site-directed mutagenesis and in vitro ^{65}Zn binding assay (345,346). However, in the model structure, we found that no histidine was present in close proximity to the proposed cysteine-rich Zn^{2+} -binding pocket. Hence, we solely focused on the cysteine residues.

After purifying minidomain RUBCa and its mutants and removing background metal ions by Chelex 100 resin, we determined the zinc content with

a colorimetric assay by using the dye 4-(2-Pyridylazo)resorcinol (PAR). The protein-bound Zn^{2+} is released by protease K digestion. The released Zn^{2+} will then form a $\text{Zn}-(\text{PAR})_2$ complex accompanied by a large increase in extinction coefficient at 500 nm ($\Delta\epsilon = 6.6 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ at pH 7.0) (Fig. 5.13A). The affinity of PAR for Zn^{2+} is high with $K_1' = 4.0 \times 10^6 \text{ M}^{-1}$ and $K_2' = 5.5 \times 10^5 \text{ M}^{-1}$ at neutral pH (387). In the presence of 100 μM PAR, 1-15 μM Zn^{2+} released from a protein is converted rapidly ($< 3\text{ms}$) and quantitatively to Zn-PAR complex (of which over 98% of the zinc is in the 1:2 complex) with a ratio of $[\text{Zn}-(\text{PAR})_2]$ to $[\text{Zn}^{2+}]_{\text{free}}$ of $\sim 2 \times 10^4$. Shown in Fig. 5.13B is the zinc to protein molar ratio determined by this method, with carbonic anhydrase as positive control ($[\text{Zn}^{2+}]/[\text{P}] = 0.88$). Except for the WT RUBCa ($[\text{Zn}^{2+}]/[\text{P}] = 0.69$) and the mutant C1167S ($[\text{Zn}^{2+}]/[\text{P}] = 0.67$), all the other cysteine mutants (C^{1175}S , C^{1178}S , C^{1225}S , C^{1227}S , $\text{C}^{1167}\text{S}/\text{C}^{1175}\text{S}$, $\text{C}^{1175}\text{S}/\text{C}^{1178}\text{S}$, $\text{C}^{1167}\text{S}/\text{C}^{1175}\text{S}/\text{C}^{1178}\text{S}$, and $\text{C}^{1225}\text{S}/\text{C}^{1227}\text{S}$) showed significantly lower metal to protein ratio ($[\text{Zn}^{2+}]/[\text{P}] = 0.08 \sim 0.29$). With all of the five cysteine residues substituted with serines (C5 mutant), the protein fails to bind zinc ions. Thus, it seems that C^{1167} is not directly involved in the coordination of zinc ions in the binding pocket.

5.3.2. Zn^{2+} -induced conformational changes

RUBCa contains three Trp residues and according to the model of the structure of the RUB NS protease (Fig. 5.12C), one of these aromatic residues, W^{1231} , is in close proximity to the proposed Zn^{2+} -binding pocket. Aromatic residue is highly sensitive to local or global conformational changes and this feature has

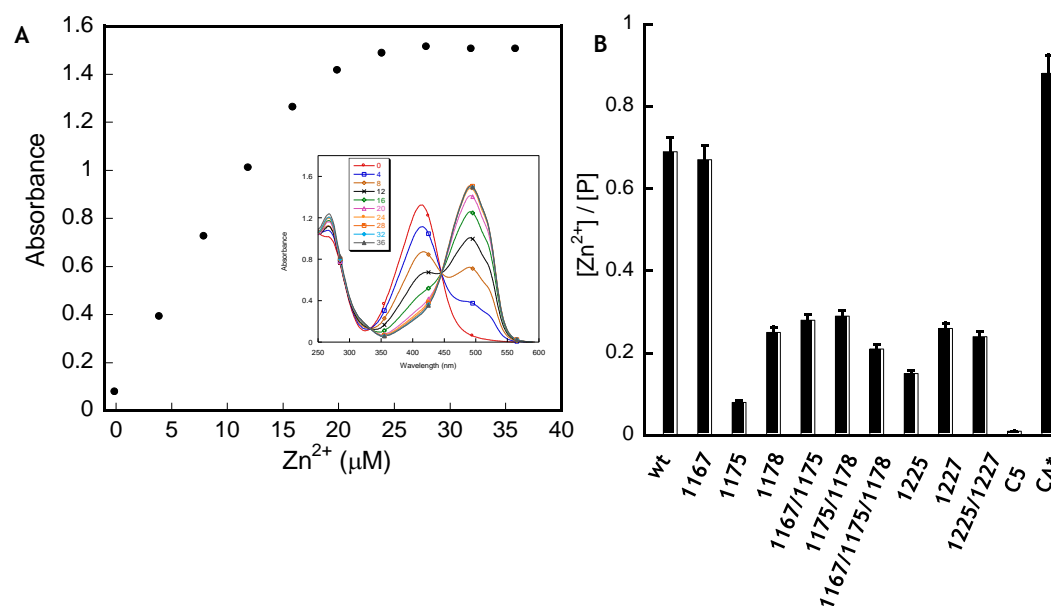


Figure 5.13. Determination of protein-bound Zn²⁺ content by PAR assay. (A), A PAR colorimetric assay is used to determine the Zn²⁺ content in proteins. The binding of zinc released from protein to PAR results in an increase of signal at 500 nm (inset). The linear range of this method lies within 0-15 μM Zn²⁺. (B), Zinc to protein ratio in wt RUBCa and its cysteine to serine mutants. The C5 mutant substitutes all of the five cysteines with serines. Carbonic anhydrase, which contains one high-affinity zinc binding site, is used as positive control (marked as CA*).

been utilized to probe Zn^{2+} -induced structural changes. As shown in Fig. 5.14A, the emission maximum of RUBCa blue-shifted from 339 nm to 336 nm with concomitant increase of fluorescence intensity by at least 15%, indicating that the aromatic residue W¹²³¹ is surrounded by a more hydrophobic environment. By monitoring this intensity change, a dissociation constant of $0.2 \pm 0.1 \mu\text{M}$ for Zn^{2+} was obtained by assuming 1:1 binding (inset, Fig. 5.14A).

Zn^{2+} -induced conformational change was further confirmed by ANS fluorescence. The anionic amphiphilic ANS is often used as a hydrophobic probe to examine the conformational properties of proteins. As shown in Fig. 5.13B, upon the addition of RUBCa, the emission peak of ANS fluorescence blue-shifted from 510 nm to 500 nm and the intensity increased by ~30%, suggesting that some of the hydrophobic regions of the purified RUBCa were exposed to the solvent and thus accessible to ANS. The addition of excessive amounts of Zn^{2+} caused redshift (by 2 nm) and ~10% decrease of the fluorescence intensity (Fig. 5.13B). This result suggests that the binding of Zn^{2+} leads to conformational changes and the subsequent shielding of some hydrophobic regions from the solvent in RUBCa. A dissociation constant of $0.8 \pm 0.2 \mu\text{M}$ was obtained by monitoring the ANS fluorescence intensity change as a function of Zn^{2+} concentration (Inset, Fig. 5.14B). Taken together, these data implied that the binding of Zn^{2+} induced significant conformational changes by rearrangement of some hydrophobic regions.

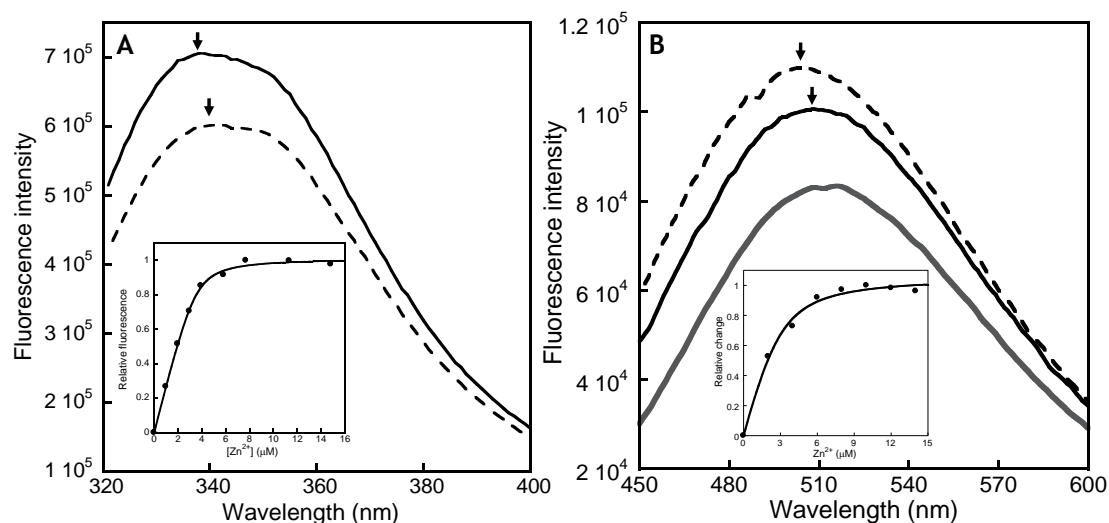


Figure 5.14. Zn²⁺-induced conformational changes. (A), Intrinsic Trp fluorescence emission spectra of RUBCa (2.5 μM) in the presence of 1 mM EGTA (dashed line) or 20 μM Zn²⁺ (solid line). Inset: the relative intrinsic Trp fluorescence intensity plotted as a function of the concentration of Zn²⁺. An average dissociation constant of 0.2 μM was obtained by assuming a 1:1 binding model. The excitation wavelength was set at 282 nm. (B), Fluorescence emission spectra of 40 μM ANS (gray line) and ANS:RUBCa complex with 1 mM EGTA (dashed line) or 20 μM Zn²⁺ (solid line). The excitation wavelength was set at 390 nm. All the buffers used in metal titration consist of 10 mM Tris-HCl, 100 mM KCl, pH 7.4.

5.3.3. Effects of Cys-to-Ser mutations on protease cleavage

We further introduced the individual or combined cysteine to serine mutation(s) into a RUBrep-HA/GFP replicon expressing HA-tagged P150 to assay the P200 cleavage in Vero cells. These experiments were done at 6 hrs post-transfection, when translation from the input transcripts is detectable but replication has not yet started. As shown in Fig. 5.15, at 35 °C P200 to P150 cleavage was efficient in wt RUBrep-HA transfected cells, as well as in RUBrep-HA/GFP-C1167S transfected cells. Vero cells transfected with RUBrep-HA/GFP constructs with single cysteine to serine mutations at residue 1175, 1178, 1225, 1227, similar to the construct with a C¹¹⁵²S substitution at the catalytic site that served as an uncleaved control, were unable to undertake cleavage. In addition, cleavage in Vero cells transfected RUBrep-HA/GFP constructs of multiple cysteine to serine mutations (C¹¹⁶⁷S/C¹¹⁷⁸S, C¹¹⁷⁵S/C¹¹⁷⁸S, C¹¹⁶⁷S/C¹¹⁷⁵S/C¹¹⁷⁸S, C¹²²⁵S/C¹²²⁷S, and the mutant C5S with all five cysteines mutated to serines) was abolished (Fig. 5.15). This finding is largely consistent with our previous *in vitro* co-transcription-translation assay, which demonstrated that single mutations of C¹¹⁷⁵, C¹¹⁷⁸, C¹²²⁷ could result in the loss of protease activity. However, in the *in vivo* assay, C¹¹⁶⁷ was able to cleave the P200 precursor and C¹²²⁵ was unable to undertake the P200 precursor cleavage, which is exactly opposite to the observations in the *in vitro* assay. This discrepancy would likely arise from the difference in reaction conditions. Specifically, *in vitro* assay is carried out under aerobic, non-reducing conditions that may not truly reflect the reducing environment of cell-based assays. Given the high number of cysteines, a



Figure 5.15. Effect of mutations of cysteines on P200 precursor cleavage. Vero cells were transfected with transcripts from WT RUBrep/GFP or RUBrep/GFP mutants containing single or multiple cysteine-to-serine substitutions. Six hours posttransfection, cells were lysed and the P200 precursor and P150 product were resolved by Western blotting probed with anti-HA antibodies following SDS-PAGE (the other product, P90, does not appear because it does not contain the HA epitope). Cells transfected with a replicon containing a C¹¹⁵²S* catalytic site mutation that cannot mediate P200 cleavage served as uncleaved control. C3S: C1167/1175/1178S; C5S: C1167/1175/1178/1225/1227S.

reducing environment is more favorable for Zn^{2+} -binding and protein folding by preventing formation of mismatched disulfide bonds. Given those, the *in vivo* assay carried out in Vero cells would be more reliable and closer to the true situation occurring in the cells.

5.3.4. Secondary and tertiary structure of RUBCa and its mutants

For site-directed mutagenesis studies, caution has to be taken for the interpretation of the observed results, since there are always concerns regarding whether single or multiple cysteine-to-serine mutations would disrupt the integrity of the protein structure, and thus leading to loss of its function. To clarify this point, we compared the secondary and tertiary structure of the minidomain RUBCa and its cysteine mutants. As shown in Fig. 5.16, no significant difference were detected among the far UV CD spectra and intrinsic Trp fluorescence spectra of RUBCa and its mutants, C3S and $\text{C}^{1225}\text{S}/\text{C}^{1227}\text{S}$, suggesting the structural integrity was retained after mutagenesis. Thus, it can be concluded that the loss of proteolytic activity in the cysteine-to-serine mutants is due to its compromised Zn^{2+} binding capability.

5.3.5. A possible iron-containing domain in RUBCa

During the expression and purification of the minidomain RUBCa, we observed a striking brown color for the protein solution. A typical UV/Vis spectrum of the purified RUBCa is shown in Fig. 5.16A. In addition to the main

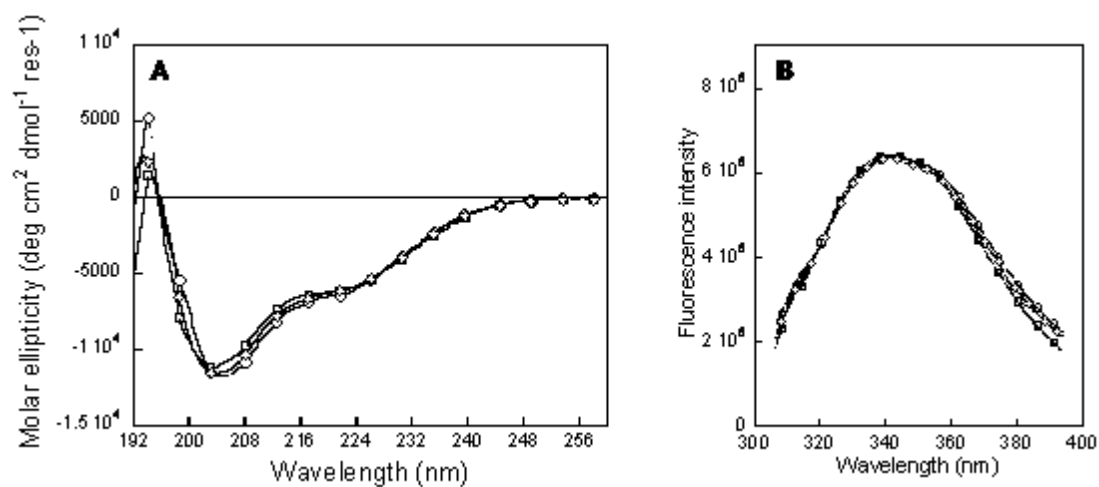


Figure 5.16. Far UV CD (A) and Trp fluorescence (B) spectra of 2 μ M RUBCa (\circ) and its mutants containing triple ($C^{1167}S/C^{1175}S/C^{1178}S$, \square) and double ($C^{1225}/C^{1227}S$, \diamond) cysteine mutations.

peak at 280 nm, two additional peaks that are typical of iron-sulfur clusters, one at 325 nm and the other at 420 nm, were recorded. In addition, the Fe to protein ratios for GST-RUBCa (n=1) and RUBCa (n=1) determined by ICP emission spectrometer were 0.05:1 and 0.18:1, respectively (Fig. 5.17). It is evident that the bacteria-expressed minidomain contains Fe. The iron is so tightly bound to the protein that it could not be chelated by EGTA, EDTA or citrate (Fig. 5.17A). However, the addition of a strong reducing agent, sodium dithionite (E_o' : ~ -420 mV at pH 7), resulted in the decrease of absorbance at 420 nm (Fig. 5.17B), suggesting the degradation of the iron-sulfur cluster. To narrow down the cysteines that are involved in the formation of the FeS cluster, we further examined the absorbance spectra of RUBCa Cys-to-Ser mutants. As shown in Fig. 5.17C, the absorbance peak at 420 nm disappeared after substituting all of the five cysteines with serines (C5S mutant). However, both the triple mutant C^{1167/1175/1178}S and the double mutant C^{1225/1227}S retained a peak at 420 nm. Given our findings at Sections 5.3.1-5.3.3, it seems that both Zn²⁺ and Fe²⁺ (or Fe³⁺) ions are sharing or competing some of the cysteine ligands. In the future, we will further map the exact ligands and further characterize the FeS cluster by using Mössbauer spectroscopy and/or electron paramagnetic resonance spectroscopy.

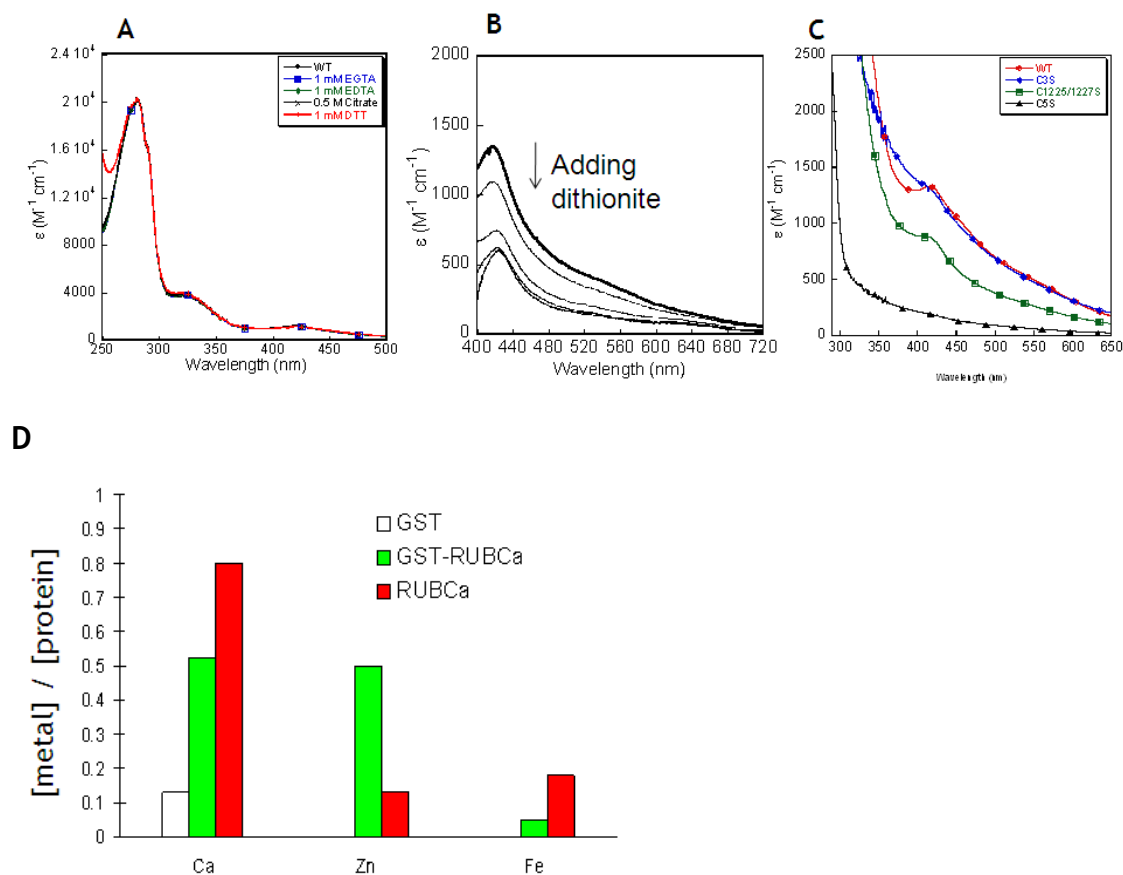


Figure 5.17. Absorbance spectra and metal content of RUBCa. (A) UV/vis spectrum of RUBCa (20 μM) with different chelators in 10 mM Tris-HCl at pH 7.5. (B) Gradual addition of 0.1 mg dithionite to the protein sample leads to decrease of absorbance signals. (C) UV/Vis spectrum of RUBCa Cys-to-Ser mutations. C3S, C^{1167/1175/1178S}, C5S, C^{1167/1175/1178/1225/1227S}. (D) Metal content of GST, GST-RUBCa, and RUBCa determined by ICP emission spectrometry.

5.3.6. Summary

In conclusion, site-directed mutagenesis studies carried out on the bacterially-expressed minidomain RUBCa and the RUB replicon constructs have defined an atypical CxxC(x)₄₈CxC type of Zn²⁺ binding motif. The coordination of Zn²⁺ by C¹¹⁷⁵, C¹¹⁷⁸, C¹²²⁵, and C¹²²⁷, with an affinity comparable to intracellular Zn²⁺ level, is required for the proteolytic activity of the protease domain in P150. The presence of Fe ions in the bacteria-expressed minidomain raises the possibility that the protease domain could also be involved in the redox reactions.

6. Prediction and identification of calmodulin binding sites in gap junction protein connexins

6.1. Calmodulin as intracellular Ca^{2+} sensor

Ca^{2+} , an important second messenger in eukaryotic cells, is regarded as a “signal for life and death” (3,388). Temporal and spatial changes of the Ca^{2+} concentration in different compartments of cells affect the regulation of cellular signaling, by modulating the activity of numerous Ca^{2+} -binding proteins. Among them, CaM is one of the most crucial players and is extensively studied due to its ubiquitous expression in eukaryotes and its versatile ability to activate or inhibit more than 100 functional enzymes, cellular receptors and ion channels (14,53,56,389-394).

CaM is a small (148 amino acids; MW: 16.7 KDa) and acidic (pI: ~4.0) EF-hand Ca^{2+} -binding protein first identified in the brain and heart as cyclic nucleotide phosphodiesterase activator protein (PAF) in the early 1960's (48). CaM consists of two globular and autonomous domains, each of which contains two helix-loop-helix EF-hand motifs. Through its reversible or irreversible binding to Ca^{2+} , the resultant conformational changes and the interaction with target proteins, CaM is capable of transducing the intracellular Ca^{2+} signal changes into a myriad of divergent cellular events, such as cell proliferation, cell differentiation and apoptosis (3).

6.1.1 Binding of metal ions to CaM: affinity, cooperativity and specificity

Calcium binding affinities and cooperativity

The coordination of Ca^{2+} in EF-hand motifs is achieved by adopting a pentagonal bipyramidal geometry with seven oxygen atoms from the sidechain carboxyl or hydroxyl groups (loop sequence positions 1, 3, 5, 12), the main chain carbonyls (position 7), and a bridged water (position 9). The Glu at position 12 serves as a bidentate ligand for the Ca^{2+} ion (Fig. 2A). In the coordination sphere, ligands at position 1 and 9 (bridged water) constitute the X axis; ligands at position 3 and 7 form the Y axis, whereas ligands at position 5 and 12 are located along the Z axis (Fig. 1.3). Position 6 of the 12-residue EF-loop is always occupied by a glycine and position 8 is highly reserved as a hydrophobic residue. Residues at position 7-9 form an antiparallel β -sheet that is involved in the pairing of two adjacent EF-hands. The exiting helix begins at position 10.

More than 50 mutations within the EF-hand loops have been made to study the key determinants of metal-binding affinities (361). The bidentate ligand at position 12 is least tolerant of mutations. An X-ray crystallographic structure that captures the conformational intermediate of CaM has been determined by locking the N-terminal domain in a “closed” state with a disulfide bond between residues 41 and 75. In this structure, the bidentate Glu residue within the C-terminal domains are about 6-7 Å farther away from its normal position to coordinate Ca^{2+} . The movement of Glu to a position 2.5 Å from the Ca^{2+} is required to couple the chelation of Ca^{2+} and the conformational change by moving the exiting helix (395). Mutations of coordinating ligands, in general, lead

to significant decreases in Ca^{2+} -binding affinity or less selectivity towards other metal ions with similar ionic radii. Based on studies of a series of mutants on site III of CaM where additional acidic residues are introduced on either the X or Z axis, Reid *et al* has proposed an “acid-pair” hypothesis to correlate the type of chelating residues with calcium affinity (396,397). According to this hypothesis, a higher Ca^{2+} -binding affinity could be expected if the chelating ligands at X or Z axis were paired by acidic residues. This simplified model could be used as a guideline to predict Ca^{2+} -binding affinity, while excluding the contribution of nonchelating residues and cooperativity. The nonchelating residue at position 8 has been reported to play an important role in the maintenance of stability of CaM. Browne *et al* reported that the substitution of the highly-conserved hydrophobic residue at position 8 by Gly results in the destabilization of CaM under both Ca^{2+} -free and Ca^{2+} -loaded states (398). These findings underscore the importance of the antiparallel β -sheet that structurally links the neighboring paired EF-hand Ca^{2+} -binding sites. The conserved hydrophobic residue at position 8 has been regarded as part of the “EF-hand β -scaffold” (395). In paired Ca^{2+} -binding sites, the bonding network forms a stable resonance structure that involves the hydrogen bonding between residues at position 8, and the interaction of the bound Ca^{2+} with the oxygen atoms from the main-chain carbonyl at position 7 (Figs. 6.1A and B). Based on this particular structural feature, an “EF-hand β -scaffold” model has been proposed by Grabarek to explicitly explain the Ca^{2+} -binding process in the EF-loop with two steps (395,399). First, the position of Ca^{2+} is defined by the EF β -scaffold and the N-

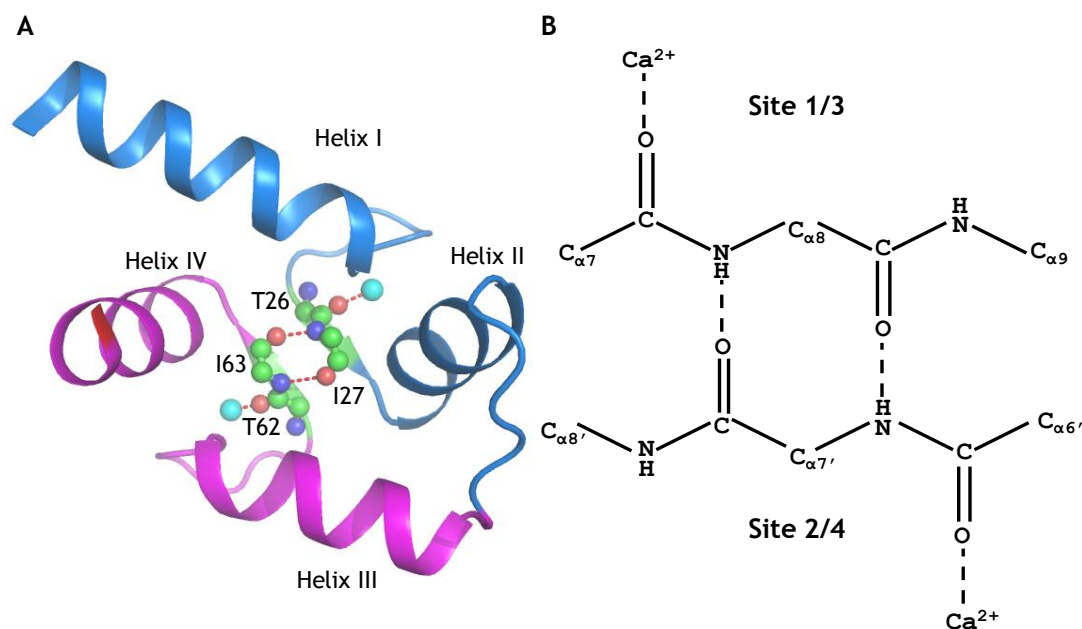


Figure 6.1. The EF-hand β -scaffold structure formed by paired adjacent EF-hand Ca^{2+} binding sites, the N-terminal domain of CaM. (A) 3-D representation of the N-terminal domain of CaM (pdb code: 3cln). The residues involved in the formation of the β -scaffold structure are labeled and shown as ball-and-stick. Ca^{2+} ions are shown as cyan spheres. (B) The formation of a resonance structure by the paired adjacent EF-hand Ca^{2+} binding sites.

terminal part of the Ca^{2+} -coordinating ligands. Subsequently, the C-terminal bidentate ligand Glu in the EF-loop approaches the immobilized Ca^{2+} with concomitant movement of the exiting F helix to complete the Ca^{2+} coordination (399).

The binding of Ca^{2+} to CaM exhibits a positive cooperativity with a Hill coefficient of 1.8-2.5. Ca^{2+} binds to CaM in a sequential order: the C-terminal domain (sites III and IV) is firstly occupied by two Ca^{2+} ions with an affinity about 10^{-7} M, followed by the binding of two additional Ca^{2+} ions to the “low-affinity” N-terminal domain (sites I and II) with the affinity around 10^{-6} . The rates of Ca^{2+} association to these Ca^{2+} binding sites range from 10^6 to $10^8 \text{ M}^{-1}\text{S}^{-1}$, whereas the rates of Ca^{2+} dissociation is on the order of 10 to 10^3 S^{-1} . The 10-fold difference in the Ca^{2+} -binding affinities between two domains, as well as the cooperative binding of Ca^{2+} to each pair of sites, enables CaM to efficiently sense the narrow range of cellular calcium concentration at rest and excited conditions, and thus switching on/off Ca^{2+} signaling cascades by reversible interaction with target proteins. The interaction of CaM with target proteins has been reported to enhance Ca^{2+} -binding affinities by 10-1000 fold (400,401) due to slower off-rates. The free energy change (ΔG) of Ca^{2+} binding to two coupled EF-hand motifs and the accompanying Ca^{2+} -induced conformational changes are governed by the following equations:

$$\Delta G = \Delta G_1 + \Delta G_2 + \Delta G_{\text{coop}} + \Delta G_{\text{conf}}$$

$$\Delta G_{\text{coop}} + \Delta G_{\text{conf}} = RT \ln \frac{K_{\text{micro } 1} K_{\text{micro } 2}}{K_{\text{micro } 1} K_{\text{micro } 2}}$$

where ΔG_1 and ΔG_2 are the intrinsic energy changes of each 1:1 metal-ligand

binding process, ΔG_{coop} and ΔG_{conf} are the energy contribution from cooperativity between the two metal binding sites and the metal-induced conformational changes, respectively. Four site specific binding constants (microscopic), K_I , K_{II} , $K_{I,II}$ and $K_{II,I}$, are used to describe the corresponding binding steps (Fig. 6.2). To determine the cooperativity of two-coupled sites, either the ratio of K_{II}/K_I plus two macroscopic association constants K_1 and K_2 , or three of the four microscopic binding constants are required to be determined. Tremendous efforts have been made to estimate the cooperativity of Ca^{2+} binding to CaM by using peptide models (362,367), enzymatic fragmentation (363,402,403) and a novel grafting approach (1,2,57,361). According to calculations on the basis of the grafting system, we are able to conclude that the energy contribution of cooperativity and conformational change from the C-terminal domain is 40% greater than that from the N-terminal domain, which could be one of the mechanisms underlying the observed high-affinity Ca^{2+} -binding sites at C-terminal domain (1).

Taken together, these studies reveal that the number and type (main chain or side chain) of ligands, the total ligand charge, the nonchelating residues and the Ca^{2+} -induced conformational changes are closely related to the Ca^{2+} binding affinity (367). No clear-cut and conclusive rules can thus far be derived, suggesting additional complexity to Ca^{2+} binding affinity beyond currently evaluated empirical determinants. Therefore, a comprehensive charge-ligand-balanced model emphasizing the accumulative effects of all these factors is more scientifically valid to explain the variation in Ca^{2+} binding affinities.

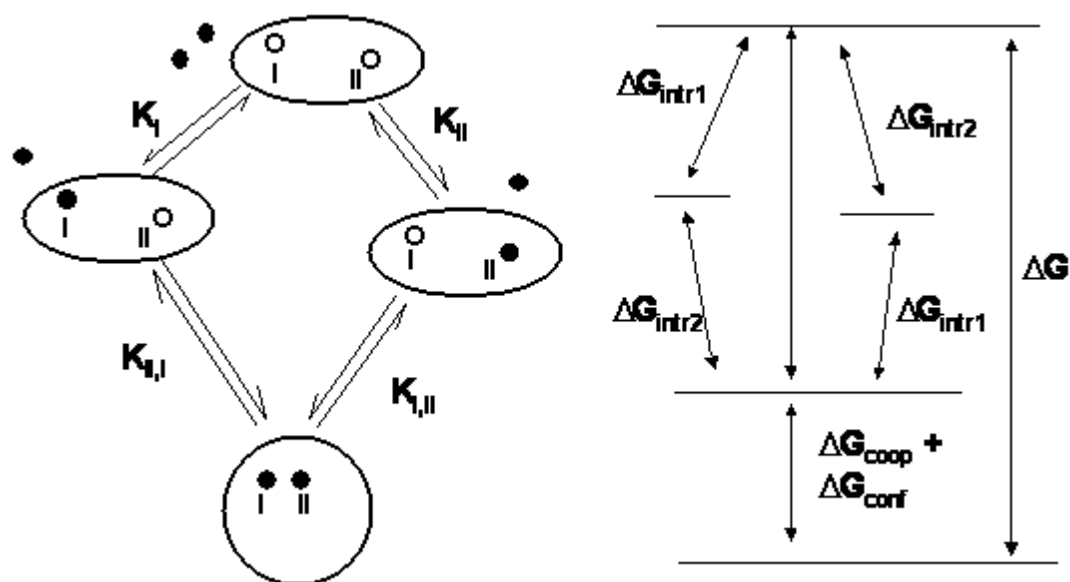


Figure 6.2. Schematic diagram of Ca^{2+} binding to two coupled EF-hand motifs and binding energetics. ΔG is the free energy change of calcium binding to two coupled EF-hand motifs. ΔG_1 and ΔG_2 are the intrinsic energy changes of each 1:1 metal-ligand binding process, ΔG_{coop} and ΔG_{conf} are the energy contribution from cooperativity between the two metal binding sites and the metal-induce conformational changes, respectively. Four site specific binding constants (microscopic), K_I , K_{II} , $K_{I,I}$ and $K_{I,II}$, are used to describe the corresponding binding steps.

Metal binding selectivity

In addition to Ca^{2+} , other metal ions, particularly those sharing similar coordination chemistry or having similar ionic radii as Ca^{2+} (0.99 Å), are capable of binding to CaM with varying binding affinities. Lanthanides (Ln^{3+}), such as Tb^{3+} , La^{3+} and Eu^{3+} , have similar ionic radii and coordination geometry as Ca^{2+} . However, they first bind to the N-terminal domain with stronger affinities (K_d : 6-12 nM) and then to the C-terminal domain with lower affinities (K_d : 2-3 μM), contradictory to the order of Ca^{2+} binding dissociation constants. However, this discrepancy is not observed in the binding constants obtained using the grafting approach (1). The site-specific binding affinities for both Ca^{2+} and Ln^{3+} decrease in the same order: I>III~II>IV. Since the grafting approach eliminates the energy contribution from conformational changes and cooperativity, it is highly possible that the eliminated factors could account for the different stepwise binding modes between Ca^{2+} and Ln^{3+} .

To specifically respond to the altered $[\text{Ca}^{2+}]_c$, it is crucial for CaM to discriminate Ca^{2+} from the overwhelmingly abundant Mg^{2+} (10^{-3} M) and monovalent ions (10^{-1} M) within the cytoplasm. With excessive concentrations of monovalent ions, the Ca^{2+} -binding affinities of CaM decrease over 50 fold due to the screening of electrostatic interactions. To achieve metal selectivity, the Ca^{2+} -binding sites in CaM evolved to adopt a pentagonal bipyramidal geometry that is optimized to favor the size and charge of Ca^{2+} over other metals, such as Mg^{2+} (0.66 Å) and K^+ (1.33 Å) (367,404). Mg^{2+} , a group IIA metal ion above the Ca^{2+} in the periodic table, binds to CaM at the same sites as Ca^{2+} with the dissociation

constants ranging from 10^{-2} to 10^{-4} M, which is ~1,000-10,000-fold weaker than the binding constants of Ca^{2+} . Although the intracellular Mg^{2+} concentration is in the mM range, previous studies have shown that such weak binding will not cause significant effects on the cooperative binding of Ca^{2+} to CaM. However, the interaction of CaM with some of its target peptides or proteins will be affected. For example, with the concentration of Mg^{2+} increased from 0 to 100 mM, the binding affinity of Ca^{2+} -CaM to its target peptide derived from MLCK is lowered by almost 40 folds (405). In view of this, the binding affinities of CaM-target complex obtained under low salt conditions need to be reevaluated with the presence of physiological amounts of Mg^{2+} (406). Furthermore, studies on the Ca^{2+} -binding affinity of the low-affinity N-terminal domain (TR_1C) reveals that Mg^{2+} can occupy almost half of the binding sites under physiological condition without triggering a major structural rearrangement (407,408).

6.1.2. Structure of CaM: conformational plasticity and structural malleability

The first structure of CaM was determined in 1985 by Babu *et al.* Since then, at least 90 structures of CaM and CaM-target complexes under Ca^{2+} -free or Ca^{2+} loaded conditions have been determined by using X-ray crystallography or NMR, which makes a thorough discussion of these structures beyond our scope. However, scrutinization of some of the landmark works conducted by Bax's, Ikura's, Brunger's and Forsen's groups are sufficient enough to reveal the key features of CaM and illustrate the striking difference between apo-CaM and Ca^{2+} -CaM (363,409-411).

Attempts to grow crystals of apo-CaM suitable for X-ray crystallography have proven to be fruitless. The available structures were determined by multidimensional NMR techniques in 1995. In the absence of Ca^{2+} , CaM consists of two sets of compactly packed, twisted, four-helix bundles, one at the N-terminus (helices I-IV) and the other at the C-terminus (helices V-VIII), which are capped by a short anti-parallel β -sheet. The central helical region connecting the two domains is interrupted by S^{81} and is slightly bent. The C-terminal domain turns out to be less stable than the N-terminal domain as indicated by the much faster backbone amide hydrogen exchange rate (410). It has been shown that the helices in apo-CaM are more mobile and less stable, which could explain the observed reduction in helical content, compared Ca^{2+} -CaM, as revealed by circular dichroism studies. Upon binding of Ca^{2+} , CaM undergoes drastic conformational changes and adopts a “dumbbell” shape: two autonomous globular domains are tethered by a flexible linker (residues 77-80). The most striking structural change is the rearrangement of helices in each domain and the increased solvent exposure of hydrophobic surface. The original anti-parallel helices in each EF-hand motifs become nearly perpendicular to each other by undergoing remarkable reorientation (Fig. 6.3). The exposed area of hydrophobic patches to the solvent is estimated to be $10 \times 12.5 \text{ \AA}$ in both the N- and C-terminal domains (409). Therefore, CaM transits from a compact “closed” state to a more relaxed “open” state, poising itself to interact with numerous target proteins (Fig. 6.3A). Comparison of the X-ray crystal structure and NMR structure of Ca^{2+} -CaM indicates that the N-terminal domain in solution is much less

opened than in crystalline, where a fifth Ca^{2+} has been reported to bridge the electrostatic interaction between Glu47 and Asp58 (411). According to the energy balance mechanism proposed by Chazin *et al*, a Ca^{2+} -induced open conformation is more favored because more energy is required to accommodate Ca^{2+} ions in the close conformation (412). A more quantitative expression of these changes is reflected by the alteration in the interhelical angles and distances (Table 6.1). The interhelical angles within each EF-hand decrease by 40° and the interhelical distances increases by 4 Å, on average, by addition of Ca^{2+} to apo-CaM.

Another unique feature of CaM is the unusual abundance of methionines (9/148, 6%) in the primary sequence. Methionine has flexible and relatively polar side chain that results in strong van der Waals interactions. Both domains contain four methionines (M^{36} , M^{51} , M^{71} , M^{72} for the N-terminal domain and M^{109} , M^{124} , M^{144} and M^{145} for the C-terminal domain) that are situated at the gate of the hydrophobic cores and contribute as much as ~50% of the total hydrophobic surface area (Fig. 6.3B). The binding of Ca^{2+} to apo-CaM increases the solvent accessible areas of methionines by 30 fold at the N-terminal domain. The solvent accessible areas of M109 and M45 at the C-terminal undergo 45-fold and 5-fold increases, respectively. The importance of the methionine in the broad spectrum of target recognition has long been realized. Substitution of M by L or Q within the hydrophobic core significantly impairs CaM's capability to activate or deactivate its target proteins, such as phosphodiesterase (413) and the RyR1 (414). Nevertheless, substitution of 8 methionines with unnatural sulfur-free

homolog of methionine, norleucine, neither significantly changes the structure of CaM, nor leads to considerable loss of its ability to activate MLCK (415). The role of methionines in stabilizing the open conformation is further corroborated by Nelson and Chazin's less biased, interaction-based studies on the Ca^{2+} -induced conformational change in CaM by using distance difference matrices and by analyzing the interresidue contacts (416).

Table 6.1. The interhelical angle and distance in apo-CaM and Ca^{2+} -CaM.

helix pair	Interhelical angle ($^{\circ}$)		Interhelical distance ^b (Å)	
	Apo-CaM	Ca^{2+} -CaM	Apo-CaM	Ca^{2+} -CaM
I/II	138 ^a /128 ^b	90 ^c /86 ^d	12.3	18.5
III/IV	130/130	86/86	12.3	16.2
II/III	126/130	111/113	11.2	11.3
I/IV	127/121	110/108	10.7	10.7
V/VI	131/137	102/101	11.9	15.1
VII/VIII	133/132	89/94	12.3	14.4
VI/VII	141/144	110/112	11.8	11.0
V/VIII	142/144	119/116	11.2	10.3

PDB codes: a. 1cfc (410); b. 1dmo (409); c. 1c1l (417); d. 3cln (418).

6.1.3. Target recognition of CaM: diversity and recurrence

Although it has been known from almost four decades that numerous proteins can interact with CaM, it was not until 1992 that the first structure of CaM with a 26-residue target peptide derived from the skeletal myosin light chain kinase (MLCK) was solved (419,420). This landmark work has unveiled several key features responsible for the interaction of CaM with its target proteins and has enabled the prediction of the CaM target sequences. This pioneering work

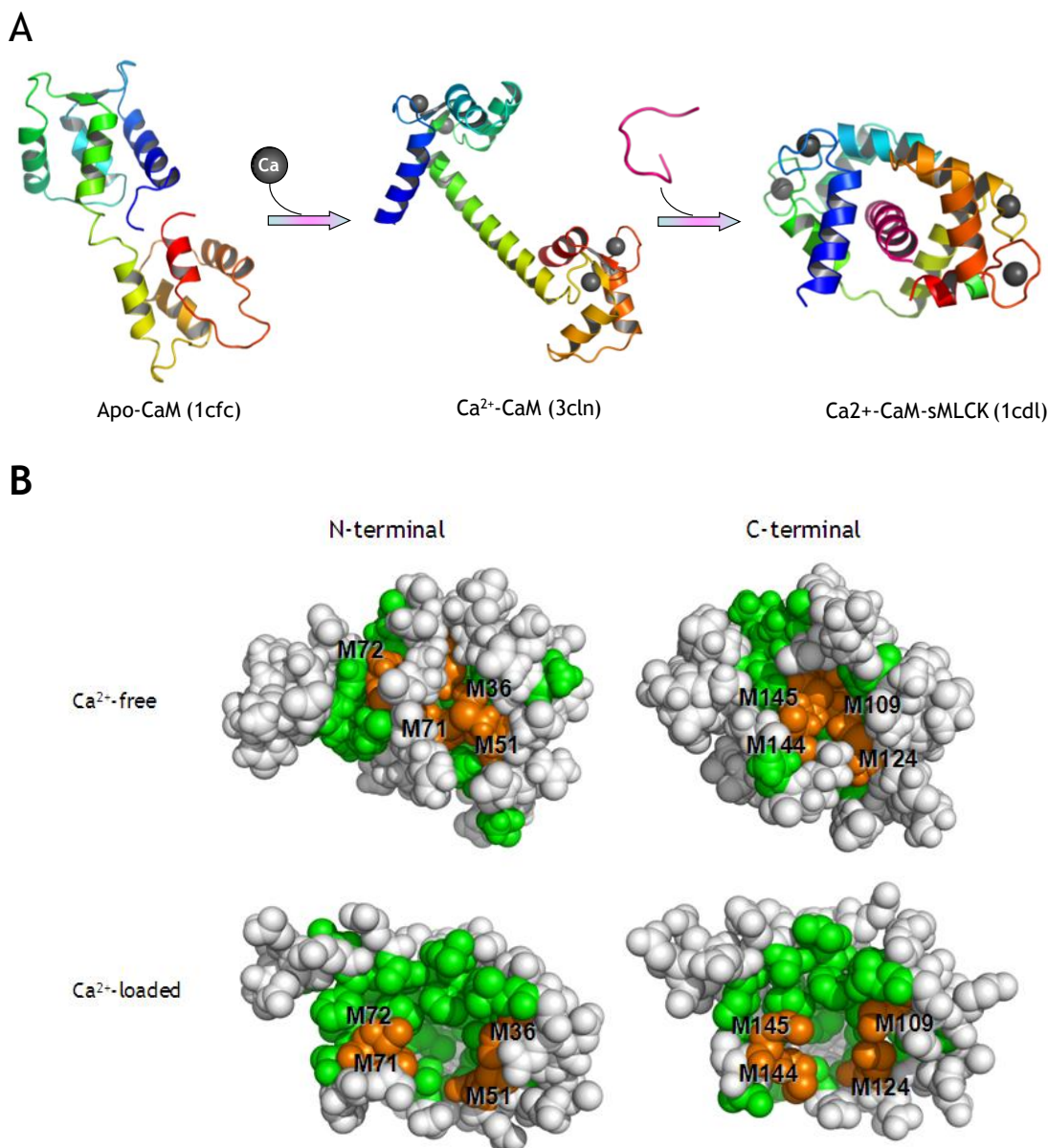


Figure 6.3. The Ca²⁺-induced conformational change and binding of a target peptide from chicken smooth muscle myosin light chain kinase to Ca²⁺-CaM. (A), The apo-CaM is comprised of two four-helix bundles linked by an interrupted central helical region. The binding of Ca²⁺ induces considerable changes in the interhelical angles and results in the exposure of a number of hydrophobic residues to the solvent. This exposed hydrophobic area is responsible for the interaction of CaM with its target proteins or peptides, such as the chicken smooth muscle light chain kinase (smMLCK). (B), Space-filling representation of distribution of hydrophobic residues (green) and methionine (orange) residue in Ca²⁺-free and Ca²⁺-ligated CaM. The binding of Ca²⁺ induced a conformational change to expose the hydrophobic residues toward the solvent, enabling its broad interaction with targets.

has thus far spurred the determination of more than 30 structures of CaM-target complexes and new modes of interaction have been discovered in recent years (see Table 6.2). At least three modes of CaM-target interaction have thus far been characterized.

The embracing mode

The embracing mode or “wrap-around” mode is well characterized as exemplified by the Ca^{2+} -CaM-sMLCK complex (419,420). Upon binding to CaM, the peptide with 1-8-14 spacing pattern of bulky hydrophobic residues and interspersed basic residues adopts an α -helical structure and is engulfed into the hydrophobic cleft of Ca^{2+} -CaM. The hydrophobic interaction is the main force driving the interaction of Ca^{2+} -CaM with its targets, whereas the electrostatic interactions between the positively charged residues of the peptide and the negatively charged clusters in both N- and C-terminal domains of CaM also contribute considerably to the interaction. The hydrophobic interaction is maximized by the unwinding of the central helix, which subsequently bends by $\sim 100^\circ$ and undergoes a $\sim 120^\circ$ rotation (Fig. 6.3A). As opposed to the dumbbell shape of Ca^{2+} -CaM, the CaM-target complex is globular with the N- and C-terminal domains coming close. The originally unstructured peptide binds to CaM in an antiparallel orientation such that the C-terminal domain of CaM contacts with the hydrophobic residue at position 1 of the peptide, whereas the N-terminal domain interacts with residues at positions 14. A similar binding mode has also been seen in Ca^{2+} -CaM-CaMKII complex with 1-5-10 anchoring spacing and

Table 6.2. List of CaM-target complexes deposited in protein data bank.

Pdb code (resolution, Å)	Target (length, aa)	Mode	Anchor spacing	CaM:Target ratio and No. of bound Ca ²⁺	Technique	Year
1A29 (2.7)	Trifluoperazine	1	others	1:2, 4Ca ²⁺	X-ray	1998
1LIN (2.0)				1:4, 4Ca ²⁺		1996
1CTR (2.5)				1:2, 4Ca ²⁺		1998
1CDL (2.4)	smMLCK (20)	1	1-8-14	1:1, 4Ca ²⁺	X-ray	1992
2BBM	skMLCK(26)				NMR	1992
1CDM (2.0)	CaMKIIalpha (18)	1	1-5-10	1:1, 4Ca ²⁺	X-ray	1994
2BDW (1.8)	AID of CaMKII	1	1-5-10	1:1, 4Ca ²⁺	X-ray	2005
1CFF	Ca ²⁺ -pump (20)	2	1-8	1:1, 4Ca ²⁺	NMR	1999
1CKK	CaMKK (26)	1	1-16	1:1, 4Ca ²⁺	NMR	1999
1IQ5 (1.8)	(24)				X-ray	
1G4Y (1.6)	SK2 channel (81)	3	1-8-14	2:2, 2Ca ²⁺ (N)	X-ray	2001
1QX5				apo-CaM		2004
1IWQ (2.0)	MARCKS (18)	1	1-5-10	1:1, 4Ca ²⁺	X-ray	2003
1K90 (2.8)	anthrax edema factor (485)	2	1-8-14	1:1, 2Ca ²⁺ (C)	X-ray	2002
1K93 (3.0)	(485)					2002
1LVC (3.6)	(485)					2002
1XFZ (3.3)	(735)					2005
1L7Z (2.3)	myristoylated CAP-23/NAP-22 (6)	1	others	1:1, 4Ca ²⁺	X-ray	2003
1MUX	W7	1	others	1:1, 4Ca ²⁺	NMR	1998
1MXE (1.7)	CaMKI (25)	1	1-5-10	1:1, 4Ca ²⁺	X-ray	1998
1NIW (2.1)	NOS (19)	1	1-5-8-14	1:1, 4Ca ²⁺	X-ray	2003
1NWD	GAD (28)	1	others	1:2, 4Ca ²⁺	NMR	2003
1QIV (2.6)	DPD	1	others	1:2, 4Ca ²⁺	X-ray	2000
1QIW (2.3)						2000
1QS7 (1.8)	rs20 (19)	1	1-8-14	1:1, 4Ca ²⁺	X-ray	2003
1QTX (1.7)	(21)					2003
1VRK (1.9)	(21)					1999
1SY9	olfactory CNG channel (19)	1	1-8-14	1:1, 4Ca ²⁺	NMR	2005
1WRZ (2.0)	DAP kinase (19)	1	1-8-14	1:1, 4Ca ²⁺	X-ray	2006
1YR5 (1.7)	(19)					2006
1XA5	KAR-2 (bis-indol alkaloid)	1	others	1:1, 4Ca ²⁺	X-ray	2004
2F2O (2.2)	calcineurin (25)	1	1-5-8-14	2:2, 4Ca ²⁺	X-ray	2006
2F2P (2.6)	(25)					2006
2F3Y (1.5)	Ca _v 1.2 channel (18)	1	IQ	1:1, 4Ca ²⁺	X-ray	2005
2F3Z (1.6)	(18)					2005
2FOT (2.5)	alphaII-spectrin (23)	1	others	1:1, 4Ca ²⁺	X-ray	2006

Ca²⁺-CaM-CaMKK complex (421) with 1-16 spacing pattern (Table 6.2). Due to the spacing variation of anchoring residues, more unwinding of the central linker region in Ca²⁺-CaM-CaMKII is observed, further demonstrating the structural malleability of CaM. From the perspective of function, this mode of interaction enables CaM to displace the targeted “autoinhibitory domain” from the active site and activate the enzyme.

The tethering mode

The second mode of interaction has been seen in the CaM-Ca²⁺ pump complex (422). Instead of the normal spacing of 1-8-14, certain isoforms of Ca²⁺-pump lack the bulky hydrophobic residue at position 14 (Table 6.2). As a result, the fragment (C20W) corresponding to the CaM binding site at the N-terminal of Ca²⁺ pump binds solely to the C-terminus of CaM and no contact between the two domains is observed. Here we use the term “tethering” to describe the binding mode since the peptide physically adheres loosely to one of the two lobes and is more exposed to the solvent than the MLCK embraced by CaM. In contrast to the collapsed globular shape of CaM-MLCK complex, CaM-Ca²⁺ pump complex remains a unique relatively extended form.

A more close-to-reality CaM-target complex structure is determined with the full sequence of anthrax edema factor (EF) (423). Though the anchoring of the CaM-binding site of EF follows the 1-8-14 spacing pattern, the extended structure of this complex does not resemble the collapsed structure typically observed in the binding mode 1. The N- and C-terminal domains of CaM in the

complex represent strikingly distinct states. The N-terminal domain remains in a close conformation without the binding of Ca^{2+} and interacts with the EF catalytic domain. In contrast, the C-terminal domain is open and binds two Ca^{2+} . The constitutive binding of CaM to the “low affinity” N-terminal domain is speculated to stabilize EF, making it more resistant to enzymatic cleavage. Meanwhile, the sole binding of two Ca^{2+} to the C-terminal domain suggests that EF may sense the changes of cellular $[\text{Ca}^{2+}]$ ahead of host CaM-activated enzymes. Overall, the enzymatic activity of EF has been enhanced by over 100 folds with the binding of CaM through remodeling of the active site.

Extensive biochemical and biophysical studies have revealed that CaM under Ca^{2+} -free condition can interact with a number of targets, such as myosin and neuromodulin that possess characteristic “IQ” motifs (424). Given the well-hidden hydrophobic core in apo-CaM, it is less likely for the interaction of apo-CaM with its targets to adopt an embracing mode that requires a large area of hydrophobic interaction. From the observation of the interaction between apo-CaM and the neuromodulin peptide, it has been suggested that the binding of peptide causes significantly fewer chemical shift changes in the N-terminal domain than in the C-terminal domain, which could be indicative of binding of the peptide to the C-lobe of CaM. In view of this, it is highly possible that the apo-CaM-target interaction will follow a similar binding mode as that of the CaM- Ca^{2+} pump complex. High-resolution structures featuring the interaction of apo-CaM with IQ motifs are needed to elucidate the true binding mode.

The oligomerizing mode

CaM-induced oligomerization has been best exemplified by the interaction of CaM with the small conductance K^+ (SK) channel (425) and the glutamate decarboxylase (GAD) fragment (426). CaM is constitutively associated with SK channel through its Ca^{2+} -free C-terminal domain. The binding of Ca^{2+} to its “low affinity” N-terminus initiates the dimerization of SK channel and ends up forming a 2:2 complex. Given the fact that the half-maximal activation of SK channel requires a $[Ca^{2+}]_i$ of $\sim 5 \times 10^{-7}$, it is intriguing to ask why the sensing of Ca^{2+} is carried out by CaM’s low affinity (10^{-6} M) domain instead of its C-terminal high affinity Ca^{2+} -binding sites (10^{-7} M). CaM-induced dimerization is also observed in its interaction with glutamate decarboxylase with the formation of 1:2 ([CaM]/[GA]) complex. The dimerized peptide, in this structure, is still embraced by Ca^{2+} -CaM.

6.1.4. Prediction and screening of potential CaM target sequences

With the vast amount of genomic information available, new CaM-binding sequences are reported each year. This raises the question as to whether the current known CaM-binding motifs are just the “tip of the iceberg”, or if we are closer to a final appreciation of the CaM-target interaction network. Development of fast and reliable strategies to predict and validate potential CaM-binding site would undoubtedly be of great help to visualize the macroscopic picture.

Though low sequence identity has been observed among different target sequences, all the target peptides (18mer to 26mer) share some key features: the hydrophobic residues with bulky sidechains that anchor the target sequence to the hydrophobic cleft of CaM are arranged in particular spacing patterns with

2-6 positively charged residues interspersed in between. According to the distribution of anchoring residues in the continuous polypeptide sequence, the binding motifs can be divided into five classes: 1-5-10 class, 1-8-14 class, 1-16 class, IQ class and others. Among them, the first three classes have been found to interact with CaM in a Ca^{2+} -dependent way in most cases. The IQ class motifs, however, are activated or inhibited by CaM in a Ca^{2+} -independent fashion and CaM seems to be a constitutive part of its functional unit. In addition, there are several unclassified CaM binding motifs that interact with CaM when the four Ca^{2+} -binding sites are only half occupied. All of these common features enable scientists to gain more confidence when confronted by the following two questions: 1) Are the proteins under investigation capable of interacting with CaM, and if so, 2) is it Ca^{2+} -dependent or independent? An online server that is conveniently accessible for scientists to predict calmodulin binding sites has been generated by Ikura *et al* (90). Nevertheless, pure *in silico* work could only serve as guidance for the research direction and always brings about researchers' concerns on its accuracy and accountability. Therefore, several groups have worked to develop robust methods to screen CaM-binding proteins (CaMBPs) on the proteome-wide scale.

A CaM binding overlay (CAMBOT) technique that requires labeling of CaM with ^{35}S and the use of SDS-PAGE was established by O'Day *et al* (427). In this technique, the bacteria-expressed ^{35}S -CaM was used as the probe to detect CaMBPs and compare the CaMBPs expression level under different conditions. Under Ca^{2+} -saturated and Ca^{2+} -free conditions, it is possible, in theory, to detect

both Ca^{2+} -dependent and Ca^{2+} -independent CaMBPs. This method can be potentially applied to verify and profile the CaMBPs, to assess the effects of drugs on CaM-CaMBP networks, and to evaluate the effects of heavy metal ion present in the environment on the CaM-CaMBP network. However, this method requires the handling of radioactive products and involves relatively labor-intensive and time-consuming procedures. In addition, the amount of proteins encoded by rare genes is far from sufficient to be detected by SDS-PAGE. An alternate way to screen CaMBP is explored by Shen *et al* (428-430). This large-scale screening method utilizes the mRNA display technique. mRNA-displayed enriched proteome library can be selected by biotinylated CaM, which will be further captured by streptavidin-agarose beads. Another round of selection involving six steps (PCR, in vitro transcription, crosslinking, cell-free translation and fusion, reverse transcription and affinity chromatography) can be repeated to diminish unspecific interactions. Screening of CaMBPs in the human proteome with this method reveals a large body of previous undetected proteins, such as ribosomal proteins, proteasome 26S subunits, deubiquitinating enzymes, and leucine zipper proteins.

The above-mentioned recurring binding modes underlying the interaction of CaM can surely be applied to some of the newly detected CaMBPs. However, with such diversity in CaM's target recognition conferred by its conformational plasticity and structural malleability, we could highly expect the emergence of new and surprising modes.

6.2. Gap junction and calmodulin

Gap junctions comprise the intercellular channels that mediate the cell-to-cell transfer of small molecules including metabolites, second messengers and ions in mammalian cells (431). Gap junctions are composed of two hemichannels (termed connexons) each comprised of six connexin (Cx) subunits. The connexon from one cell joins in mirror symmetry with another connexon of the apposing cell. To date, at least 20 connexin genes have been identified in the human genome with connexin43 (Cx43) and the most ubiquitous connexin (432).

Gap junctions have been shown to be regulated by intracellular Ca^{2+} concentration ($[\text{Ca}^{2+}]_i$) (433) that Peracchia and others have shown to be likely mediated by CaM interacting directly with the connexin proteins (434,435). We have shown previously that cell-to-cell communication in lens epithelial cell cultures is inhibited by elevated $[\text{Ca}^{2+}]_i$. Specifically, cell-to-cell transfer of the fluorescent dye AlexaFluor594 was half-maximally inhibited at ~300 nM $[\text{Ca}^{2+}]_i$ in lens cell cultures and this inhibition was prevented by preincubation of these cultures with CaM antagonists (75,436), consistent with earlier reports that elevated $[\text{Ca}^{2+}]_i$ increased internal electrical resistance in the lens that was prevented by preincubation with CaM antagonists (436). Indeed, this action of Ca^{2+} in lens cell cultures is due in part to the inhibition of Cx43, the major connexin in these cell cultures. It has been demonstrated that Cx43-transfected HeLa cells exhibiting a similar Ca^{2+} -dependent inhibition appears to be CaM-mediated (68). The rapid onset of this Ca^{2+} -dependent inhibition of cell-to-cell communication (within seconds) suggests that this is mediated by a direct

interaction of CaM with the connexin protein rather than being mediated via the action of a CaM-dependent protein kinase. Indeed, the Ca^{2+} -dependent binding of CaM to rat Cx32 (435), fish Cx35, mouse Cx36 (69) and Cx50 (437) have been reported. Two cytoplasmic CaM binding domains, with one site ($K_d = 27 \text{ nM}$) in the N-terminus and the other site ($K_d = 1.2 \text{ }\mu\text{M}$) in the C-terminal region have been identified in Cx32 (435), whereas a single CaM binding site ($K_d: 11\text{-}72 \text{ nM}$) was identified in the C-terminus of Cx35 and Cx36 (69).

Each Cx43 or Cx44 monomer consists of four highly-conserved transmembrane segments, a short N-terminal cytoplasmic region, one intracellular and two extracellular loops, as well as a C-terminal tail. (Fig. 6.4). Variability in sequence homology across different connexin types is greatest in the intracellular loop and C-terminus. Efforts to map the potential CaM binding sites in Cx43 have led to conflicting results. Torok *et al*/ reported the binding of the fluorescent CaM derivative (TA-CaM) to the N-terminal (aa 1-16) region of Cx43 with a dissociation constant of $1.2 \text{ }\mu\text{M}$ (435). However, Duffy *et al*/ were unable to detect any interaction of CaM with a peptide spanning the first 21 amino acids of the N-terminus (438). Both groups, nevertheless, failed to detect the binding of CaM to the peptides derived from the C-terminal tail (aa 314-325; aa 336-350; aa 346-360) or the intracellular loop (aa 95-114; aa 123-136; aa 119-144). They suggest that the interaction of CaM with Cx43 might occur via other regions of the intracellular loop. Interestingly, by using the CaM binding database server (90), we predicted a potential CaM binding site with high predictive score in the second half of the intracellular loop of Cx43 (aa 136-158) or Cx44 (aa 129-150)

that has not been tested before. In the present study, we have applied a variety of biophysical approaches to examine the binding of CaM to this sequence. Our findings strongly suggest a Ca^{2+} -dependent interaction between the CaM and the intracellular loop region of Cx43 or Cx44.

6.2.1. Prediction of calmodulin binding sites in connexins

Using a search program for CaM-binding sites (90), a putative CaM binding region with high predictive score was identified in the intracellular loop of Cx43 (Ac-¹³⁶KYGIEEHGKVKMRGGLRTYIIS¹⁵⁸-NH₂) or Cx44 (Ac-¹²⁹VRDDRGGKVRIAGALLRTYVFNI¹⁵⁰-NH₂) (Fig. 6.4). In addition, two other stretches of sequences with lower predictive scores, one in the N-terminal part of the intracellular loop region (Ac-⁸⁶SVPTLLYLAHVIFYVMRKEEKLN¹⁰⁷-NH₂) and the other in the C terminus region (Ac-²²⁴NIIELFYVFFKGVKDRVKGSRSDPY²⁴⁷-NH₂), were detected in Cx43 but not in Cx44. A common feature of the CaM binding regions is a number of hydrophobic residues interspersed with positively charged residues (≥ 3), typically forming a 1-5-10, 1-8-14 or 1-16 pattern of hydrophobic residues (90,439). These regions are normally comprised of 12-30 continuous amino acids with a strong propensity to form an amphipathic helical structure (90,439). The putative CaM binding region of Cx43 exhibits a 1-5-10 hydrophobic residue motif as reported in other well-characterized CaM targeting proteins such as CaM Kinase I (440), CaM Kinase II (441), MARCKS (442) and synapsin (443). The randomized control peptide (Ac-LGGEYLVTMESKIHKGKRIGYR-NH₂) with the same composition of amino acids

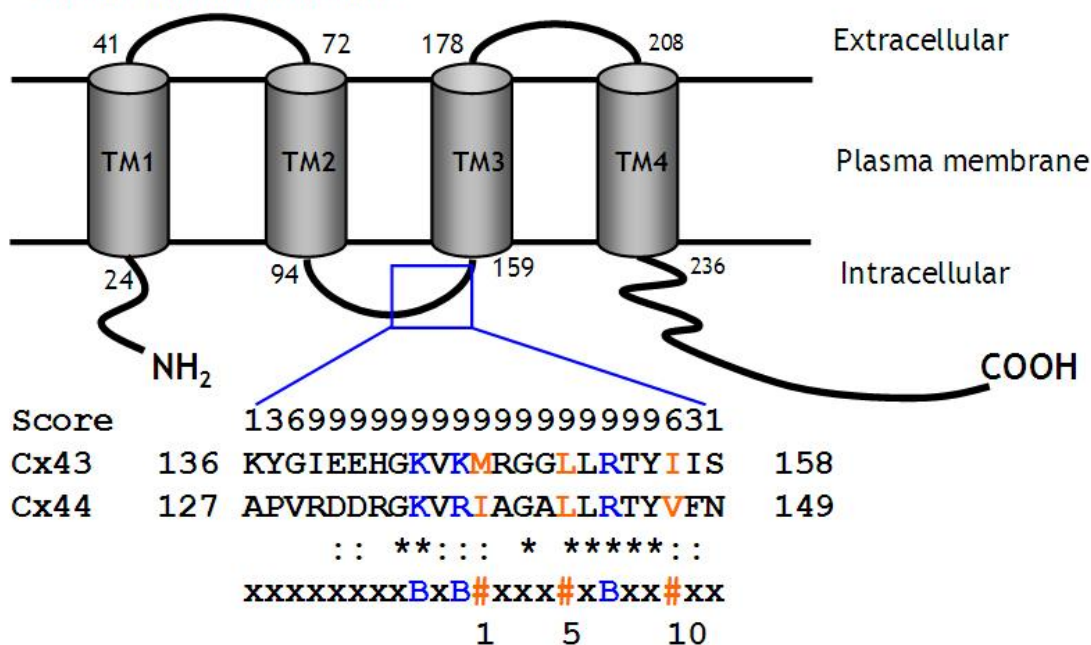


Figure 6.4. Membrane topology and predicted CaM binding sequences in Cx43 and Cx44. Primary sequence and transmembrane (TM) topology of Cx43. The integral membrane protein Cx is composed of four TM regions, two extracellular loops, one cytoplasmic loop, a short N-terminus, and a longer C-terminal tail. The predicted high-affinity CaM binding site (aa 136-158 in Cx43 and aa 129-150 in Cx44) of highest predicted score is located in the second half of the intracellular loop between TM2 and TM3. Two other regions (boxes I and III) with lower predictive scores are also identified. The numeric score ranges from 1-9, representing the probability of an accurate prediction of a high affinity CaM binding site. #, hydrophobic residues. :, highly conserved residues.

as the Cx43 peptide but arranged in a different order with no predicted CaM binding capacity is used in the study as control.

The ability of the peptides Cx43₁₃₆₋₁₅₈ or Cx44₁₂₉₋₁₅₀ to form an α -helical structure was evaluated by monitoring its secondary structure using far UV CD in the presence of varying concentrations of trifluoroethanol (TFE) (Fig.6.5). In aqueous solution, the free peptides were largely unstructured with a negative maximum at 197 nm. However, when the TFE concentration was increased to 20% it started to form α -helical structure (12%), with two major troughs at 208 nm and 222 nm observed in the CD spectra. The helical contents increased up to 50% in the presence of 40% TFE (Fig. 6.5). TFE, with a dielectric constant one third that of water, is capable of strengthening intramolecular hydrogen bonding (444), and is known to induce and stabilize the intrinsic secondary structures in peptides, possibly by mimicking the hydrophobic environment of the peptide in the intact proteins (445-447). Several CaM binding peptides have been reported to adopt similar α -helical structures in TFE solvent as in the CaM-peptide complexes (448,449). These results support our prediction that this Cx43₁₃₆₋₁₅₈ peptide possesses a strong α -helical-forming propensity.

6.2.2. Secondary and tertiary structure change induced by peptide binding

As described above, the binding of CaM to CaM-binding peptides typically induces the formation of α -helical structure in these peptides (420). As seen in Fig. 6.6A and Fig. 6.7A, the addition of a 1:1 molar equivalent of peptides to Ca²⁺-CaM results in an approximately 10% more negative signal in the spectrum.

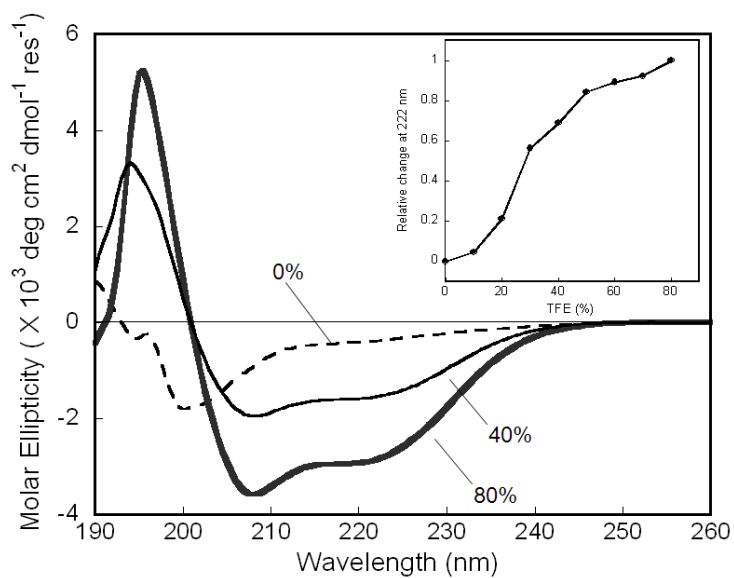
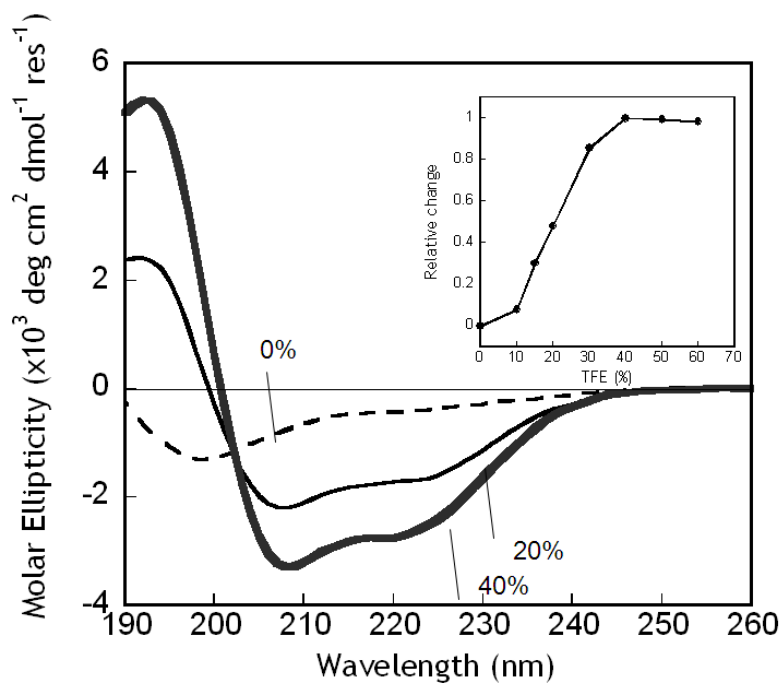
A**B**

Figure 6.5. Far UV CD spectra of the peptides Cx43₁₃₆₋₁₅₈ (A) and Cx44₁₂₉₋₁₅₀ (B) with different concentrations of TFE. The *inset* shows the changes of molar ellipticity at 222 nm and 208 nm as a function of TFE concentration.

Since the α -helical content of CaM typically does not increase upon peptide binding (420), the observed net increase in the CD signal could be reasonably attributed to CaM-bound Cx peptides. The difference spectrum, obtained by subtracting the Ca^{2+} -CaM spectrum from the Ca^{2+} -CaM-peptide spectrum, showed two major troughs at 208 nm and 222 nm, which is strikingly different from the random coiled structure of the peptide alone (Fig. 6.6B and Fig. 6.7B). Deconvolution of this difference spectra revealed that the CaM-bound peptides had ~30-40% helical structure, which is comparable to the helical content of this peptide in 20%-40% TFE (Fig. 6.5). The Cx43₁₃₆₋₁₅₈ binding-induced CD signal change enabled us to measure the peptide binding affinity of CaM. As shown in the inset of Fig. 6.6A, the signal changes reach the maxima when the Cx43₁₃₆₋₁₅₈:CaM ratio was approximately 1:1 and the fitting curves provided an apparent K_d of 100 ± 20 nM ($n = 3$) in 10 mM KCl and 750 ± 120 (n = 3) nM in 100 mM KCl (Table 6.3) in the presence of 1 mM Ca^{2+} . The decrease in K_d at the higher salt concentration suggests that electrostatic interactions might play an important role in the interaction of Cx43₁₃₆₋₁₅₈ with CaM. In the presence of 1 mM EGTA, no significant difference in the far UV CD spectra was detected after the addition of a 1:1 molar equivalent of Cx43₁₃₆₋₁₅₈ (Fig. 6.6A) or Cx44₁₂₉₋₁₅₀ (Fig. 6.7A) to CaM suggesting that the peptide is unable to interact with CaM or the binding is much weaker in the absence of Ca^{2+} .

Near UV CD spectroscopy was conducted to examine the tertiary packing around aromatic residues of CaM, Cx peptides and the CaM-Cx peptide complexes (Fig. 6.6C and Fig. 6.7C). Both Cx43₁₃₆₋₁₅₈ and Cx44₁₂₉₋₁₅₀ exhibited

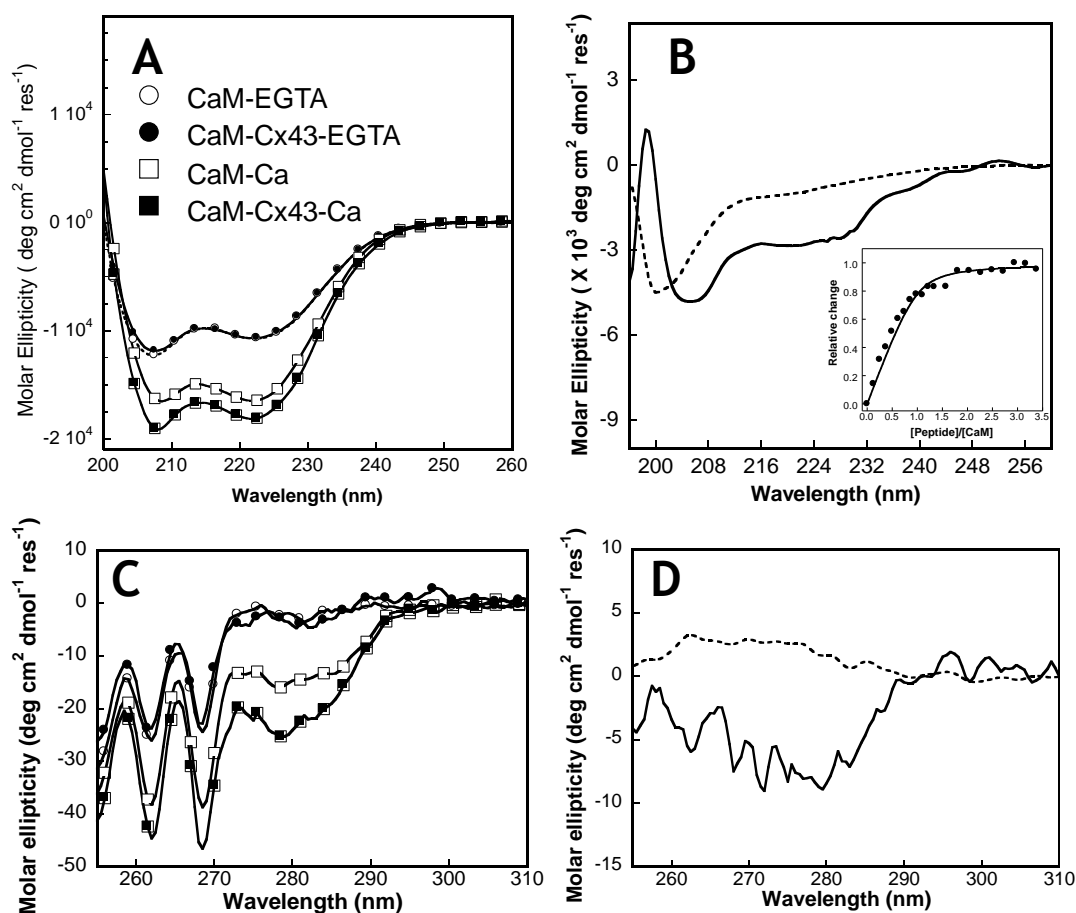


Figure 6.6. Circular Dichroism studies of the interaction between Cx43₁₃₆₋₁₅₈ and CaM. (A), Far UV Circular Dichroism spectrum of CaM in the presence of 1 mM EGTA (○) or 1 mM CaCl₂ (□), and a 1:1 CaM-synthetic peptide mixture with 1 mM EGTA (●) or 1 mM CaCl₂ (■) after subtracting the contribution from buffer and added peptides. (B), The far UV circular dichroism spectra of Cx43₁₃₆₋₁₅₈ (*dashed line*) and the calculated difference spectrum (*solid line*) by subtracting the spectrum of Ca²⁺-CaM from that of the Ca²⁺-CaM-Cx43₁₃₆₋₁₅₈ mixture with 1 mM Ca²⁺ in a buffer consisting of 100 mM KCl, 10 mM Tris-HCl, pH 7.4. The *inset* reports the relative change of the circular dichroism molar ellipticity at 222 nm as a function of synthetic peptide concentration. (C), Near UV circular dichroism spectra of CaM in the presence of 1 mM EGTA (○) or 1 mM CaCl₂ (□), and 1:1 CaM- Cx43₁₃₆₋₁₅₈ mixture with 1 mM EGTA (●) or 1 mM CaCl₂ (■) after subtracting the contribution from buffer and added peptides. (D), The near UV circular dichroism spectra of Cx43₁₃₆₋₁₅₈ (*dashed line*) and the calculated difference spectrum by subtracting the Ca²⁺-CaM spectrum from the Ca²⁺-CaM- Cx43₁₃₆₋₁₅₈ spectrum (*solid line*) in a buffer consisting of 1 mM Ca²⁺, 100 mM KCl, 10 mM Tris-HCl, pH 7.4.

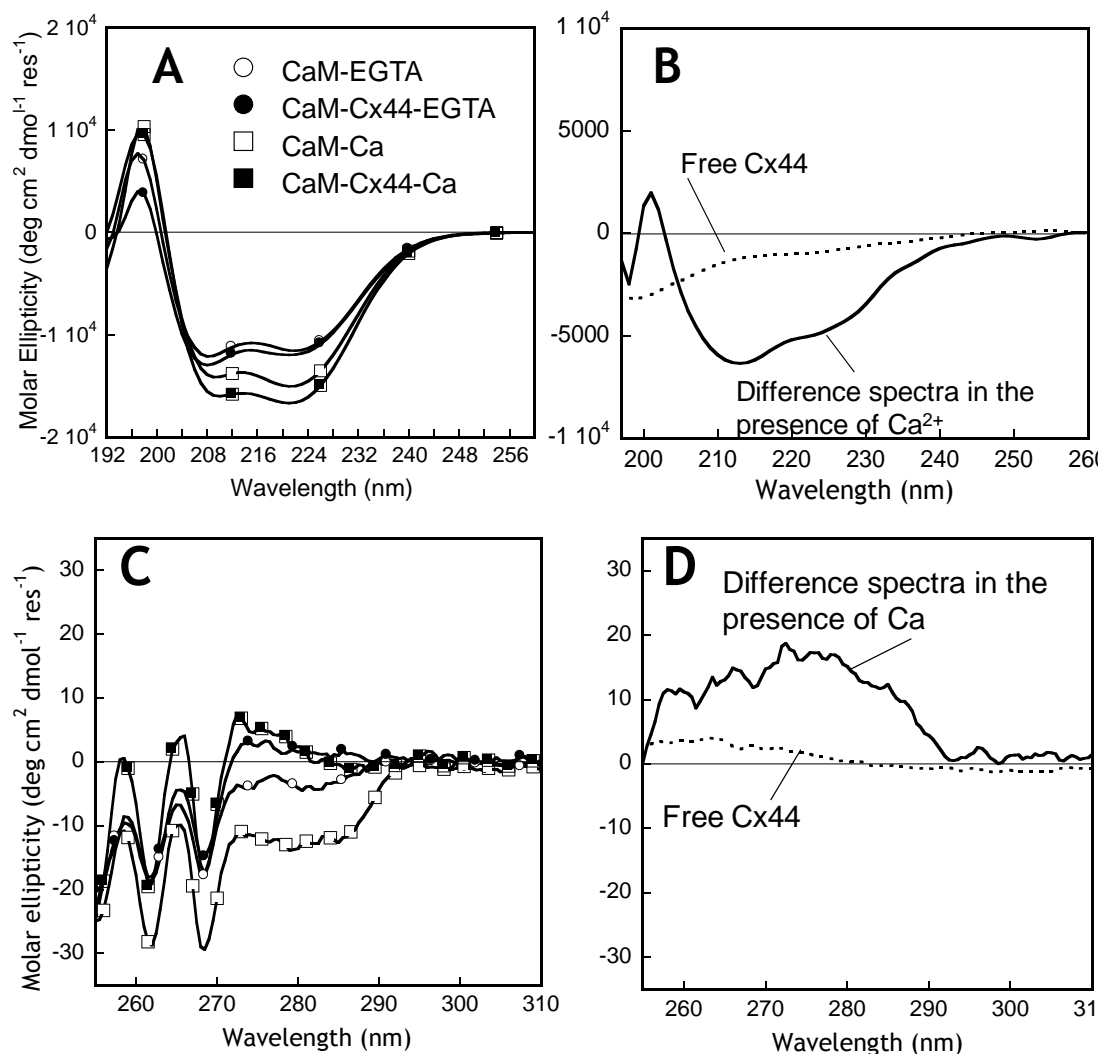


Figure 6.7. Circular Dichroism studies of the interaction between Cx44₁₂₉₋₁₅₀ and CaM. (A), Far UV Circular Dichroism spectrum of CaM in the presence of 1 mM EGTA (○) or 1 mM CaCl₂ (□), and a 1:1 CaM-synthetic peptide mixture with 1 mM EGTA (●) or 1mM CaCl₂ (■) after subtracting the contribution from buffer and added peptides. (B), The far UV circular dichroism spectra of Cx44₁₂₉₋₁₅₀ (*dashed line*) and the calculated difference spectrum (*solid line*) by subtracting the spectrum of Ca²⁺-CaM from that of the Ca²⁺-CaM-Cx44₁₂₉₋₁₅₀ mixture with 1 mM Ca²⁺ in a buffer consisting of 100 mM KCl, 10 mM Tris-HCl, pH 7.4. (C), Near UV circular dichroism spectra of CaM in the presence of 1 mM EGTA (○) or 1 mM CaCl₂ (□), and 1:1 CaM- Cx44₁₂₉₋₁₅₀ mixture with 1 mM EGTA (●) or 1mM CaCl₂ (■) after subtracting the contribution from buffer and added peptides. (D), The near UV circular dichroism spectra of Cx44₁₂₉₋₁₅₀ (*dashed line*) and the calculated difference spectrum by subtracting the Ca²⁺-CaM spectrum from the Ca²⁺-CaM- Cx44₁₂₉₋₁₅₀ spectrum (*solid line*) in a buffer consisting of 1 mM Ca²⁺, 100 mM KCl, 10 mM Tris-HCl, pH 7.4.

negligible signal in aqueous solution by itself, but the addition of Ca^{2+} -CaM to Cx peptides resulted in the appearance of more negative (Cx43₁₃₆₋₁₅₈) or more positive (Cx44₁₂₉₋₁₅₀) ellipticity. Both CaM and the synthetic peptides do not contain any tryptophan but each contains two tyrosines. The prominent band above 275 nm is likely due to the tyrosine residues in the C-terminus region of CaM (Y⁹⁹ and Y¹³⁸), while the two peaks at 262 nm and 268 nm are likely due to the nine phenylalanine residues in CaM (450). The difference spectrum obtained by subtracting the Ca^{2+} -CaM spectrum from that of Ca^{2+} -CaM-Cx43₁₃₆₋₁₅₈ (Fig. 6.6D) or Ca^{2+} -CaM-Cx44₁₂₉₋₁₅₀ (Fig. 6.7D) showed a prominent band between 275 and 279 nm, which is most likely due to the immobilization of tyrosine residues in the Cx peptides (which contains no tryptophan) following its binding to CaM. The different near UV difference spectra (Figs. 6.6D and 6.7D) could arise from the difference in the local environment around the Try residue(s) in these two peptides. The contribution of signals from the CaM tyrosine residue is less likely to be dominant although both CaM and Cx peptides contain two tyrosine residues making it difficult to unambiguously differentiate the contribution from each macromolecule in the spectra. The Y⁹⁹ and Y¹³⁸ in CaM are located in the C-terminus EF-loops where the local structures of the Ca^{2+} -bound EF-loops have not been reported to be altered upon peptide binding in the currently known CaM-peptide complex despite the rearrangement of tertiary structures. Thus, the local environment of the CaM tyrosine is less likely to be significantly altered upon Cx peptides binding. No significant differences were observed between the spectra of apo-CaM and the 1:1 mixture of apo-CaM:Cx peptide (Fig. 6.6C and

Fig. 6.7C), further supporting the Ca^{2+} -dependent binding of Cx peptide to Ca^{2+} -CaM and the resulting conformational changes in Cx peptides revealed by the far UV CD studies.

6.2.3. Monitoring the CaM-Cx interaction by SPR

The binding of Cx43₁₃₆₋₁₅₈ to CaM was further confirmed using surface plasmon resonance spectroscopy with CaM immobilized on the CM5 sensor chip. Fig. 6.8 shows the sensorgrams of successive peptide binding experiments. The peptide Cx43₁₃₆₋₁₅₈ rapidly associated with CaM in the presence of Ca^{2+} . At the end of peptide injection, dissociation of the complex was observed during flushing of the chip with peptide-free running buffer. Analysis of the kinetic data inferred that the apparent K_d for the binding of Cx43₁₃₆₋₁₅₈ to CaM is approximately 1 μM with a k_{off} rate of $5 \times 10^{-4} \text{ s}^{-1}$ in the presence of saturating amount of Ca^{2+} and 100 mM KCl. In contrast, no significant binding of Cx43₁₃₆₋₁₅₈ to CaM was detected in the presence of EGTA. Furthermore, the control randomized Cx43₁₃₆₋₁₅₈ peptide that contained the same composition of amino acids as Cx43₁₃₆₋₁₅₈ did not exhibit specific binding in the presence of either Ca^{2+} or EGTA; the response units of these injections were not significantly higher than that of injection of the peptide through the underivatized CM5 chip surface. These results further confirm the Ca^{2+} -dependent binding and specificity of Cx43₁₃₆₋₁₅₈ binding to CaM.

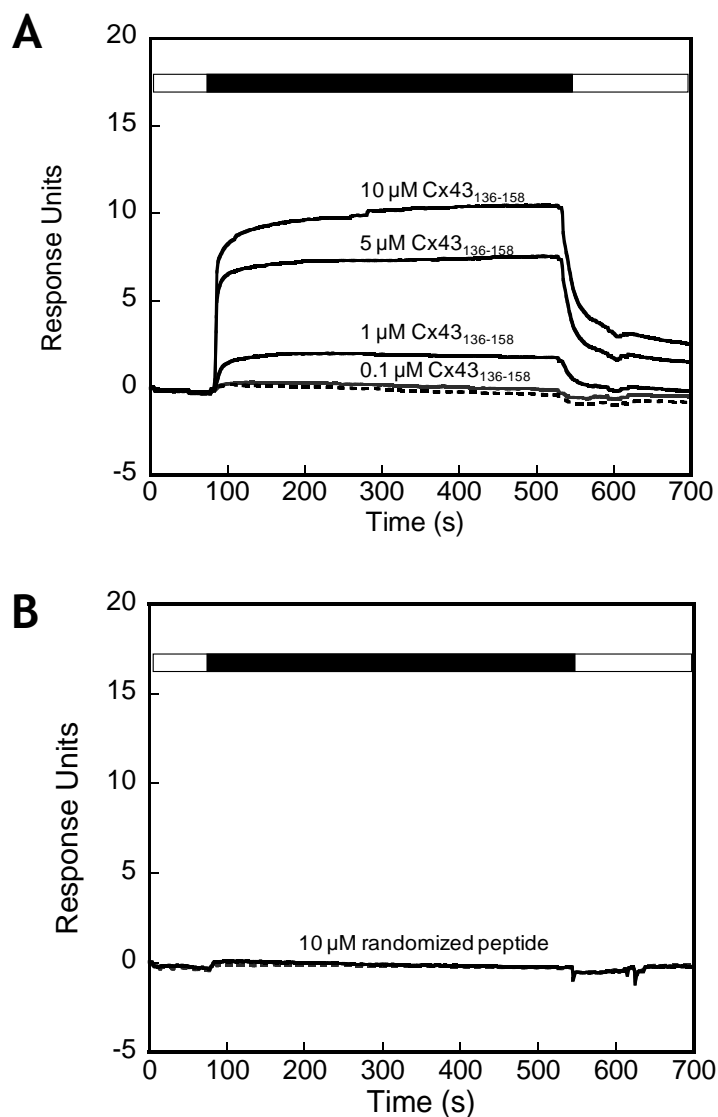


Figure 6.8. Monitoring the interaction of CaM with Cx43₁₃₆₋₁₅₈ by surface plasmon resonance spectroscopy. The sensorgrams represent the binding of 0.1, 1, 5 and 10 μM Cx43₁₃₆₋₁₅₈ (A) or 10 μM of the randomized Cx43 peptide (B) to the CaM-immobilized chip in the presence of 5 mM Ca²⁺ (solid line) or EGTA (dashed line). CaM was immobilized to the sensor chip CM with a response unit of approximately 3500. The flow rate was set at 5 μl/min. During recording, running buffer contained 5 mM CaCl₂ or 5 mM EGTA in 100 mM KCl, 50 mM Tris-HCl at pH 7.4 or the same buffer plus the peptides (solid bar). Subsequently, peptide-free buffer was injected to monitor the dissociation process.

6.2.4. Structural changes induced by Cx binding

Next, it was important to demonstrate that there is in fact a direct interaction between the Cx peptides and CaM. This was confirmed using high resolution NMR. As shown in Fig. 6.9A-B, in the presence of Ca^{2+} , a number of dispersed peaks of CaM underwent chemical shifts of greater than 0.05 ppm following the addition of Cx43₁₃₆₋₁₅₈ or Cx44₁₂₉₋₁₅₀. In contrast, only insignificant changes in several residues were observed upon the addition of Cx43₁₃₆₋₁₅₈ to CaM in the presence of 10 mM EGTA (Fig. 6.9C), again supporting the absence of any of interaction at physiological concentrations of K^+ in the absence of Ca^{2+} . In addition, the addition of the control randomized Cx43₁₃₆₋₁₅₈ peptide resulted in only very minor changes in the CaM ^{15}N - ^1H HSQC spectra even in the presence of Ca^{2+} (Fig. 6.9A, blue). This latter result would appear to exclude the possibility of non-specific binding between the randomized peptide and CaM arising from the charge or hydrophobic interactions of certain residues, indicating the necessity of the specific arrangement of the Cx43 or Cx44 peptide sequences.

The NMR chemical shift resonances of Ca^{2+} -CaM have been assigned by several groups (420,451,452). By following their movements during titration of CaM with Cx43₁₃₆₋₁₅₈ peptide, the dispersed peaks of the peptide-CaM complex have been unambiguously assigned. CaM has been known to activate downstream proteins by displacing autoinhibitory domains, remodeling the active sites, or inducing oligomerization in these proteins (310). The binding of CaM to target proteins and peptides has been reported to involve the N- and/or C-terminal domains of CaM (14,310,420,422). Analysis of the effects of peptide

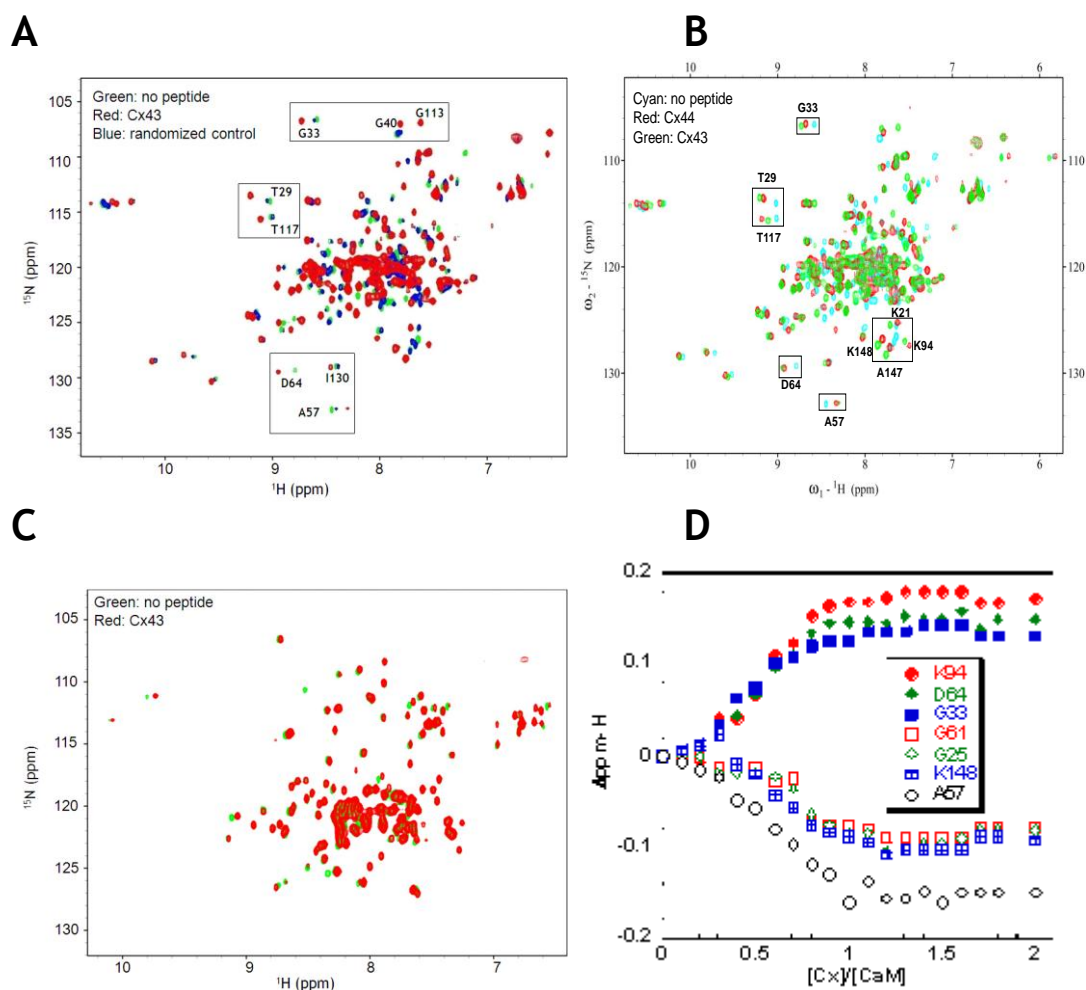


Figure 6.9. Interaction between CaM and Cx peptides revealed by NMR. (A), Overlaid HSQC spectra of Ca^{2+} -CaM (green) with Ca^{2+} -CaM-Cx43₁₃₆₋₁₅₈ (red) or Ca^{2+} -CaM:control randomized Cx43₁₃₆₋₁₅₈ peptide (blue). A subset of assigned peaks displaying significant movement upon peptide binding were framed in boxes. (B), Overlaid HSQC spectra of Ca^{2+} -CaM (cyan) with Ca^{2+} -CaM-Cx43₁₃₆₋₁₅₈ (green) or Ca^{2+} -CaM-Cx44₁₂₉₋₁₅₀ (red). (C), Overlay of HSQC spectrum of apo-CaM (green) with apo-CaM-Cx43₁₃₆₋₁₅₈ (red). (D), Changes in chemical shift of selected well-dispersed resonances as a function of the Cx₁₃₆₋₁₅₈: Ca^{2+} -CaM ratio in the presence of saturating Ca^{2+} at physiological K^+ condition.

binding on the backbone amide chemical shifts of CaM could shed light on the underlying mode of interaction of the Cx peptides with CaM. The resonances that underwent chemical shifts of greater than 0.05 ppm in the peptide titration were present in both the N- to C-termini of CaM, including residues G²⁵, T²⁹, G³³, G⁴⁰, A⁵⁷, G⁶¹, D⁶⁴, K⁷⁷, K⁹⁴, G⁹⁸, G¹¹³, T¹¹⁷, G¹³⁴, N¹³⁷, A¹⁴⁷ and K¹⁴⁸ and possibly others since the assignment of the overlapped regions is still in progress. The movements of the residues that we were able to assign during the titration followed a similar chemical shift trend (Fig. 6.9D), suggesting a single binding process. Furthermore, the chemical shifts saturated when the Cx43₁₃₆₋₁₅₈ or Cx44₁₂₉₋₁₅₀ to CaM ratio exceeded 1.0. These results demonstrated that the 1:1 Ca²⁺-dependent binding of Cx43₁₃₆₋₁₅₈ or Cx44₁₂₉₋₁₅₀ to CaM induces conformational changes in both the N- and C-terminal domains of CaM.

6.2.5. Peptide binding affinities and thermodynamics of the CaM-Cx interaction

The dansyl fluorescence emission maxima, at ~510 nm, are well-separated from any intrinsic fluorescence arising from aromatic amino acid residues. Because of this, as well as dansyl's great sensitivity to environmental changes, dansylated proteins including dansylated CaM have been widely used in studying the effect of protein-target interaction of ions, peptides, or drugs (65).

Binding of Cx43₁₃₆₋₁₅₈ to dansyl-CaM

As shown in Fig. 6.10A-B, Ca²⁺ binding to dansyl-CaM resulted in an increase of the fluorescence intensity and an emission blue-shift from 510 to 497

nm indicating that the dansyl group moved to a more hydrophobic environment. The addition of Cx43₁₃₆₋₁₅₈ to Ca²⁺-dansyl-CaM resulted in a further 80% increase in dansyl fluorescence intensity and a further emission blue-shift to 477 nm, suggesting that the dansyl groups are located in an even more hydrophobic environment in the complex. In the absence of Ca²⁺, the dansyl fluorescence remained nearly unaltered when Cx43₁₃₆₋₁₅₈ was added to dansyl-CaM, indicating a Ca²⁺-dependent interaction of Cx43₁₃₆₋₁₅₈ with dansyl-CaM. Titration data with Cx43₁₃₆₋₁₅₈ resulted in apparent dissociation constants of 240 ± 10 nM in 10 mM KCl and 860 ± 20 nM in 100 mM KCl (Table 6.3) assuming a 1:1 binding mode (Fig. 6.10A). In addition, the titration of dansyl-CaM with two other Cx43 peptides (Cx43₈₆₋₁₀₇ corresponding to residues 86-107 in Cx43 in the N-terminal region of the intracellular loop, and Cx43₂₂₄₋₂₄₇ corresponding to residues 224-247 in Cx43 in the C-terminus region) that have significantly lower predictive scores for CaM binding (Fig. 6.4) resulted in no significant fluorescence signal changes (Figs. 6.10C and D) indicating these peptides did not bind CaM.

Binding of Cx44₁₂₉₋₁₅₀ to dansyl-CaM

As shown in Figs. 6.10E and F, in the presence of EGTA or Ca²⁺, dansylated CaM exhibits its fluorescence maxima at 510 nm and 500 nm respectively. The addition of Cx44₁₂₉₋₁₅₀ to Ca²⁺-loaded dansyl-CaM resulted in ~50% increase in the dansyl fluorescence intensity and brought about 17 nm blue-shift of the fluorescence maximal peak (Inset, Fig. 6.10E), strongly suggesting that the dansyl groups are located in more hydrophobic environment

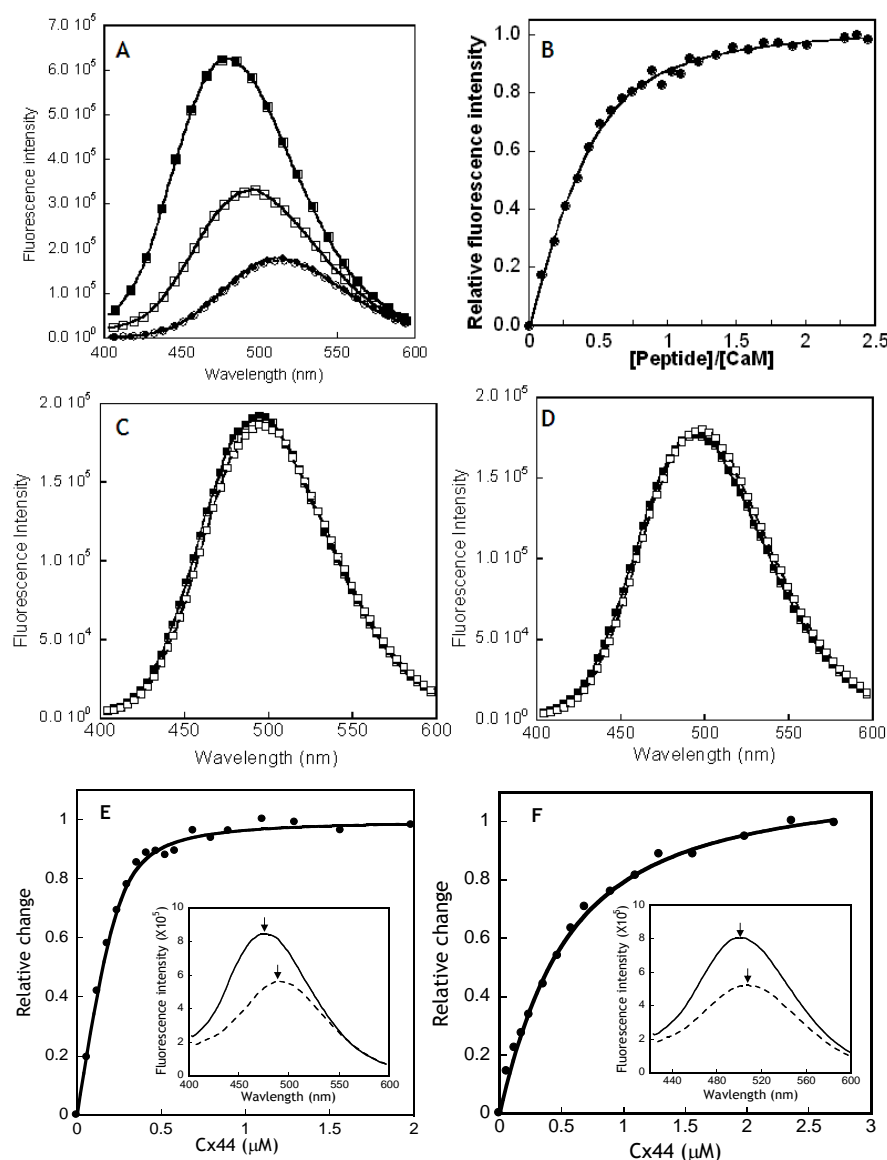


Figure 6.10. Interaction of Cx peptides with dansyl-CaM. (A) Dansyl fluorescence spectra of apo-dansyl-CaM in the absence (\circ) and presence (\bullet) of the Cx43₁₃₈₋₁₅₈ peptide, and Ca²⁺-dansyl-CaM in the absence (\square) and presence (\blacksquare) of the Cx43₁₃₈₋₁₅₈ peptide (2.5 μ M). (B), The dansyl fluorescence intensity plotted as a function of Cx43₁₃₈₋₁₅₈ peptide. (C), Dansyl fluorescence spectra of Ca²⁺-dansyl-CaM (1.25 μ M) in the absence (\square) and presence (\blacksquare) of the Cx43₈₈₋₁₀₇ peptide (2.5 μ M) in 1 mM CaCl₂. (D), Dansyl fluorescence spectra of Ca²⁺-dansyl-CaM (1.25 μ M) in the absence (\square) and presence (\blacksquare) of the Cx43₂₂₂₋₂₄₇ peptide (2.5 μ M) in 1 mM CaCl₂. (E-F), The dansyl fluorescence intensity plotted as a function of Cx44₁₂₉₋₁₅₀ peptide in the presence of Ca²⁺ or EGTA. The inset showed the dansyl fluorescence spectra of dansyl-CaM with (solid line) and without (dashed line) the addition of the Cx44₁₂₉₋₁₅₀ peptide with 1 mM CaCl₂ (E) or EGTA (F). The buffer consists of 50 mM Tris, 100 mM KCl, pH 7.4.

in the complex. During titration, the fluorescence intensity reached saturation when the peptide to CaM ratio was around 1:1. In the absence of Ca^{2+} , the dansyl fluorescence intensity also increased by ~50%, but the fluorescence maxima was only blue-shifted by 5 nm after addition of Cx44₁₂₉₋₁₅₀ (Inset, Fig. 6.10F), strikingly different from the behavior of Ca^{2+} -loaded dansyl-CaM. Under salt-free conditions, the Cx44₁₂₉₋₁₅₀ titration of dansyl-CaM provided a peptide binding apparent dissociation constants of 43 ± 8 nM with Ca^{2+} and 414 ± 20 nM in the presence of EGTA, respectively. However, under physiological salt condition (100 mM KCl), the peptide binding affinity of EGTA treated dansyl-CaM (>5 μM) was at least 100-fold weaker than Ca^{2+} -dansyl-CaM (Table 6.3). Moreover, with even higher concentration of salt (200 mM KCl), the weaker interaction in the presence of EGTA can be hardly detected by this method. Given the fact that the intracellular CaM concentration was about 9 μM (453) and CaM interacted with over 300 target proteins *in vivo* (3), such weak or unspecific interaction between the peptide and Ca^{2+} -depleted CaM would be of little physiological relevance. These results support a model in which the interaction of Cx peptides with CaM is Ca^{2+} -dependent.

Thermodynamics of the CaM-Cx interaction

The stoichiometry, binding affinity and thermodynamics of the interaction between CaM and Cx44₁₂₉₋₁₅₀ were further characterized by ITC. Representative calorimetric traces of titration are shown in Fig. 6.11. For the peptide Cx44₁₂₉₋₁₅₀, the binding event was found to be exothermic ($\Delta H = -1.15$ kcal mol⁻¹) and

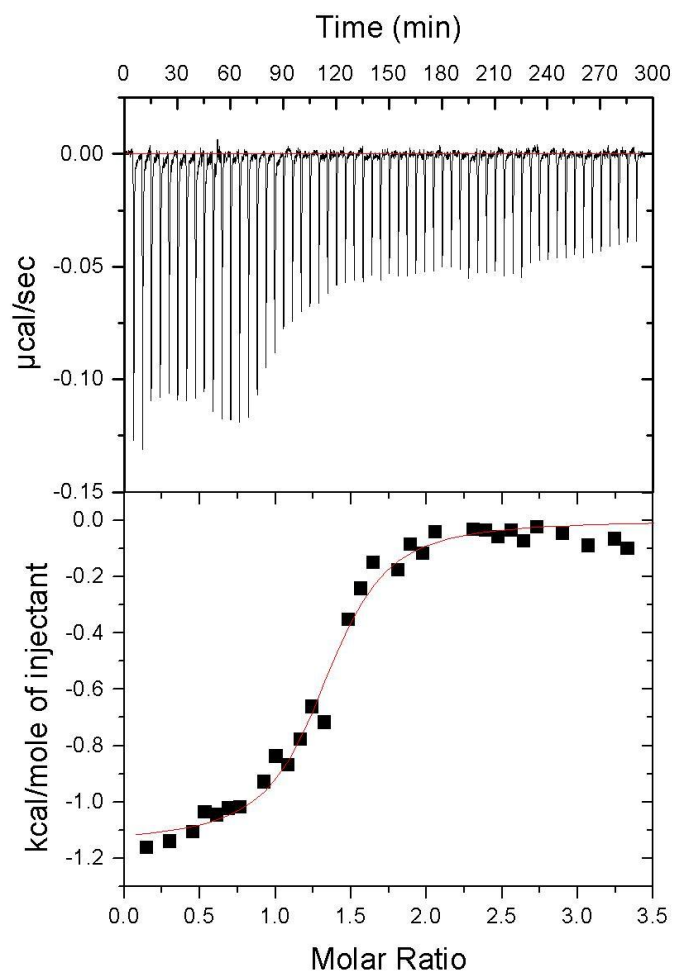


Figure 6.11. ITC microcalorimetric traces and the derived isotherms of CaM titrated with Cx peptide. Samples are prepared and dialyzed in a buffer consisting of 20 mM PIPES, 100 mM KCl, 2 mM CaCl_2 , pH 6.8. All the solutions were degassed for at least 15 min prior to experiments. 4-6 μL aliquots of peptides (400-600 μM) were injected from the syringe into the reaction cell containing 25 μM CaM in the same buffer at 5 min intervals at 25 $^\circ\text{C}$.

entropically favorable ($\Delta S = 23.97 \text{ J mol}^{-1} \text{ K}^{-1}$) with an association constant of $K_a = 1.2 \pm 0.2 \times 10^6 \text{ M}^{-1}$ ($K_d = 8.3 \pm 1.4 \times 10^{-7} \text{ M}$). Thus, the interaction seems to be both enthalpically and entropically driven with a ΔG of -8.3 kJ mol^{-1} . Since the enthalpic terms represent the forces of hydrogen bonds, van der Waals and electrostatic interaction, whereas entropy mainly reflects the hydrophobic interaction, it is evident that all these forces are contributing in concert to promote the interaction between CaM and Cx44₁₂₉₋₁₅₀. In consistent with our peptide titration data with dansyl-CaM (Fig. 6.10E), the peptide specifically bound to Ca^{2+} -CaM with a ratio near 1:1 ($n = 1.3 \pm 0.1$). The binding affinity obtained with ITC was, nonetheless, notably different from the values derived from titration data with dansyl-CaM. The discrepancy in the binding affinities using two separate methods could mainly arise from the modification to CaM by the dansyl moiety and the difference in buffering and pH conditions (6.8 in ITC vs 7.4 in dansyl-CaM titration).

6.2.6. Effect of ionic strength and pH on the CaM-Cx interaction

CaM ($pI = 4.2$) and the synthetic peptide CaM-Cx44₁₂₉₋₁₅₀ ($pI = 10.8$) were oppositely charged with a net charge of about -15 (82) and +3, respectively, at neutral pH. Such high opposite charges makes it reasonable to anticipate that electrostatic interaction could be an important determinant in the formation of CaM and Cx44₁₂₉₋₁₅₀ complex. To gain an insight into the effects of electrostatic interaction in the CaM-Cx44₁₂₉₋₁₅₀ interaction, we further analyzed the effects of varying salt and pH on the binding affinity. During the complex formation between

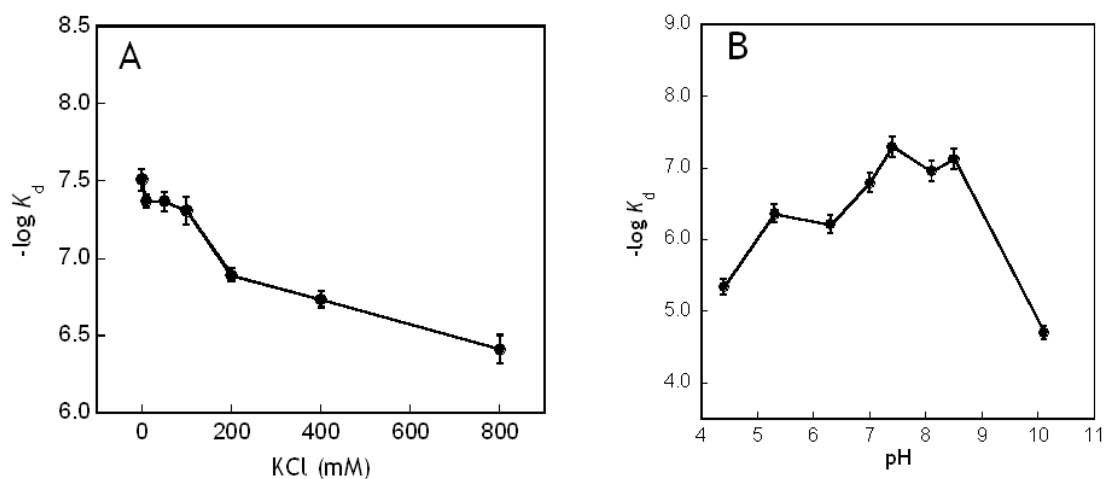


Figure 6.12. Effects of salt and pH on the binding affinity of Cx44₁₂₉₋₁₅₀. (A), Plot of $-\log K_d$ as a function of varying amount of KCl in 5 mM CaCl₂, 50 mM Tris, pH 7.4. (B), pH dependence of the binding affinity of Cx44₁₂₉₋₁₅₀. The binding affinities were derived from the peptide titration curve of dansyl-CaM. (n=2-4)

two oppositely charged macromolecules, an increase in the ionic strength is expected to decrease the binding affinity mainly due to screening of electrostatic interactions. As shown in Fig. 6.12A, the CaM-binding affinity of Cx₁₂₉₋₁₅₀ decreased by almost one magnitude of order as the salt concentration varied from 0 to 800 mM with a notable plateau stage between 10 to 100 mM. In addition, the CaM-binding affinity of the peptide exhibited a pH-dependent increase between pH 5.0 to 9.0 (Fig. 6.12B). As expected, the binding affinity at pH values near the isoelectric points of CaM (4.2) or the peptide (10.8) was drastically weakened and the difference in the affinity was ~2 orders of magnitude. Taken together, both the salt- and pH-dependence of the interaction between CaM and Cx₁₂₉₋₁₅₀ indicate that electrostatic interaction is one of the main forces governing the complex formation, which is in agreement with the observed negative (favorable) enthalpic change in ITC studies.

6.2.7. Effects of Cx binding on the metal binding properties of CaM

To examine the effect of Cx peptides on the Ca²⁺ binding properties of CaM, Ca²⁺ titrations were performed in the presence of the Ca²⁺ indicator Oregon Green 488 BAPTA-5N. Monitored by the intrinsic tyrosine fluorescence, the fluorescence change in the presence of Cx peptides was significantly left-shifted when compared to CaM alone (Fig. 6.13). The binding affinity of the peptide-bound CaM for Ca²⁺ was 1.5-3.5 fold stronger than that of CaM alone (Table 6.4). Furthermore, the Ca²⁺-dependent change in the tyrosine fluorescence of the CaM:Cx peptide mixtures exhibited stronger positive co-operativity than that of

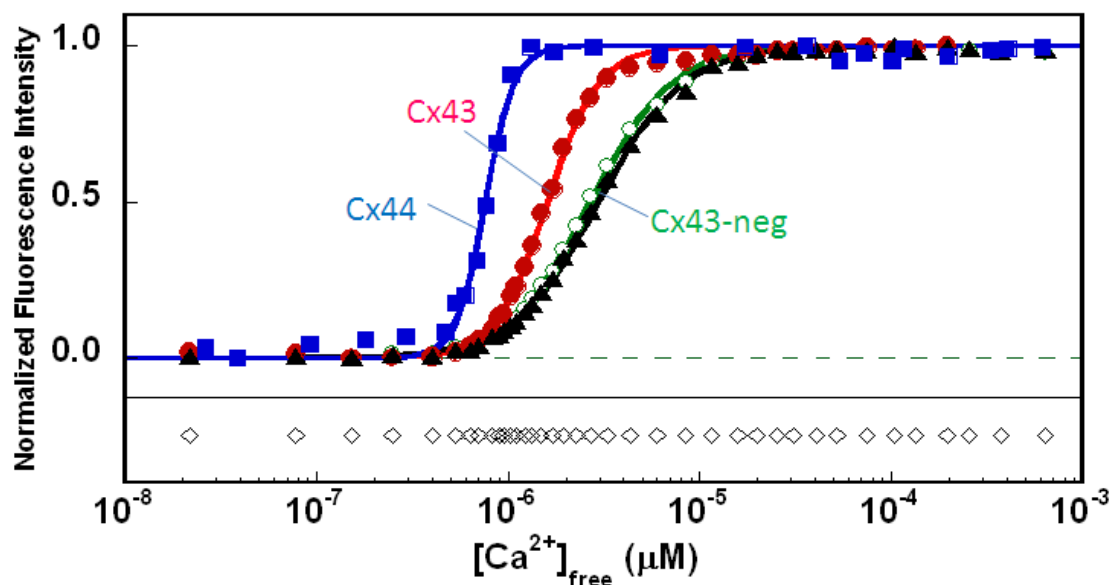


Figure 6.13. Ca^{2+} titration of CaM and CaM-Cx peptides complexes. Ca^{2+} titration of CaM (\circ), the 1:1 CaM:Cx43 randomized peptide mixture (\blacktriangle), the 1:1 CaM:Cx43₁₃₆₋₁₅₈ mixture (\bullet), and the 1:1 CaM: Cx44₁₂₉₋₁₅₀ mixture (\blacksquare) in 100 mM KCl, 50 mM Tris-HCl, pH 7.4. The intrinsic tryrosine fluorescence emission intensity of CaM or CaM-peptide mixture (8 μM) was monitored at 307 nm with fluorescence excitation at 277 nm. The Ca^{2+} indicator dye Oregon Green 488 BAPTA-5N was used to calibrate the ionized Ca^{2+} concentration. Each titration point is indicated as an open diamond at the bottom of the figure. Data from a single experiment is representative of at least triplicate experiments.

Table 6.3. The binding affinities of Cx peptides to CaM or dansyl-CaM

sample	KCl (mM)	dissociation constants (K_d , nM)	
		Ca^{2+}	EGTA
Cx43	10	100 ± 20^a	ND ^c
		240 ± 10^b	
	100	750 ± 120^a	ND ^c
		860 ± 20^b	
Cx44	10	43 ± 8^b	414 ± 20^b
	100	49 ± 3^b	$>5,000^b$

Data obtained from ^afar UV CD data or ^bDansyl-CaM fluorescence (n=3-6). ^cND: not detectable.

Table 6.4. Effect of Cx peptides binding on the Ca^{2+} binding affinity of CaM.

sample	fold		
	enhancement of tyrosine fluorescence	K_d (10^{-6} M)	n_{Hill}
CaM	2.3 ± 0.2	2.9 ± 0.1	2.1 ± 0.1
CaM-randomized control peptide	1.9 ± 0.1	2.7 ± 0.1	1.9 ± 0.2
CaM-Cx43 ₁₃₆₋₁₅₈	1.7 ± 0.1	1.6 ± 0.1	3.3 ± 0.5
CaM-Cx44	1.8 ± 0.1	0.8 ± 0.1	5.8 ± 0.4

Data obtained from Ca^{2+} titration curve (n=3).

CaM alone (Hill coefficients: 3.3 and 5.8 *versus* 2.1) for Ca^{2+} binding. It is well recognized that the Ca^{2+} affinity of CaM is significantly enhanced on binding to its receptor protein as a result of a slower dissociation rate of Ca^{2+} from CaM-receptor complex than from CaM alone (454-456). The dissociation of Ca^{2+} from CaM typically leads to the inactivation or inhibition of its bound protein (14) so slowing this dissociation would ensure that the CaM:Cx complex remains associated even during periods of frequent oscillation of $[\text{Ca}^{2+}]_i$.

To further confirm that, we measured the rapid kinetics of EGTA-induced Ca^{2+} dissociation from CaM-Cx44₁₂₉₋₁₅₀ complex by monitoring the CD signals change at 222 nm on a Jasco-810 CD spectropolarimeter with a BioLogic stopped-flow apparatus. A k_{obs} of $5.3 \pm 0.2 \text{ s}^{-1}$ was obtained by fitting the stopped-flow traces with a single-component exponential decay function (Fig. 6.14A). The observed rate constant for the EGTA-induced dissociation of Ca^{2+} from CaM could not be measured by this instrument (a minimal sampling interval of 15 ms) since the release of Ca^{2+} is too fast to be sampled (Fig. 6.14B). Bayley's group had reported an observed rate constant of $>800 \text{ s}^{-1}$ for EGTA-induced Ca^{2+} -dissociation from the CaM mutant T26W with fluorescence stopped-flow techniques (456). Conceivably, the dissociation of Ca^{2+} from CaM is drastically slowed when bound with the synthetic peptide Cx44₁₂₉₋₁₅₀. The dissociation of Ca^{2+} from CaM typically leads to the inactivation or inhibition of its bound protein (14) so slowing this dissociation would ensure that the complex remains unaltered and that the CaM-Cx44 gap junction intracellular loop maintains its functional form even during periods of frequent oscillation of $[\text{Ca}^{2+}]_i$.

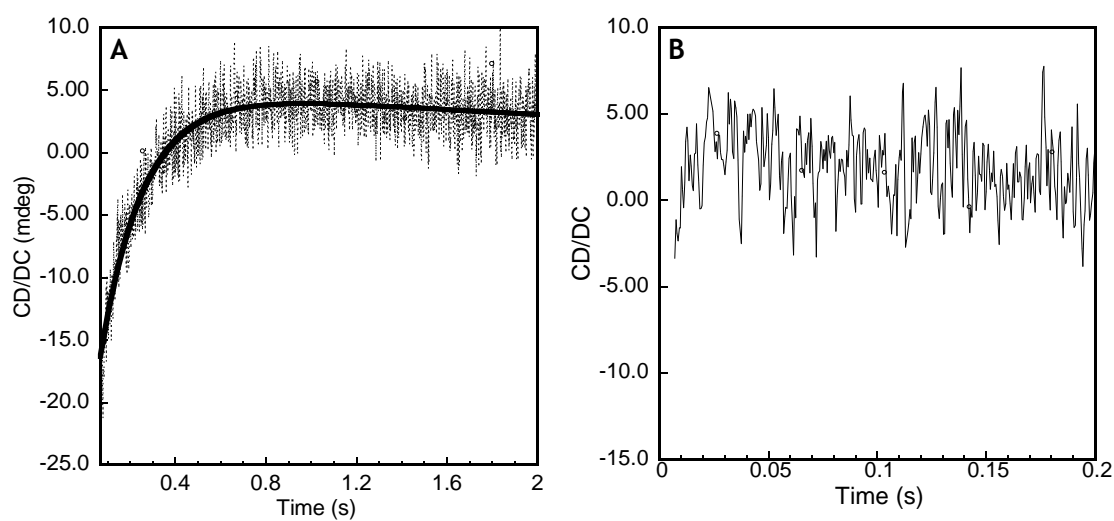


Figure 6.14. Stopped-flow trace for the EGTA-induced dissociation of Ca^{2+} from the CaM-Cx44₁₂₉₋₁₅₀ complex (A) and CaM (B). Syringe 1 contained 2 μM CaM-peptide (1:2) mixture with 0.1 mM Ca and syringe 2 consisted of an equal volume of 10 mM EGTA.

Furthermore, higher positive cooperativity enables the complex to respond to a narrower window of $[Ca^{2+}]_i$ with higher sensitivity.

6.2.8. Physiological effects of knocking out the putative CaM-binding site in Cx43. *(contribution from Dr. Monica Lurtz)*

Physiological experiments were conducted to confirm the results obtained using various biophysical approaches. Cell-to-cell dye transfer under resting and elevated intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) in communication-deficient HeLa cells transiently transfected individually with either wtCx43-EYFP or one of two Cx43-EYFP intracellular loop mutants (CaM binding site knockouts) was used to demonstrate that CaM mediates the Ca^{2+} -dependent inhibition of Cx43 by directly associating with the aa 136-158 region in Cx43.

HeLa cells transiently transfected with wtCx43-EYFP expressed large gap junction plaques at the cell-cell interface (Figs. 6.15A & E, left panels) as has been previously described; there was no significant intracellular expression of wtCx43-EYFP making it relatively straightforward to identify cells making gap junctions with adjacent cells (Fig. 6.15A). In contrast both the Cx43_{K146E, R148E}-EYFP and Cx43_{M147Q, L151E, I156E}-EYFP CaM binding-deficient mutants expressed an abundance of protein when transfected in HeLa cells, most of which was in non-plasma membrane locations (Figs. 6.15B&F, C&G left panels, respectively) making it harder to identify cells making gap junctions with adjacent HeLa cells. Under resting $[Ca^{2+}]_i$ conditions (1.8 mM extracellular Ca^{2+} concentration and no ionomycin), cell-to-cell communication (i.e., AlexaFluor 594 cell-to-cell dye

transfer) was measured between each pair of cells expressing the wtCx43-EYFP; every cell that contained EYFP fluorescence exhibited cell-to-cell transfer of injected dye indicating that Cx43-EYFP formed functional gap junctions (Fig. 6.15I). In contrast, only 31% of the double mutant Cx43_{K146E, R148E}-EYFP transfected HeLa cells and 32% of the triple mutant Cx43_{M147Q, L151E, I156E}-EYFP transfected HeLa cells exhibited cell-to-cell dye transfer (Fig. 6.15I). These results indicated that a significant fraction of the mutant Cx43-EYFP transfected cells that exhibited EYFP fluorescence did not express functional gap junctions. This fits with the EYFP fluorescence data (Figs. 6.15B&C) suggesting a significant fraction of the mutant Cx43-EYFP protein accumulates in non-plasma membrane locations in the transfected HeLa cells. Furthermore, mock-transfected HeLa cells did not exhibit cell-to-cell transfer of injected AlexaFluor 594 dye confirming that the cell-to-cell transfer exhibited by the wt Cx43-EYFP and mutant Cx43-EYFP transfected HeLa cells is gap junction mediated (Fig. 6.15D). Addition of the membrane permeant polyether antibiotic ionomycin (1 μ M) and subsequent elevation of extracellular Ca^{2+} from 1.8 mM to 21.8 mM resulted in a significant ($p < 0.001$) increase of the $[\text{Ca}^{2+}]_i$ to concentrations known to inhibit Cx43-mediated cell-to-cell dye transfer in a CaM-dependent manner. As shown in Fig. 6.15I, elevation of $[\text{Ca}^{2+}]_i$ in wtCx43-EYFP transiently transfected HeLa cells results in the inhibition of cell-to-cell dye transfer in all cells tested indicating that the addition of EYFP on the C-terminus of Cx43 did not affect the ability of elevated $[\text{Ca}^{2+}]_i$ to inhibit Cx43 gap junctions. In contrast, elevation of $[\text{Ca}^{2+}]_i$ in HeLa cells expressing the Cx43_{K146E, R148E}-EYFP double

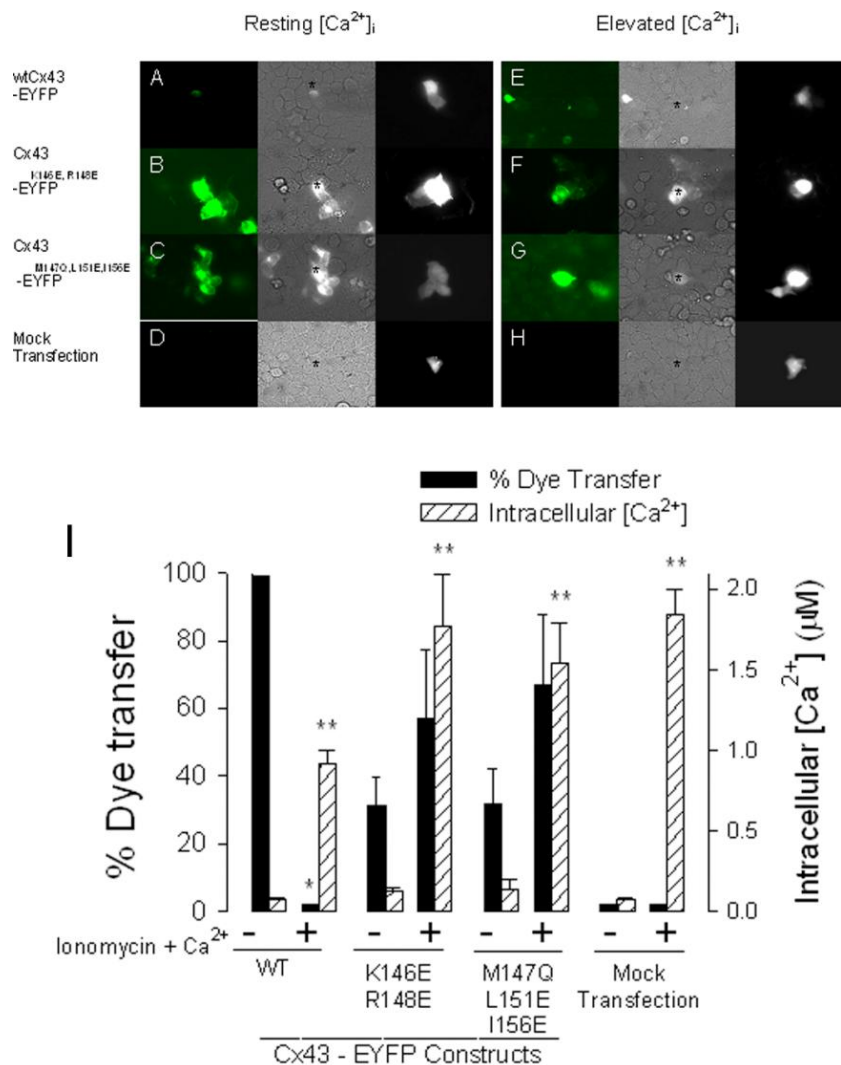


Figure 6.15. Functional effect of a sustained elevation in intracellular $[Ca^{2+}]$ on CaM binding-deficient Cx43-EYFP mutants. Cell-to-cell transfer of injected AlexaFluor 594 dye between HeLa cells transiently transfected with wtCx43-EYFP or the CaM-binding Cx43 mutants (Cx43_{K146E, R148E}-EYFP or Cx43_{M147Q, L151E, I156E}-EYFP) was measured in confluent monolayers of cells. Panels A-H show images of transiently transfected wt or mutant Cx43-EYFP or mock-transfected HeLa cells. The left image shows the EYFP fluorescence of the injected and adjacent cells, the middle image is the fluorescence of EYFP plus brightfield, and the images on the right are of AlexaFluor 594 dye transfer. Panels A&E: wtCx43-EYFP transiently transfected HeLa cells; Panels B&F: Cx43_{K146E, R148E}-EYFP transiently transfected HeLa cells; Panels C&G: Cx43_{M147Q, L151E, I156E}-EYFP transiently transfected HeLa cells; Panels D&H: mock-transfected HeLa cells. Panel I: Summary data for cell-to-cell dye transfer (solid bars) and $[Ca^{2+}]_i$ (hatched bars) in the absence or presence of elevated $[Ca^{2+}]_i$. * indicates cell-to-cell dye transfer was significantly lower than cell-to-cell dye transfer determined in low $[Ca^{2+}]_i$; ** indicates $[Ca^{2+}]_i$ was significantly higher than the low $[Ca^{2+}]_i$ value. (contribution from Dr. Monica Lurtz)

mutant or the Cx43_{M147Q, L151E, I156E}-EYFP triple mutant did not result in the inhibition of cell-to-cell dye transfer (57% and 67% [respectively] of attempts resulted in dye transfer). The apparent increase in the percentage of injected cells inhibiting cell-to-cell dye transfer in elevated $[Ca^{2+}]_i$ versus resting $[Ca^{2+}]_i$ was not statistically significant ($p>0.1$). There was no cell-to-cell transfer of dye when AlexaFluor 594 dye was injected into mock HeLa transfected cells under either resting or elevated $[Ca^{2+}]_i$ conditions (Fig. 6.15H).

6.2.9. Indication of in vivo interaction

In this study we provide evidence that a peptide derived from the intracellular loop of the major gap junction protein Cx43 and Cx44, which has the highest predictive score of any region of Cx to bind CaM (Fig. 6.4), binds CaM with high affinity in a Ca^{2+} -dependent manner (Table 6.3). Given that intracellular CaM concentration is approximately 9 μM (453), such an affinity would imply that this CaM-binding region of Cx43 would be saturated with CaM when intracellular Ca^{2+} concentration is elevated and that this association accounts for the well described ability of Ca^{2+} to inhibit gap junctions (75). As indicated by our circular dichroism, fluorescence and NMR data, the stoichiometry of Cx peptides:CaM binding is close to 1:1. To our knowledge, this report presents the first demonstration of a direct CaM-Cx43 or CaM-Cx44 interaction *in vitro*, which would explain our previous *in vivo* observation that preincubation of lens cell cultures or Cx43-transfected HeLa cells with CaM antagonists counteracts the intracellular Ca^{2+} -dependent inhibition of Cx43-containing gap junctions (75,436).

The results clearly demonstrate that the Ca^{2+} -dependent inhibition of Cx43 gap junctions is mediated via the association of Ca^{2+} -CaM with the intracellular portion of Cx43. In addition, our results support a mechanism of Cx43 or Cx44 regulation by CaM in which CaM first binds to Ca^{2+} , and then the Ca^{2+} -CaM complex associates with Cx43 or Cx44 to inhibit the gap junction, rather than a stable CaM-Cx complex in which Ca^{2+} binds to the complex which then results in gap junction inhibition.

A different model for the gating of Cx43 has been proposed by Delmar's laboratory (438) in which an intracellular gating element within the C-terminal domain of Cx43 interacts with a region of the Cx43 cytoplasmic loop. In that report Duffy *et al* used surface plasmon resonance, ELISA and NMR approaches to demonstrate that a portion of the C-terminal of Cx43 comprising amino acids 255-382 associated with a peptide corresponding to the second half of the Cx43 intracellular loop comprising amino acids 119-144, which was significantly enhanced by low pH. The results reported here for the inhibition of Cx43 by Ca^{2+} -CaM has an interesting corollary to this pH-dependent inhibition described by Duffy *et al* (438) in that binding of a portion of Cx43 C-terminus region or CaM to the cytoplasmic region of Cx43 inhibits this gap junction protein. However, while Sorgen *et al* (457) reported that dimerization of the Cx43 C-terminal region may be one of the structural changes involved in the pH regulation of Cx43, our data reported here indicate that Ca^{2+} -CaM and the intracellular loop peptide associate as a 1:1 complex; i.e. that a single Ca^{2+} -CaM binds to Cx43 (i.e. 6 Ca^{2+} -CaM per connexon hemichannel) to effect gap junction closure. Indeed

while there is partial overlap of the C-terminal binding domain on the intracellular loop (residues 119-144) and the CaM binding domain (residues 136-158), it is clear that CaM will associate with a portion of Cx43 that is very close to the third transmembrane loop of this connexins so it is well positioned to physically obstruct the gap junction pore.

As in any molecular study conducted *in vitro*, direct extrapolation between the data presented in this report using a peptide, and the mechanisms of Ca^{2+} -CaM gating of Cx43 in intact cells needs to be made with caution. However, it is worth noting that there are several examples in the literature demonstrating that two molecular domains that associate in living cells also associate as separate domains *in vitro* (60,458). In the studies of CaM-target interaction, synthetic peptides have been widely used due to their excellent ability to mimic specific domains of native proteins (401,419,459-462). Indeed Peracchia *et al* (391,434) have shown that calmodulin is associated with gap junctions and plays a direct role in the chemical gating of Cx32-containing gap junctions. This was supported by data for CaM binding to Cx32 in which Török *et al* (435) identified two distinct CaM binding amino acid sequences. The sequence of one of these sites in the N-terminal domain of Cx32 contains a CaM binding motif common to a large class of CaM-dependent proteins (439); notably this N-terminal CaM binding sequence is absent in Cx43 and Cx44.

6.2.10. Indication of potential CaM-Cx43₁₃₆₋₁₅₈ binding mode

Through its reversible binding of Ca^{2+} and the resultant conformational changes, CaM is capable of interacting with over 300 target proteins to regulate a

range of cellular events (3). Although the sequence identity among the CaM targeted sequences is low, all the consensus peptides for such sequences (18 to 26-mer peptides) possess common features including their ability to form amphipathic α -helices, containing 2-6 positively charged residues, and their pattern of bulky hydrophobic residues that anchor the peptides to the hydrophobic cleft of CaM. Our results demonstrate that the Cx43₁₃₆₋₁₅₈ sequence forms an α -helical structure in TFE as well as in the CaM-peptide complex, similar to other CaM-bound peptides (442,447,449,463,464). Two lysine and two arginine residues are contained in the Cx43₁₃₆₋₁₄₈ sequence which has a high predictive score for CaM binding. The basic residues in the N-terminus of CaM-binding sequences have been proposed to ensure an antiparallel orientation of the peptide with respect to CaM (421), which is believed to optimize the electrostatic attraction between the basic residues of the peptide with negatively charged residues in CaM. In addition, the hydrophobic residues M¹⁴⁷-L¹⁵¹-I¹⁵⁶ of the Cx43 peptide (I¹³⁹-L¹⁴³-V¹⁴⁸) in the Cx44 peptide) reflect the 1-5-10 type of CaM-binding pattern (Fig. 6.4), which is similar to the CaM-binding regions of CaMKI and CaMKII (90,439,463).

In addition to the formation of α -helical structure by the Cx peptides, the hydrophobicity of the peptide environment as well as the tertiary arrangement of CaM are also changed following the formation of CaM:Cx complexes. As indicated by the near UV CD studies, the chemical environment around the tyrosine residues is significantly perturbed following the binding of Cx43₁₃₆₋₁₅₈ to CaM. Furthermore, the significant blue shift and concomitant enhancement of

fluorescence intensity induced by Cx peptides binding to dansyl-CaM (Fig. 6.10) suggests that the dansyl group in dansyl-CaM is shielded from the solvent and moves into a highly hydrophobic environment on binding to Cx peptides; similar changes have been observed for the binding of dansyl-CaM to other receptors (465-467). A number of the currently-assigned CaM resonances in the NMR ^1H - ^{15}N HSQC spectra exhibited chemical shift movements ≥ 0.05 ppm. These residues were spread in both the N- and C-terminal domains of CaM as well as in the linker region, indicating that a global conformational change occurred upon binding of the Cx43₁₃₆₋₁₅₈ peptide to CaM. Such global changes of amide chemical shifts have also been reported in other CaM complexes such as Ca^{2+} -CaM-skeletal muscle myosin light chain kinase and Ca^{2+} -CaM-CaMKI (420,468). Together these observations suggest that the Ca^{2+} -CaM-Cx interaction might adopt the commonly seen wrapping-around mode of action (469).

6.2.11. In vivo functional analysis of the putative Cx43 CaM-binding site

Transient expression of a Cx43-EYFP construct without and with mutations in communication-deficient HeLa cells provided a system by which to test our biophysical data in a physiological manner. Because CaM interaction with a protein requires both an electrostatic interaction and a hydrophobic interaction, we generated two Cx43 mutants - Cx43_{K146E, R148E}-EYFP that knocks out the Cx43 electrostatic interaction with CaM, and Cx43_{M147Q, L151E, I156E}-EYFP that knocks out the hydrophobic interaction of Cx43 with CaM. Transient expression of these EYFP-tagged Cx43 constructs was consistent with the

anticipated results. The wtCx43-EYFP formed very large gap junction plaques and dye transfer was restricted to those cells expressing EYFP fluorescence (see Fig. 6.15A&E). With elevated $[Ca^{2+}]_i$, these junctions were no longer dye coupled as reported by us previously with wt Cx43 that lacked the EYFP label.

As CaM has been implicated in Cx32 assembly at two stages of oligomerization, it was anticipated that the Cx43_{K146E, R148E}-EYFP and Cx43_{M147Q, L151E, I156E}-EYFP mutants would probably encounter problems reaching the plasma membrane, and perhaps fail to form functional gap junctions. Indeed, although both of the mutant Cx43 proteins did express, the majority of the expressed protein was trapped inside the cell, very little of this protein appeared to reside in the plasma membrane, and no obvious gap junction plaques were observed. Thus it was not anticipated that either of the mutants would exhibit cell-to-cell transfer of injected dye as reproducibly as we observed with the Cx43-EYFP. This proved to be the case because while in low ($\sim 1 \mu M$) $[Ca^{2+}]_i$ 100% of the wt Cx43-EYFP transfected HeLa cells exhibited cell-to-cell transfer of injected dye, only approximately 30% of each of the mutant-transfected HeLa cells appeared to be functional and exhibited cell-to-cell transfer of injected dye (Fig. 6.15I). However, while elevated $[Ca^{2+}]_i$ completely prevented cell-to-cell dye transfer between wt Cx43-EYFP transfected HeLa cells, elevated $[Ca^{2+}]_i$ was now unable to inhibit cell-to-cell dye transfer between HeLa cells transfected with these Cx43 mutants. Thus, knocking out the CaM binding capability of the residue 136-158 region of Cx43 abolished the Ca^{2+} -dependent inhibition of Cx43 gap junctions supporting our biophysical data that demonstrates that CaM

mediates the Ca^{2+} -dependent inhibition of Cx43 via its interaction with residues 136-158 in this connexin.

6.3. Summary

In summary, we have identified a CaM binding sequence in the ubiquitous gap junction protein Cx43 and Cx44. This sequence resides in a juxtamembrane region of the only intracellular loop of Cx43 and Cx44. Our results demonstrate a 1:1 Ca^{2+} -dependent CaM-Cx peptides interaction with an affinity in the submicromolar range. The binding of this peptide to CaM enhances the Ca^{2+} affinity of CaM 2-4 fold. These results explain the molecular basis of our previously reported Ca^{2+} -CaM-dependent regulation of both lens and Cx43-containing gap junctions via a direct interaction of CaM with this connexin protein (75,436). The data reported here confirms that this regulation is effected via the Ca^{2+} -dependent association of CaM with residues 136-158 of Cx43 (129-150 of Cx44) that in turn affects a change in the structural organization of Cx43 or Cx44 such that gap junction permeability is significantly inhibited. Further elucidation of the structural changes within both CaM and Cx43 is the subject of ongoing investigations in this laboratory.

7. Conclusions and major discoveries

We have developed a computational tool for the accurate prediction of EF-hand or EF-hand-like Ca^{2+} -binding proteins from primary sequences that will benefit the whole metalloprotein field. To our knowledge, our server (<http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm>) is the first of its kind to provide such service for scientists. More importantly, the comprehensive prediction and detailed analysis on putative prokaryotic and viral EF-hand and EF-hand like Ca^{2+} -binding proteins will contribute to improved understanding of versatile roles Ca^{2+} ions play in various biological systems, and should encourage more intensive studies toward this direction.

The studies on the rubella virus (RUB) nonstructural (NS) protease have established a model in which both Ca^{2+} and Zn^{2+} are required for the optimal protease activity and efficient virus replication.

The studies on the interaction between CaM and connexins 43 and 44 have confirmed a Ca^{2+} -dependent CaM-Cx interaction model and advanced the understanding of the structural basis underlying Ca^{2+} -dependent gating of gap junction.

All of the completed work discussed in this dissertation combined computational prediction methods, followed by extensive experimentation to validate predictions through protein engineering and peptide approaches. More importantly, as demonstrated by results from cell-based assays, the predicted

motifs have been tightly linked to biological functions. Such three-pronged “prediction to protein engineering to function” approaches could be extended to the study of many other subjects.

Publications and manuscripts

Zhou Y, Yang W, Lurtz MM, Ye Y, Huang Y, Lee HW, Chen Y, Louis CF, Yang JJ. Identification of the calmodulin binding domain of connexin43. *J Biol Chem*. 2007 Sep 27 [Epub ahead of print].

Huang Y*, **Zhou Y***, Yang W, Butters R, Lee HW, Li S, Castiblanco A, Brown EM, Yang JJ. Identification and dissection of Ca²⁺-binding sites in the extracellular domain of Ca²⁺-sensing receptor. *J Biol Chem*. 2007; 282(26):19000-10. (*Co-first author)

Zhou Y, Tzeng WP, Yang W, Zhou Y, Ye Y, Lee HW, Frey TK, Yang J. Identification of a Ca²⁺-binding domain in the rubella virus nonstructural protease. *J Virol*. 2007; 81(14):7517-28.

Liu Y, Tong Q, **Zhou Y**, Lee HW, Yang JJ, Buhring HJ, Chen YT, Ha B, Chen CX, Yang Y, Zen K. Functional elements on SIRPα IgV domain mediate cell surface binding to CD47. *J Mol Biol*. 2007;19;365(3):680-93.

Zhou Y, Yang W, Kirberger M, Lee HW, Ayalasomayajula G, Yang JJ. Prediction of EF-hand calcium-binding proteins and analysis of bacterial EF-hand proteins. *Proteins*. 2006; 65(3):643-55.

MANUSCRIPT IN REVISION OR IN PREPARATION

Yubin Zhou, Wei Yang, Monica Lurtz, Yun Huang, Charles Louis, and Jenny Yang. Spectroscopic characterization of the interaction between calmodulin and connexin44. *(In preparation)*

Yubin Zhou, Wen-Pin Tzeng, Teryl Frey and Jenny Yang. Functional and conformational studies on a cysteine-rich metal-binding domain in rubella virus nonstructural protease. *(In preparation)*

Yubin Zhou, Teryl Frey, Jenny Yang. Calcium and virus infection. *(In preparation)*

Yubin Zhou, Jenny Yang. Calmodulin: a versatile Ca^{2+} -sensing protein. *(In preparation)*

Appendices

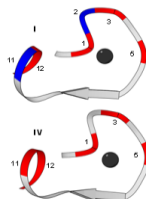
This section includes: 1) analyzing key factors governing the metal binding events by inserting well-characterized Ca^{2+} -binding sequences from cellular proteins into CD2.D1 with the grafting approach; 2) predicting and identifying a putative CaM binding site in rubella virus nonstructural protease; 3) list of CD2 variants I made and their primary sequences and physiochemical properties.

A1. Analysis of factors affecting the metal-binding affinities of isolated Ca^{2+} -binding motifs

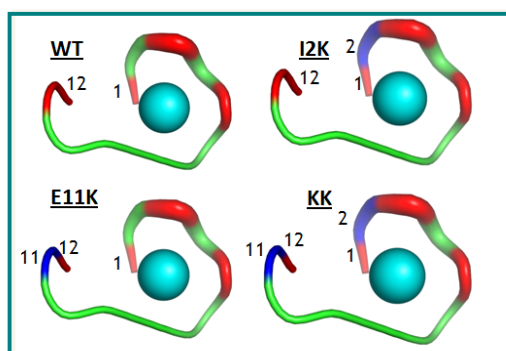
The established grafting approach enables us to study factors governing the metal binding properties of isolated EF-hand motifs without the complexity encountered in multiple cooperative Ca^{2+} -binding processes. Specifically, we have generated a series of engineered CD2 proteins grafted with i) the 12-residue EF-loop and modified loops from the Ca^{2+} -binding site III of calmodulin to study the effects of charge distribution on metal binding affinities. ii) the pseudo (denoted as CD2.Cal1) and canonical EF-hand loops (denoted as CD2.Cal2) from calbindin D9k to examine the effects of coordinating residues on metal binding affinities. At present, these constructs have been made and some preliminary studies have been carried out. However, this project is still in its early stage and will be further explored.

A

	1	2	3	4	5	6	7	8	9	10	11	12	Ligand Charge	Overall Charge	$K_{d,Tb}$ (μ M)
LOOP I	D	K	D	G	D	G	T	I	T	T	K	E	-4	-2	5 \pm 1
LOOP II	D	A	D	G	N	G	T	I	D	F	P	E	-4	-4	25 \pm 6
LOOP III	D	K	D	G	N	G	Y	I	S	A	A	E	-3	-2	20 \pm 6
LOOP IV	D	I	D	G	D	G	Q	V	N	Y	E	E	-4	-5	72 \pm 10



B



	1	2	3	4	5	6	7	8	9	10	11	12	Ligand Charge	Overall Charge
WT	D	I	D	G	D	G	Q	V	N	Y	E	E	-4	-5
I2K	D	K	D	G	D	G	Q	V	N	Y	E	E	-4	-4
E11K	D	I	D	G	D	G	Q	V	N	Y	K	E	-4	-3
KK	D	K	D	G	D	G	Q	V	N	Y	K	E	-4	-2

Figure A1. The charge-ligand-balanced model. Both the number of negatively charged ligand residues and the balanced electrostatic dentate-dentate repulsion by the adjacent charged residues (positions 2 and 11) are two determinants for the relative affinities of the EF-loops (1,2). Experimental design to test the “charge-ligand-balanced model”. Substitutions of position 2 (residue Ile) and 11 (residue Glu) with Lys residues in the EF-loop IV of calmodulin is expected to mimic the EF-loop I of calmodulin and has similar affinity.

Effects of charge distribution

Previous studies carried out by Dr. Yiming Ye and Dr. Hsiau-Wei Lee have characterized the metal binding affinities of 4 EF-loops from calmodulin (1,2). It has been found that the Tb^{3+} binding constants follow the order of: I > II ~ III > IV, which is correlated to the overall charge of the isolated EF-loops (-2, -2, -4, -5, respectively) (Fig. A1A). Based on this observation, a “Charge-ligand-balanced model” has been hypothesized, in which both the number of negatively charged ligand residues and the balanced electrostatic dentate-dentate repulsion by the adjacent charged residues are two determinants for the relative affinities of the EF-loops (1,2). To further test this hypothesis, we further engineered the “lowest-affinity” site IV of EF-loop (denoted as CD2.IV5G) by mutating residues at position 2 and 11 individually or in combination to positively-charged residue K (denoted as I²K, E¹¹K and KK, respectively) (Fig. A1B). These proteins have been expressed and purified to homogeneity (Fig. A2A). The grafting of these sequences will not disrupt the host protein structure (Fig. A2B & C). We will carry out metal titration experiments to obtain metal binding affinities for each variant in the future.

Effects of coordinating residues

EF-hand motifs are divided into two major types: the canonical EF-hands and the pseudo EF-hands. The S100 protein Calbindin D9k contains a pseudo EF-hand Ca^{2+} -binding sites at its N terminus and a canonical EF-hand Ca^{2+} -binding sites at its C terminus (Fig. A3A). The major difference between these

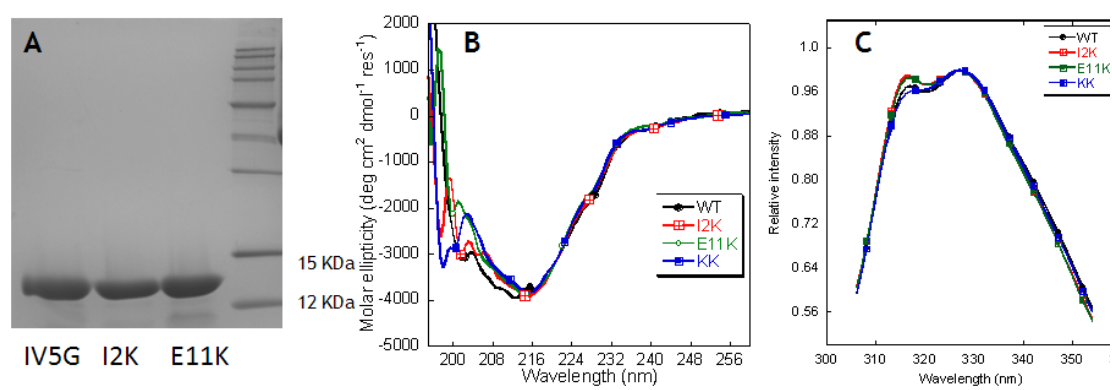


Figure A2. The grafting of EF-loop does not disrupt the host protein structure. (A) SDS-PAGE of proteins under study. Far UV CD (B) and Trp fluorescence (C) spectra of engineered proteins are also shown.

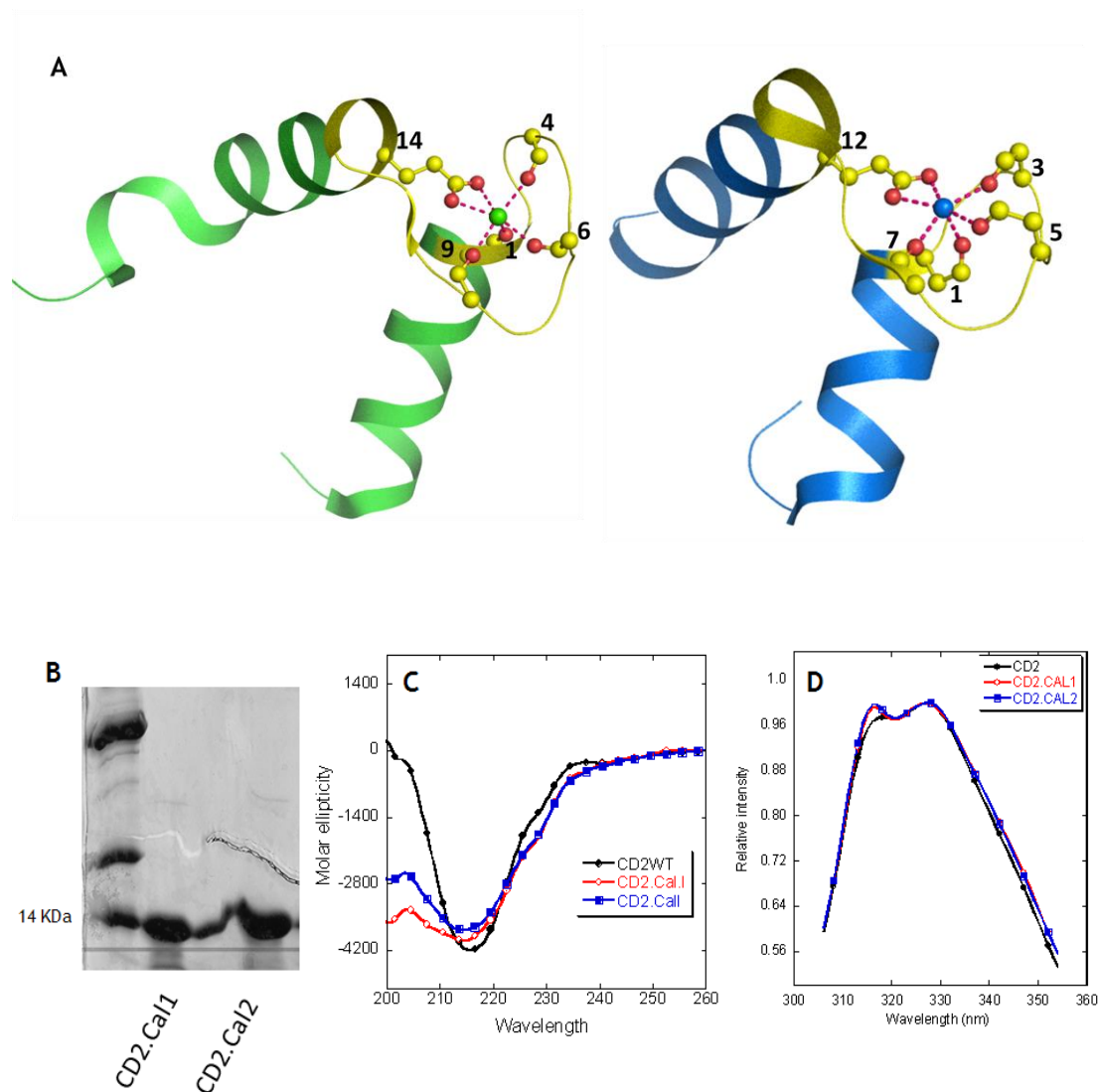


Figure A3. The pseudo and canonical EF-hands in calbindin D9k. (A) 3D cartoon representation of Ca^{2+} -binding sites (highlighted as yellow) in Calbindin D9k (PDB code: 1b1g). (left) Pseudo EF-hand loop coordinates Ca^{2+} by mainchain carbonyls (position 1,4,6,9) and sidechain carboxylates at position 14. (right) Canonical EF-hand loop chelates Ca^{2+} by sidechain carboxylates (position 1,3,5,12) and one mainchain carbonyl (position 7). The water ligand is not shown here. (B) Grafting the EF-loops to CD2.D1. After the insertion of 12- or 14-residue EF-loops, the host protein retains its native structure, providing the foundation to study the effects of coordinating residues on metal binding properties. (left) SDS-PAGE of engineered proteins CD2.Cal1 and CD2.Cal2. (middle) Far UV CD spectrum of engineered proteins. The observed difference below 210 nm arises from the insertion of 12 residues that increases the random coil structure by ~10%. (right) Tryptophan fluorescence spectra of engineered proteins.

two groups lies in the Ca^{2+} -binding loop: the 12-residue canonical EF-hand loop binds Ca^{2+} mainly via sidechain carboxylates (loop sequence positions 1, 3, 5, 12), whereas the 14-residue pseudo EF-hand loop chelates Ca^{2+} primarily via backbone carbonyls (positions 1, 4, 6, 9). The residue at the position 9 (11 in pseudo EF-hand) coordinates the Ca^{2+} ion through a bridged water molecule. The pseudo and canonical EF-hand loops have a bidentate ligand (Glu) at position 14 and 12, respectively. It is worth asking how the coordination of metal ion by main chain carbonyls or side chain carboxylates will affect metal binding properties. The grafting approach, which obviates the metal binding cooperativity encountered in natural proteins containing multiple EF-hand Ca^{2+} -binding sites, would provide a solution to this question. Given that, we inserted the 14-residue pseudo EF-loop (denoted as CD2.Cal1) and the 12-residue canonical EF-loop (denoted as CD2.Cal2) into CD2.D1 between Ser⁵² and Gly⁵³ with –GGG– linkers. Now we have successfully designed and purified the proteins (Fig. A3B-D) and will soon address this interesting question.

A2. Prediction of a putative CaM binding site in the RUB NS protease

Using the search program for CaM-binding sites (90) as described in Section 6.2.1, a putative “IQ-like” CaM binding region with high predictive score was identified in the rubella virus nonstructural protease. The location of the CaM binding site is immediately prior to the entering helix of the helix-loop-helix EF-hand Ca^{2+} -binding motif (Fig A4). To test whether the predicted sequence is

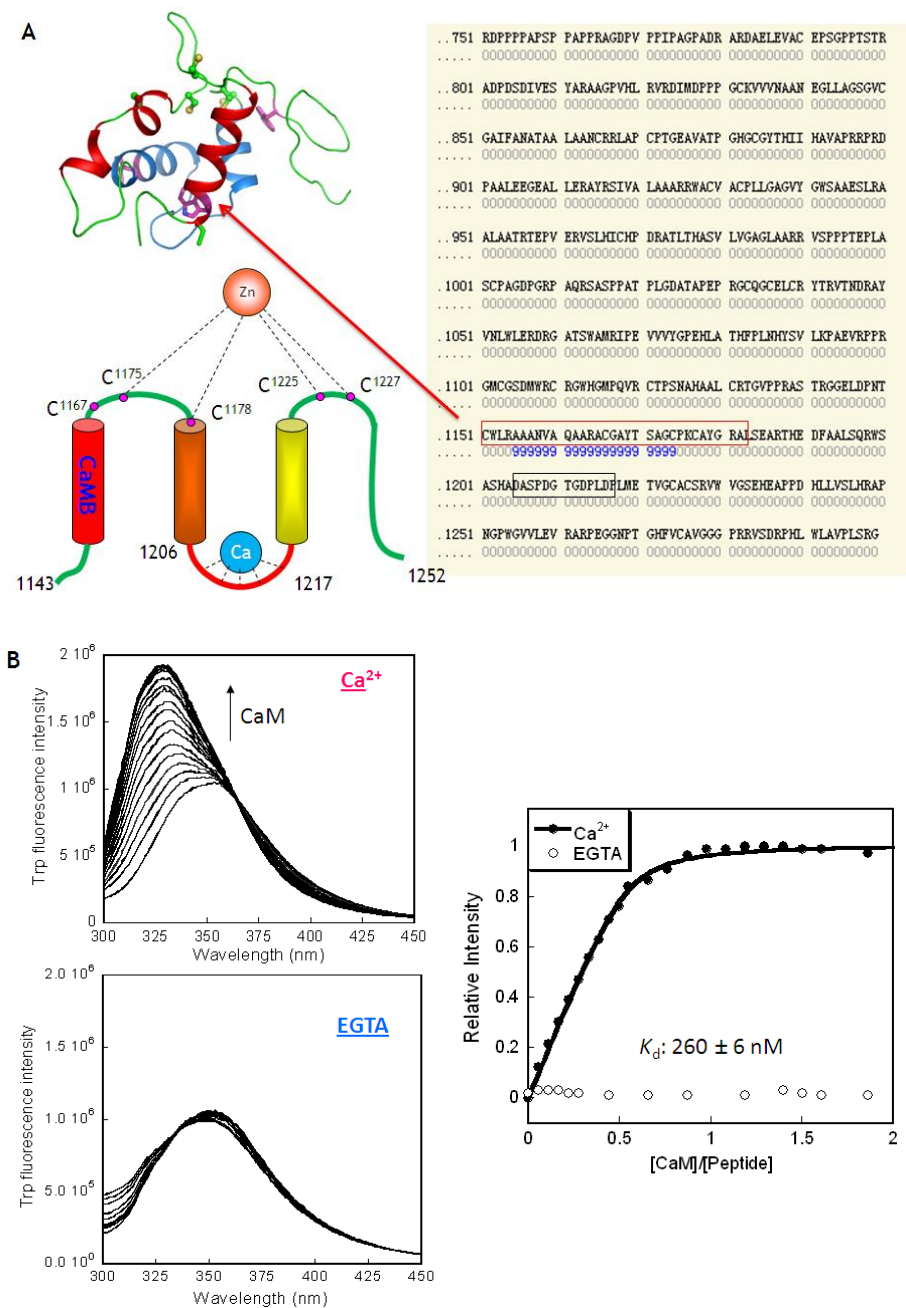


Figure A4. Prediction of a CaM binding site in RUB NS protease. (A) The location of the predicted CaM binding site (CaMB). The CaMB is situated prior to the entering helix of the EF-hand motif (black box). (B) Titration of CaM into a synthetic peptide (red box) encompassing the putative CaMB in the presence of Ca^{2+} (left upper) or EGTA (left lower). Right panel showed the Trp intensity as a function of the $[\text{CaM}]/[\text{Peptide}]$ ratio.

capable of interacting with CaM *in vitro*, we synthesized a 32-mer peptide encompassing the CaM binding region (¹¹⁵²CWLRAAANVAQAARACGAYTSAGCPKCAYGRA¹¹⁸³). Since the peptide contains a Trp residue and CaM contains zero Trp and two Tyr residues, the Trp fluorescence signal from the peptide can be used to monitor the interaction between the peptide and CaM. As shown in Fig. A4B, the Trp fluorescence emission spectrum of the free peptide exhibited a major peak at 354 nm, indicating that the peptide is largely unstructured and the only Trp residue is fully exposed to the solvent. Following the addition of CaM in the presence of 1 mM Ca²⁺, the fluorescence intensity increased by nearly 2 fold and the emission peak underwent a 16-nm blueshift, strongly suggesting that the Trp residue is now buried within a highly hydrophobic environment due to the peptide-CaM interaction. An apparent dissociation constant of 260 ± 6 nM is obtained by fitting the curve with a 1:1 binding process. In the absence of Ca²⁺, however, neither fluorescence intensity enhancement nor emission peak shift is observed. Thus, the interaction between the CaM binding site in RUB NS protease and CaM seems to be a Ca²⁺-dependent event, which is commonly seen in many other CaM-target interactions (90). Since the putative CaM binding site is within the NS protease, it is expected that its effect would be on protease activity, either through its interaction with the neighboring EF-hand Ca²⁺-binding motif or by a mechanism independent of the EF-hand motif. It is also possible that the CaM binding event has an effect on other unknown aspects of P150 function. The physiological relevance of this CaM binding domain will be further characterized.

A3. Primary sequences and physiochemical properties of engineered CD2 variants

Primary sequences of engineered CD2 variants:

>CD2

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGAFEI
LANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2 . RUBCA

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGDA
SPDGTGDPLDGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2 . RUBCA . AA

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGDA
SPAGTGDPLAGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2 . RUBCA . EF

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGSO
RWSASHADASPDGTGDPLDPLMETVGAGGGGA FEILANGDLKIKNLTRDDSGTYNVTVYS
TNGTRILNKALDLRILE

>CD2 . SIN

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGDL
DNEKMLGTREGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2 . POX

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGDY
DGNGTETRGEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2 . SHR

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLXAEFKRKMKPFLKSGGGDL
DGDGKLSKTEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2 . CAL1

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGAA
KEGDPNQLSKEEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRI
LE

>CD2 . CAL2

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGGDK
NGDGEVSFEEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2.III5G.F

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**DK**
DGNGYISAAELRHVMTNLGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNK
 ALDLRILE

>CD2.III5G.EF

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**EE**
EIREAFRVFDKDGNGYISAAELRHVMTNLGGGA FEILANGDLKIKNLTRDDSGTYNVTV
 YSTNGTRILNKALDLRILE

>CD2.III5G.SKEAA

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**SE**
EEKREAERVFDKDGNGYISAAELRHAATNLGGGA FEILANGDLKIKNLTRDDSGTYNVT
 VYSTNGTRILNKALDLRILE

>CD2.IV5G

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**DI**
DGDGQVNYEEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2.IV5G.I2K

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**DK**
DGDGQVNYEEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2.IV5G.E11K

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**DI**
DGDGQVNYKEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2.IV5G.KK

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**DK**
DGDGQVNYKEGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRILNKALDLRILE

>CD2.STIM1

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**SE**
EAVRN IHKLMDDDANGDV DVEESDEFLREDLGGGA FEILANGDLKIKNLTRDDSGTYNV
 TVYSTNGTRILNKALDLRILE

>CD2.PMR1

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**SV**
DEALEKLD TDKNGLRSSNEANNRRSLYGGGA FEILANGDLKIKNLTRDDSGTYNVTVY
 STNGTRILNKALDLRILE

>CD2.CX43

GSRD SGTVWGALGHGINLNIPNFQMTDDIDEVRWERGSTLVAEFKRKMKPFLKSGGG**KY**
GIEEHGKV KMRGGLLR TYIISGGGA FEILANGDLKIKNLTRDDSGTYNVTVYSTNGTRI
 LNKALDLRILE

Table A1. The physiochemical properties of engineered CD2 variants.

CD2 variants	pI	MW (Da)	ϵ ($M^{-1} cm^{-1}$)	Description
CD2	8.5	11286.8	13940	wt CD2 domain 1
CD2.RUBCa	5.3	12713.1	13940	EF-hand loop from rubella virus
CD2.RUBCa.AA	6.1	12625.1	13940	EF-hand motif mutant from rubella virus
CD2.RUBCa.EF	5.6	14523.1	19630	EF-hand motif from rubella virus
CD2.SIN	6.2	12974.6	13940	EF-hand loop from rubella virus
CD2.POX	5.6	12867.3	15220	EF-hand loop from rubella virus
CD2.SHR	6.2	12732.2	13940	EF-hand loop from rubella virus
CD2.CAL.SITE1	6.2	13069.6	13940	pseudo EF-hand loop from calbindin D9k
CD2.CAL.SITE2	5.3	12879.3	13940	canonical EF-hand loop from calbindin D9k
CD2.IV5G	5.1	12907.3	15220	EF-hand loop from calmodulin site 4
CD2.IV5G.I2K	5.3	12922.3	15220	EF-hand loop mutant from calmodulin site 4
CD2.IV5G.E11K	5.6	12906.4	15220	EF-hand loop mutant from calmodulin site 4
CD2.IV5G.KK	6.2	12921.4	15220	EF-hand loop mutant from calmodulin site 4
CD2.III5G	6.1	12793.3	15220	EF-hand loop from calmodulin site 3
CD2.III5G.F	7.4	13758.4	15220	loop-F motif from calmodulin site 3
CD2.III5G.EF	6.0	15165.0	15220	EF-hand motif from calmodulin site 3
CD2.III5G.SKEAA	6.0	15160.9	15220	EF-hand motif mutant from calmodulin site 3
CD2.STIM1	4.9	15378.1	13940	EF-hand motif from STIM1
CD2.PMR1	7.2	14906.6	15220	EF-hand motif from yeast ATPase
CD2.CX43	9.4	14203.1	16500	Calmodulin binding site from connexin43

References

1. Ye, Y., Lee, H. W., Yang, W., Shealy, S., and Yang, J. J. (2005) *J Am Chem Soc* **127**(11), 3743-3750
2. Ye, Y., Shealy, S., Lee, H. W., Torshin, I., Harrison, R., and Yang, J. J. (2003) *Protein Eng* **16**(6), 429-434
3. Berridge, M. J., Bootman, M. D., and Lipp, P. (1998) *Nature* **395**(6703), 645-648
4. Michiels, J., Xi, C., Verhaert, J., and Vanderleyden, J. (2002) *Trends Microbiol* **10**(2), 87-93
5. Rigden, D. J., Jedrzejewski, M. J., Moroz, O. V., and Galperin, M. Y. (2003) *Trends Microbiol* **11**(7), 295-297
6. Zhou, Y., Yang, W., Kirberger, M., Lee, H. W., Ayalasomayajula, G., and Yang, J. J. (2006) *Proteins* **65**(3), 643-655
7. Tsien, R. W., and Tsien, R. Y. (1990) *Annu Rev Cell Biol* **6**, 715-760
8. Levine, B. A., Dalgarno, D. C., Esnouf, M. P., Klevit, R. E., Scott, G. M., and Williams, R. J. (1983) *Ciba Found Symp* **93**, 72-97
9. Schroder, B., Schlumbohm, C., Kaune, R., and Breves, G. (1996) *J Physiol* **492** (Pt 3), 715-722
10. Buchanan, J. D., Corbett, R. J., and Roche, R. S. (1986) *Biophys Chem* **23**(3-4), 183-199
11. Kawasaki, H., Nakayama, S., and Kretsinger, R. H. (1998) *Biomaterials* **11**(4), 277-295

12. Kretsinger, R. H., and Nockolds, C. E. (1973) *J Biol Chem* **248**(9), 3313-3326
13. Ravasi, T., Hsu, K., Goyette, J., Schroder, K., Yang, Z., Rahimi, F., Miranda, L. P., Alewood, P. F., Hume, D. A., and Geczy, C. (2004) *Genomics* **84**(1), 10-22
14. Ikura, M., and Ames, J. B. (2006) *Proc Natl Acad Sci U S A* **103**(5), 1159-1164
15. Perret, C., Lomri, N., and Thomasset, M. (1990) *Adv Exp Med Biol* **269**, 17-20
16. Persechini, A., Moncrief, N. D., and Kretsinger, R. H. (1989) *Trends Neurosci* **12**(11), 462-467
17. Rigden, D. J., and Galperin, M. Y. (2004) *J Mol Biol* **343**(4), 971-984
18. Pal, G. P., Elce, J. S., and Jia, Z. (2001) *J Biol Chem* **276**(50), 47233-47238
19. Raser, K. J., Buroker-Kilgore, M., and Wang, K. K. (1996) *Biochim Biophys Acta* **1292**(1), 9-14
20. Ravulapalli, R., Diaz, B. G., Campbell, R. L., and Davies, P. L. (2005) *Biochem J* **388**(Pt 2), 585-591
21. Hunter, M. J., and Chazin, W. J. (1998) *J Biol Chem* **273**(20), 12427-12435
22. Yap, K. L., Ames, J. B., Swindells, M. B., and Ikura, M. (1999) *Proteins* **37**(3), 499-507
23. Rigden, D. J., Jedrzejewski, M. J., and Galperin, M. Y. (2003) *FEMS Microbiol*

- Lett* **221**(1), 103-110
24. Vyas, N. K., Vyas, M. N., and Quiocho, F. A. (1987) *Nature* **327**(6123), 635-638
 25. Yang, K. (2001) *J Mol Microbiol Biotechnol* **3**(3), 457-459
 26. Momma, K., Mikami, B., Mishima, Y., Hashimoto, W., and Murata, K. (2002) *J Mol Biol* **316**(5), 1051-1059
 27. Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H., and Liddington, R. C. (1997) *Nature* **385**(6619), 833-838
 28. Lytle, B. L., Volkman, B. F., Westler, W. M., Heckman, M. P., and Wu, J. H. (2001) *J Mol Biol* **307**(3), 745-753
 29. Ye, Y., Lee, H. W., Yang, W., and Yang, J. J. (2005) *J Inorg Biochem* **99**(6), 1376-1383
 30. Chami, M., Oules, B., and Paterlini-Brechot, P. (2006) *Biochim Biophys Acta* **1763**(11), 1344-1362
 31. Haynes, J. I., 2nd, Chang, D., and Consigli, R. A. (1993) *Journal of virology* **67**(5), 2486-2495
 32. Adamec, T., Palkova, Z., Velkova, K., Stokrova, J., and Forstova, J. (2005) *FEMS Yeast Res* **5**(4-5), 331-340
 33. Pattanayek, R., Elrod, M., and Stubbs, G. (1992) *Proteins* **12**(2), 128-132
 34. Li, P. P., Naknanishi, A., Tran, M. A., Ishizu, K., Kawano, M., Phillips, M., Handa, H., Liddington, R. C., and Kasamatsu, H. (2003) *Journal of virology* **77**(13), 7527-7538
 35. Ishizu, K. I., Watanabe, H., Han, S. I., Kanesashi, S. N., Hoque, M., Yajima,

- H., Kataoka, K., and Handa, H. (2001) *Journal of virology* **75**(1), 61-72
36. Dong, Y., Zeng, C. Q., Ball, J. M., Estes, M. K., and Morris, A. P. (1997) *Proc Natl Acad Sci U S A* **94**(8), 3960-3965
 37. Dormitzer, P. R., Greenberg, H. B., and Harrison, S. C. (2000) *Virology* **277**(2), 420-428
 38. Emslie, K. R., Coukell, M. B., Birch, D., and Williams, K. L. (1996) *Journal of biotechnology* **50**(2-3), 149-159
 39. Gajardo, R., Vende, P., Poncet, D., and Cohen, J. (1997) *Journal of virology* **71**(3), 2211-2216
 40. Choi, Y., Gyoo Park, S., Yoo, J. H., and Jung, G. (2005) *Virology* **332**(1), 454-463
 41. Laakso, M. M., and Heaton, L. A. (1993) *Virology* **197**(2), 774-777
 42. Lin, B., and Heaton, L. A. (1999) *Virology* **259**(1), 34-42
 43. Gallagher, W. H., and Lauffer, M. A. (1983) *J Mol Biol* **170**(4), 921-929
 44. Ebenbichler, C. F., Stoiber, H., Schneider, R., Patsch, J. R., and Dierich, M. P. (1996) *Journal of virology* **70**(3), 1723-1728
 45. Dimitrov, D. S., Broder, C. C., Berger, E. A., and Blumenthal, R. (1993) *Journal of virology* **67**(3), 1647-1652
 46. Burmeister, W. P., Cusack, S., and Ruigrok, R. W. (1994) *J Gen Virol* **75** (Pt 2), 381-388
 47. Chong, A. K., Pegg, M. S., and von Itzstein, M. (1991) *Biochim Biophys Acta* **1077**(1), 65-71
 48. Cheung, W. Y. (1970) *Adv Biochem Psychopharmacol* **3**, 51-65

49. Lu, K. P., and Means, A. R. (1993) *Endocr Rev* **14**(1), 40-58
50. Means, A. R. (1994) *FEBS Lett* **347**(1), 1-4
51. Persechini, A., and Stemmer, P. M. (2002) *Trends Cardiovasc Med* **12**(1), 32-37
52. Bachs, O., Agell, N., and Carafoli, E. (1992) *Biochim Biophys Acta* **1113**(2), 259-270
53. Hook, S. S., and Means, A. R. (2001) *Annu Rev Pharmacol Toxicol* **41**, 471-505
54. Yamniuk, A. P., and Vogel, H. J. (2004) *Mol Biotechnol* **27**(1), 33-57
55. Saimi, Y., and Kung, C. (1994) *FEBS Lett* **350**(2-3), 155-158
56. Saimi, Y., and Kung, C. (2002) *Annu Rev Physiol* **64**, 289-311
57. Ye, Y., Lee, H. W., Yang, W., Shealy, S. J., Wilkins, A. L., Liu, Z. R., Torshin, I., Harrison, R., Wohlhueter, R., and Yang, J. J. (2001) *Protein Eng* **14**(12), 1001-1013
58. Pugachev, K. V., Tzeng, W. P., and Frey, T. K. (2000) *Journal of virology* **74**(22), 10811-10815
59. Tzeng, W. P., and Frey, T. K. (2003) *Journal of virology* **77**(17), 9502-9510
60. Mayer, B. J. (2001) *J Cell Sci* **114**(Pt 7), 1253-1263
61. Driscoll, P. C., Cyster, J. G., Campbell, I. D., and Williams, A. F. (1991) *Nature* **353**(6346), 762-765
62. Gill, S. C., and von Hippel, P. H. (1989) *Anal Biochem* **182**(2), 319-326
63. Fruen, B. R., Black, D. J., Bloomquist, R. A., Bardy, J. M., Johnson, J. D., Louis, C. F., and Balog, E. M. (2003) *Biochemistry* **42**(9), 2740-2747

64. Wallace, R. W., Tallant, E. A., and Cheung, W. Y. (1983) *Methods Enzymol* **102**, 39-47
65. Johnson, J. D., and Wittenauer, L. A. (1983) *Biochem J* **211**(2), 473-479
66. Whitmore, L., and Wallace, B. A. (2004) *Nucleic Acids Res* **32**(Web Server issue), W668-673
67. Maniccia, A. W., Yang, W., Li, S. Y., Johnson, J. A., and Yang, J. J. (2006) *Biochemistry* **45**(18), 5848-5856
68. Lurtz, M. M., and Louis, C. F. (2003) *Am J Physiol Cell Physiol* **285**(6), C1475-1482
69. Burr, G. S., Mitchell, C. K., Keflemariam, Y. J., Heidelberger, R., and O'Brien, J. (2005) *Biochem Biophys Res Commun* **335**(4), 1191-1198
70. Lee, H. W., Yang, W., Ye, Y., Liu, Z. R., Glushka, J., and Yang, J. J. (2002) *Biochim Biophys Acta* **1598**(1-2), 80-87
71. Tzeng, W. P., Matthews, J. D., and Frey, T. K. (2006) *Journal of virology* **80**(8), 3966-3974
72. Wang, C. Y., Dominguez, G., and Frey, T. K. (1994) *Journal of virology* **68**(6), 3550-3557
73. Pugachev, K. V., Abernathy, E. S., and Frey, T. K. (1997) *Journal of virology* **71**(1), 562-568
74. Tzeng, W. P., and Frey, T. K. (2002) *Journal of virology* **76**(7), 3189-3201
75. Churchill, G. C., Lurtz, M. M., and Louis, C. F. (2001) *Am J Physiol Cell Physiol* **281**(3), C972-981
76. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids*

Res **22**(22), 4673-4680

77. Page, R. D. (1996) *Comput Appl Biosci* **12**(4), 357-358
78. Eddy, S. R. (1998) *Bioinformatics* **14**(9), 755-763
79. Schuster-Bockler, B., Schultz, J., and Rahmann, S. (2004) *BMC Bioinformatics* **5**(1), 7
80. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) *Bioinformatics* **16**(4), 404-405
81. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) *Bioinformatics* **14**(10), 892-893
82. Andre, I., Kesvatera, T., Jonsson, B., Akerfeldt, K. S., and Linse, S. (2004) *Biophys J* **87**(3), 1929-1938
83. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) *Nucleic Acids Res* **31**(13), 3381-3385
84. Guarne, A., Hampoelz, B., Glaser, W., Carpena, X., Tormo, J., Fita, I., and Skern, T. (2000) *J Mol Biol* **302**(5), 1227-1240
85. Deng, H., Chen, G., Yang, W., and Yang, J. J. (2006) *Proteins*
86. Mitaku, S., Hirokawa, T., and Tsuji, T. (2002) *Bioinformatics* **18**(4), 608-616
87. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) *J Mol Biol* **305**(3), 567-580
88. Jones, D. T. (1998) *FEBS Lett* **423**(3), 281-285
89. Tusnady, G. E., and Simon, I. (2001) *Bioinformatics* **17**(9), 849-850
90. Yap, K. L., Kim, J., Truong, K., Sherman, M., Yuan, T., and Ikura, M. (2000) *J Struct Funct Genomics* **1**(1), 8-14

91. Hunt, J. B., Neece, S. H., and Ginsburg, A. (1985) *Anal Biochem* **146**(1), 150-157
92. Rashidi, H. H., Bauer, M., Patterson, J., and Smith, D. W. (1999) *J Mol Microbiol Biotechnol* **1**(1), 175-182
93. Dragani, B., and Aceto, A. (1999) *Arch Biochem Biophys* **368**(2), 211-213
94. Krieg, P., Schuppler, M., Koesters, R., Mincheva, A., Lichter, P., and Marks, F. (1997) *Genomics* **43**(3), 339-348
95. Makino, T., Takaishi, M., Morohashi, M., and Huh, N. H. (2001) *J Biol Chem* **276**(50), 47445-47452
96. Huber, M., Siegenthaler, G., Mirancea, N., Marenholz, I., Nizetic, D., Breitskreutz, D., Mischke, D., and Hohl, D. (2005) *J Invest Dermatol* **124**(5), 998-1007
97. Lee, S. C., Kim, I. G., Marekov, L. N., O'Keefe, E. J., Parry, D. A., and Steinert, P. M. (1993) *J Biol Chem* **268**(16), 12164-12176
98. Donato, R. (1999) *Biochim Biophys Acta* **1450**(3), 191-231
99. Donato, R. (2001) *Int J Biochem Cell Biol* **33**(7), 637-668
100. Heizmann, C. W., and Cox, J. A. (1998) *Biometals* **11**(4), 383-397
101. Tanaka, T., Miwa, N., Kawamura, S., Sohma, H., Nitta, K., and Matsushima, N. (1999) *Protein Eng* **12**(5), 395-405
102. Miwa, N., Kobayashi, M., Takamatsu, K., and Kawamura, S. (1998) *Biochem Biophys Res Commun* **251**(3), 860-867
103. Leung, A. K., Duewel, H. S., Honek, J. F., and Berghuis, A. M. (2001) *Biochemistry* **40**(19), 5665-5673

104. van Asselt, E. J., and Dijkstra, B. W. (1999) *FEBS Lett* **458**(3), 429-435
105. Wu, C. H., Yeh, L. S., Huang, H., Arminski, L., Castro-Alvear, J., Chen, Y., Hu, Z., Kourtesis, P., Ledley, R. S., Suzek, B. E., Vinayaka, C. R., Zhang, J., and Barker, W. C. (2003) *Nucleic Acids Res* **31**(1), 345-347
106. Aitio, H., Annila, A., Heikkinen, S., Thulin, E., Drakenberg, T., and Kilpelainen, I. (1999) *Protein Sci* **8**(12), 2580-2588
107. Gangola, P., and Rosen, B. P. (1987) *J Biol Chem* **262**(26), 12570-12574
108. Herbaud, M. L., Guiseppi, A., Denizot, F., Haiech, J., and Kilhoffer, M. C. (1998) *Biochim Biophys Acta* **1448**(2), 212-226
109. Kalckar, H. M. (1976) *Biochimie* **58**(1-2), 81-85
110. Tisa, L. S., and Adler, J. (1992) *Proc Natl Acad Sci U S A* **89**(24), 11804-11808
111. Carvalho, A. L., Dias, F. M., Prates, J. A., Nagy, T., Gilbert, H. J., Davies, G. J., Ferreira, L. M., Romao, M. J., and Fontes, C. M. (2003) *Proc Natl Acad Sci U S A* **100**(24), 13809-13814
112. Lin, J., and Ficht, T. A. (1995) *Infect Immun* **63**(4), 1409-1414
113. Teixeira-Gomes, A. P., Cloeckaert, A., and Zygmunt, M. S. (2000) *Infect Immun* **68**(5), 2954-2961
114. Leadly, P. F., Roberts, G., and Walker, J. E. (1984) *FEBS Letters* **178**(1), 157-160
115. Natsume, M., Yasui, K., and Marumo, S. (1989) *J Antibiot (Tokyo)* **42**(3), 440-447
116. Tossavainen, H., Permi, P., Annila, A., Kilpelainen, I., and Drakenberg, T.

- (2003) *Eur J Biochem* **270**(11), 2505-2512
117. Yonekawa, T., Ohnishi, Y., and Horinouchi, S. (2005) *FEMS Microbiol Lett* **244**(2), 315-321
 118. Xi, C., Schoeters, E., Vanderleyden, J., and Michiels, J. (2000) *Proc Natl Acad Sci U S A* **97**(20), 11114-11119
 119. Spurway, T. D., Morland, C., Cooper, A., Sumner, I., Hazlewood, G. P., O'Donnell, A. G., Pickersgill, R. W., and Gilbert, H. J. (1997) *J Biol Chem* **272**(28), 17523-17530
 120. Chen, L., Liu, M. Y., and Le Gall, J. (1991) *Biochem Biophys Res Commun* **180**(1), 238-242
 121. Lytle, B. L., Volkman, B. F., Westler, W. M., and Wu, J. H. (2000) *Arch Biochem Biophys* **379**(2), 237-244
 122. Kohler, U., Cerff, R., and Brinkmann, H. (1996) *Plant Mol Biol* **30**(1), 213-218
 123. Laarmann, S., and Schmidt, M. A. (2003) *Microbiology* **149**(Pt 7), 1871-1882
 124. Dean, M., and Allikmets, R. (1995) *Curr Opin Genet Dev* **5**(6), 779-785
 125. Higgins, C. F. (1992) *Annu Rev Cell Biol* **8**, 67-113
 126. van Veen, H. W., and Konings, W. N. (1997) *Semin Cancer Biol* **8**(3), 183-191
 127. Andrade, M. A., Ciccarelli, F. D., Perez-Iratxeta, C., and Bork, P. (2002) *Genome Biol* **3**(9), RESEARCH0047
 128. Bates, C. S., Montanez, G. E., Woods, C. R., Vincent, R. M., and

- Eichenbaum, Z. (2003) *Infect Immun* **71**(3), 1042-1055
129. Day, I. S., Reddy, V. S., Shad Ali, G., and Reddy, A. S. (2002) *Genome Biol* **3**(10), RESEARCH0056
 130. Julenius, K., Robblee, J., Thulin, E., Finn, B. E., Fairman, R., and Linse, S. (2002) *Proteins* **47**(3), 323-333
 131. Kawasaki, H., and Kretsinger, R. H. (1995) *Protein Profile* **2**(4), 297-490
 132. Franchini, P. L., and Reid, R. E. (1999) *J Theor Biol* **199**(2), 199-211
 133. Berggard, T., Julenius, K., Ogard, A., Drakenberg, T., and Linse, S. (2001) *Biochemistry* **40**(5), 1257-1264
 134. Wojcik, J., Goral, J., Pawlowski, K., and Bierzynski, A. (1997) *Biochemistry* **36**(4), 680-687
 135. Berggard, T., Thulin, E., Akerfeldt, K. S., and Linse, S. (2000) *Protein Sci* **9**(11), 2094-2108
 136. Kretsinger, R. H. (1976) *Int Rev Cytol* **46**, 323-393
 137. Kretsinger, R. H., and Moncrief, N. D. (1989) *Va Explor* **5**(5), 7-9
 138. Nakayama, S., and Kretsinger, R. H. (1994) *Annu Rev Biophys Biomol Struct* **23**, 473-507
 139. Ridinger, K., Ilg, E. C., Niggli, F. K., Heizmann, C. W., and Schafer, B. W. (1998) *Biochim Biophys Acta* **1448**(2), 254-263
 140. Franz, C., Durussel, I., Cox, J. A., Schafer, B. W., and Heizmann, C. W. (1998) *J Biol Chem* **273**(30), 18826-18834
 141. Fohr, U. G., Heizmann, C. W., Engelkamp, D., Schafer, B. W., and Cox, J. A. (1995) *J Biol Chem* **270**(36), 21056-21061

142. Schafer, B. W., Fritschy, J. M., Murmann, P., Troxler, H., Durussel, I., Heizmann, C. W., and Cox, J. A. (2000) *J Biol Chem* **275**(39), 30623-30630
143. Filipek, A., Heizmann, C. W., and Kuznicki, J. (1990) *FEBS Lett* **264**(2), 263-266
144. Heizmann, C. W. (2002) *Methods Mol Biol* **172**, 69-80
145. Clohessy, P. A., and Golden, B. E. (1996) *J Leukoc Biol* **60**(5), 674
146. Contzler, R., Favre, B., Huber, M., and Hohl, D. (2005) *J Invest Dermatol* **124**(5), 990-997
147. Maki, M., Narayana, S. V., and Hitomi, K. (1997) *Biochem J* **328** (Pt 2), 718-720
148. Palczewska, M., Groves, P., Batta, G., Heise, B., and Kuznicki, J. (2003) *Protein Sci* **12**(1), 180-184
149. Swan, D. G., Hale, R. S., Dhillon, N., and Leadlay, P. F. (1987) *Nature* **329**(6134), 84-85
150. Yonekawa, T., Ohnishi, Y., and Horinouchi, S. (2001) *Biosci Biotechnol Biochem* **65**(1), 156-160
151. Grishin, N. V. (2001) *J Struct Biol* **134**(2-3), 167-185
152. Fritz, G., Heizmann, C. W., and Kroneck, P. M. (1998) *Biochim Biophys Acta* **1448**(2), 264-276
153. Fritz, G., Mittl, P. R., Vasak, M., Grutter, M. G., and Heizmann, C. W. (2002) *J Biol Chem* **277**(36), 33092-33098
154. Berridge, M. J., Bootman, M. D., and Roderick, H. L. (2003) *Nat Rev Mol*

Cell Biol **4**(7), 517-529

155. Liu, M., Yang, Y., Gu, C., Yue, Y., Wu, K. K., Wu, J., and Zhu, Y. (2007)
Faseb J **21**(7), 1586-1596
156. Li, Y., Boehning, D. F., Qian, T., Popov, V. L., and Weinman, S. A. (2007)
Faseb J
157. Benali-Furet, N. L., Chami, M., Houel, L., De Giorgi, F., Vernejoul, F.,
Lagorce, D., Buscail, L., Bartenschlager, R., Ichas, F., Rizzuto, R., and
Paterlini-Brechot, P. (2005) *Oncogene* **24**(31), 4921-4933
158. Bergqvist, A., Sundstrom, S., Dimberg, L. Y., Gylfe, E., and Masucci, M. G.
(2003) *J Biol Chem* **278**(21), 18877-18883
159. Griffin, S. D., Beales, L. P., Clarke, D. S., Worsfold, O., Evans, S. D.,
Jaeger, J., Harris, M. P., and Rowlands, D. J. (2003) *FEBS Lett* **535**(1-3),
34-38
160. Gong, G., Waris, G., Tanveer, R., and Siddiqui, A. (2001) *Proc Natl Acad
Sci U S A* **98**(17), 9599-9604
161. Christen, V., Treves, S., Duong, F. H., and Heim, M. H. (2007) *Hepatology*
162. Dhawan, R., Chaturvedi, U. C., Khanna, M., Mathur, A., Tekwani, B. L.,
Pandey, V. C., and Rai, R. N. (1991) *International journal of experimental
pathology* **72**(1), 31-39
163. Khanna, M., Chaturvedi, U. C., Dhawan, R., Tekwani, B. L., and Pandey, V.
C. (1991) *Immunology* **72**(1), 73-78
164. Srivastava, S., Khanna, N., Mathur, A., and Chaturvedi, U. C. (1994) *Acta
virologica* **38**(2), 111-116

165. Irurzun, A., Arroyo, J., Alvarez, A., and Carrasco, L. (1995) *Journal of virology* **69**(8), 5142-5146
166. Aldabe, R., Irurzun, A., and Carrasco, L. (1997) *Journal of virology* **71**(8), 6214-6217
167. Campanella, M., de Jong, A. S., Lanke, K. W., Melchers, W. J., Willems, P. H., Pinton, P., Rizzuto, R., and van Kuppeveld, F. J. (2004) *J Biol Chem* **279**(18), 18440-18450
168. van Kuppeveld, F. J., de Jong, A. S., Melchers, W. J., and Willems, P. H. (2005) *Trends Microbiol* **13**(2), 41-44
169. van Kuppeveld, F. J., Hoenderop, J. G., Smeets, R. L., Willems, P. H., Dijkman, H. B., Galama, J. M., and Melchers, W. J. (1997) *Embo J* **16**(12), 3519-3532
170. Gazina, E. V., Harrison, D. N., Jefferies, M., Tan, H., Williams, D., Anderson, D. A., and Petrou, S. (2005) *Antiviral research* **67**(2), 98-106
171. Peterhans, E., Haenggeli, E., Wild, P., and Wyler, R. (1979) *Journal of virology* **29**(1), 143-152
172. Ubol, S., Park, S., Budihardjo, I., Desnoyers, S., Montrose, M. H., Poirier, G. G., Kaufmann, S. H., and Griffin, D. E. (1996) *Journal of virology* **70**(4), 2215-2220
173. Mohamadzadeh, M., Coberley, S. S., Olinger, G. G., Kalina, W. V., Ruthel, G., Fuller, C. L., Swenson, D. L., Pratt, W. D., Kuhns, D. B., and Schmaljohn, A. L. (2006) *Journal of virology* **80**(14), 7235-7244
174. Brask, J., Owe-Larsson, B., Hill, R. H., and Kristensson, K. (2001) *Brain*

research bulletin **55**(3), 421-429

175. Hartshorn, K. L., Collamer, M., White, M. R., Schwartz, J. H., and Tauber, A. I. (1990) *Blood* **75**(1), 218-226
176. Hartshorn, K. L., Collamer, M., Auerbach, M., Myers, J. B., Pavlotsky, N., and Tauber, A. I. (1988) *J Immunol* **141**(4), 1295-1301
177. Shainkin-Kestenbaum, R., Winikoff, Y., Chaimovitz, C., Zimlichman, S., and Sarov, I. (1993) *Israel journal of medical sciences* **29**(1), 2-6
178. Hallett, M. B., Fuchs, P., and Campbell, A. K. (1982) *Biochem J* **206**(3), 671-674
179. Volsky, D. J., and Loyter, A. (1978) *J Cell Biol* **78**(2), 465-479
180. Maehlen, J., Wallen, P., Love, A., Norrby, E., and Kristensson, K. (1991) *Brain research* **540**(1-2), 123-130
181. Owe-Larsson, B., Andersson, T., Kristensson, K., and Hill, R. H. (1997) *Journal of neurovirology* **3**(5), 369-379
182. Manninen, A., and Saksela, K. (2002) *The Journal of experimental medicine* **195**(8), 1023-1032
183. Foti, M., Cartier, L., Piguet, V., Lew, D. P., Carpentier, J. L., Trono, D., and Krause, K. H. (1999) *J Biol Chem* **274**(49), 34765-34772
184. Contreras, X., Bennasser, Y., Chazal, N., Moreau, M., Leclerc, C., Tkaczuk, J., and Bahraoui, E. (2005) *Virology* **332**(1), 316-328
185. Self, R. L., Mulholland, P. J., Nath, A., Harris, B. R., and Prendergast, M. A. (2004) *Brain research* **995**(1), 39-45
186. Haughey, N. J., Holden, C. P., Nath, A., and Geiger, J. D. (1999) *Journal of*

neurochemistry **73**(4), 1363-1374

187. Sheng, W. S., Hu, S., Hegg, C. C., Thayer, S. A., and Peterson, P. K.
(2000) *Clinical immunology (Orlando, Fla)* **96**(3), 243-251
188. Campbell, G. R., Watkins, J. D., Singh, K. K., Loret, E. P., and Spector, S.
A. (2007) *Journal of virology* **81**(11), 5919-5928
189. Koller, H., Schaal, H., Freund, M., Garrido, S. R., von Giesen, H. J., Ott,
M., Rosenbaum, C., and Arendt, G. (2001) *The European journal of*
neuroscience **14**(11), 1793-1799
190. Zocchi, M. R., Rubartelli, A., Morgavi, P., and Poggi, A. (1998) *J Immunol*
161(6), 2938-2943
191. Holden, C. P., Haughey, N. J., Nath, A., and Geiger, J. D. (1999)
Neuroscience **91**(4), 1369-1378
192. Melar, M., Ott, D. E., and Hope, T. J. (2007) *Journal of virology* **81**(4),
1773-1785
193. Lannuzel, A., Lledo, P. M., Lamghitnia, H. O., Vincent, J. D., and Tardieu,
M. (1995) *The European journal of neuroscience* **7**(11), 2285-2293
194. Nath, A., Padua, R. A., and Geiger, J. D. (1995) *Brain research* **678**(1-2),
200-206
195. Dayanithi, G., Yah, N., Baghdiguian, S., and Fantini, J. (1995) *Cell*
calcium **18**(1), 9-18
196. Nye, K. E., and Pinching, A. J. (1990) *AIDS (London, England)* **4**(1), 41-45
197. Nokta, M. A., Hassan, M. I., Morgan, J. A., Loesch, K. A., and Pollard, R. B.
(1994) *Proceedings of the Society for Experimental Biology and Medicine.*

- Society for Experimental Biology and Medicine (New York, N.Y.* **207**(3), 284-291
198. Ding, W., Albrecht, B., Kelley, R. E., Muthusamy, N., Kim, S. J., Altschuld, R. A., and Lairmore, M. D. (2002) *Journal of virology* **76**(20), 10374-10382
 199. Nair, A. M., Michael, B., Datta, A., Fernandez, S., and Lairmore, M. D. (2006) *Virology* **353**(2), 247-257
 200. D'Agostino, D. M., Ranzato, L., Arrigoni, G., Cavallari, I., Belleudi, F., Torrisi, M. R., Silic-Benussi, M., Ferro, T., Petronilli, V., Marin, O., Chieco-Bianchi, L., Bernardi, P., and Ciminale, V. (2002) *J Biol Chem* **277**(37), 34424-34433
 201. D'Agostino, D. M., Silic-Benussi, M., Hilaragi, H., Lairmore, M. D., and Ciminale, V. (2005) *Cell Death Differ* **12 Suppl 1**, 905-915
 202. D'Agostino, D. M., Zotti, L., Ferro, T., Franchini, G., Chieco-Bianchi, L., and Ciminale, V. (2000) *AIDS research and human retroviruses* **16**(16), 1765-1770
 203. Ruiz, M. C., Diaz, Y., Pena, F., Aristimuno, O. C., Chemello, M. E., and Michelangeli, F. (2005) *Virology* **333**(1), 54-65
 204. Brunet, J. P., Cotte-Laffitte, J., Linxe, C., Quero, A. M., Geniteau-Legendre, M., and Servin, A. (2000) *Journal of virology* **74**(5), 2323-2332
 205. Perez, J. F., Ruiz, M. C., Chemello, M. E., and Michelangeli, F. (1999) *Journal of virology* **73**(3), 2481-2490
 206. Tian, P., Estes, M. K., Hu, Y., Ball, J. M., Zeng, C. Q., and Schilling, W. P. (1995) *Journal of virology* **69**(9), 5763-5772

207. Bounous, D. I., Goodwin, M. A., Brooks, R. L., Jr., Lamichhane, C. M., Campagnoli, R. P., Brown, J., and Snyder, D. B. (1995) *Avian diseases* **39**(1), 135-140
208. Lupescu, A., Bock, C. T., Lang, P. A., Aberle, S., Kaiser, H., Kandolf, R., and Lang, F. (2006) *Journal of virology* **80**(22), 11370-11380
209. Sharon-Friling, R., Goodhouse, J., Colberg-Poley, A. M., and Shenk, T. (2006) *Proc Natl Acad Sci U S A* **103**(50), 19117-19122
210. Nokta, M., Eaton, D., Steinsland, O. S., and Albrecht, T. (1987) *Virology* **157**(2), 259-267
211. Cheshenko, N., Del Rosario, B., Woda, C., Marcellino, D., Satlin, L. M., and Herold, B. C. (2003) *J Cell Biol* **163**(2), 283-293
212. Arimoto, E., Iwai, S., Sumi, T., Ogawa, Y., and Yura, Y. (2006) *Virology* **3**, 62
213. Lee, B. S., Lee, S. H., Feng, P., Chang, H., Cho, N. H., and Jung, J. U. (2005) *Journal of virology* **79**(19), 12173-12184
214. Nakano, K., Isegawa, Y., Zou, P., Tadagaki, K., Inagi, R., and Yamanishi, K. (2003) *Archives of virology* **148**(5), 871-890
215. Feng, P., Park, J., Lee, B. S., Lee, S. H., Bram, R. J., and Jung, J. U. (2002) *Journal of virology* **76**(22), 11491-11504
216. Dugas, B., Mencia-Huerta, J. M., Braquet, P., Galanaud, P., and Delfraissy, J. F. (1989) *European journal of immunology* **19**(10), 1867-1871
217. Miller, C. L., Longnecker, R., and Kieff, E. (1993) *Journal of virology* **67**(6), 3087-3094
218. Newcomb, T. G., Mullins, R. D., and Siskin, J. E. (1993) *Cell calcium*

- 14(7)**, 539-549
219. Garrett, L. R., Coder, D. M., and McDougall, J. K. (1991) *Cell calcium*
12(5), 343-349
 220. Sanderson, C. M., Parkinson, J. E., Hollinshead, M., and Smith, G. L.
(1996) *Journal of virology* **70(2)**, 905-914
 221. Oliver, K. G., Buller, R. M., Hughes, P. J., Putney, J. W., Jr., and Palumbo,
G. J. (1992) *J Biol Chem* **267(35)**, 25098-25103
 222. Oh, J. C., Jeong, D. L., Kim, I. K., and Oh, S. H. (2003) *Exp Mol Med*
35(4), 301-309
 223. Chami, M., Ferrari, D., Nicotera, P., Paterlini-Brechot, P., and Rizzuto, R.
(2003) *J Biol Chem* **278(34)**, 31745-31755
 224. Bouchard, M. J., Wang, L. H., and Schneider, R. J. (2001) *Science*
294(5550), 2376-2378
 225. de Jong, A. S., Visch, H. J., de Mattia, F., van Dommelen, M. M., Swarts,
H. G., Luyten, T., Callewaert, G., Melchers, W. J., Willems, P. H., and van
Kuppeveld, F. J. (2006) *J Biol Chem* **281(20)**, 14144-14150
 226. Kamoshita, K., Shiota, M., Sasaki, M., Koga, Y., Okumura, Y., and Kido, H.
(1995) *J Biochem (Tokyo)* **117(6)**, 1244-1253
 227. Oldridge, J., and Marsh, M. (1998) *Trends in cell biology* **8(8)**, 302-305
 228. Kinoshita, S., Su, L., Amano, M., Timmerman, L. A., Kaneshima, H., and
Nolan, G. P. (1997) *Immunity* **6(3)**, 235-244
 229. Ruiz, M. C., Cohen, J., and Michelangeli, F. (2000) *Cell calcium* **28(3)**,
137-149

230. Parekh, A. B., and Penner, R. (1997) *Physiological reviews* **77**(4), 901-930
231. Berkova, Z., Morris, A. P., and Estes, M. K. (2003) *Cell calcium* **34**(1), 55-68
232. Dugas, B., Delfraissy, J. F., Calenda, A., Peuchmaur, M., Wallon, C., Rannou, M. T., and Galanaud, P. (1988) *J Immunol* **141**(12), 4344-4351
233. Fruehling, S., and Longnecker, R. (1997) *Virology* **235**(2), 241-251
234. Cesarman, E., Chang, Y., Moore, P. S., Said, J. W., and Knowles, D. M. (1995) *The New England journal of medicine* **332**(18), 1186-1191
235. Chang, Y., Cesarman, E., Pessin, M. S., Lee, F., Culpepper, J., Knowles, D. M., and Moore, P. S. (1994) *Science* **266**(5192), 1865-1869
236. Dormitzer, P. R., and Greenberg, H. B. (1992) *Virology* **189**(2), 828-832
237. Burmeister, W. P., Ruigrok, R. W., and Cusack, S. (1992) *Embo J* **11**(1), 49-56
238. Hogle, J. M., Maeda, A., and Harrison, S. C. (1986) *J Mol Biol* **191**(4), 625-638
239. Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L., and Harrison, S. C. (1991) *Nature* **354**(6351), 278-284
240. Matsumura, T., and Yamashita, H. (1978) *Microbiology and immunology* **22**(12), 803-807
241. Giranda, V. L., Heinz, B. A., Oliveira, M. A., Minor, I., Kim, K. H., Kolatkar, P. R., Rossmann, M. G., and Rueckert, R. R. (1992) *Proc Natl Acad Sci U S A* **89**(21), 10213-10217
242. Zhao, R., Pevear, D. C., Kremer, M. J., Giranda, V. L., Kofron, J. A., Kuhn,

- R. J., and Rossmann, M. G. (1996) *Structure* **4**(10), 1205-1220
243. Xiao, C., Bator-Kelly, C. M., Rieder, E., Chipman, P. R., Craig, A., Kuhn, R. J., Wimmer, E., and Rossmann, M. G. (2005) *Structure* **13**(7), 1019-1033
244. Muckelbauer, J. K., Kremer, M., Minor, I., Diana, G., Dutko, F. J., Groarke, J., Pevear, D. C., and Rossmann, M. G. (1995) *Structure* **3**(7), 653-667
245. Bishop, N. E., and Anderson, D. A. (1997) *Archives of virology* **142**(11), 2161-2178
246. Stapleton, J. T., Frederick, J., and Meyer, B. (1991) *The Journal of infectious diseases* **164**(6), 1098-1103
247. Burroughs, S. E., Eisenman, G., and Horrocks, W. D., Jr. (1992) *Biophys Chem* **42**(3), 249-256
248. Jones, T. A., and Liljas, L. (1984) *J Mol Biol* **177**(4), 735-767
249. Montelius, I., Liljas, L., and Unge, T. (1990) *J Mol Biol* **212**(2), 331-343
250. Zhou, Y., Tzeng, W. P., Yang, W., Zhou, Y., Ye, Y., Lee, H. W., Frey, T. K., and Yang, J. (2007) *Journal of virology*
251. Hopper, P., Harrison, S. C., and Sauer, R. T. (1984) *J Mol Biol* **177**(4), 701-713
252. Campbell, J. W., Clifton, I. J., Greenhough, T. J., Hajdu, J., Harrison, S. C., Liddington, R. C., and Shrive, A. K. (1990) *J Mol Biol* **214**(3), 627-632
253. Kruse, J., Kruse, K. M., Witz, J., Chauvin, C., Jacrot, B., and Tardieu, A. (1982) *J Mol Biol* **162**(2), 393-414
254. Satheshkumar, P. S., Lokesh, G. L., Sangita, V., Saravanan, V., Vijay, C. S., Murthy, M. R., and Savithri, H. S. (2004) *J Mol Biol* **342**(3), 1001-1014

255. Tars, K., Zeltins, A., and Liljas, L. (2003) *Virology* **310**(2), 287-297
256. Erickson, J. W., Silva, A. M., Murthy, M. R., Fita, I., and Rossmann, M. G. (1985) *Science* **229**(4714), 625-629
257. Silva, A. M., and Rossmann, M. G. (1987) *J Mol Biol* **197**(1), 69-87
258. Sherman, M. B., Guenther, R. H., Tama, F., Sit, T. L., Brooks, C. L., Mikhailov, A. M., Orlova, E. V., Baker, T. S., and Lommel, S. A. (2006) *Journal of virology* **80**(21), 10395-10406
259. Oda, Y., Saeki, K., Takahashi, Y., Maeda, T., Naitow, H., Tsukihara, T., and Fukuyama, K. (2000) *J Mol Biol* **300**(1), 153-169
260. Suarez, T., Gomara, M. J., Goni, F. M., Mingarro, I., Muga, A., Perez-Paya, E., and Nieva, J. L. (2003) *FEBS Lett* **535**(1-3), 23-28
261. Kaukinen, P., Koistinen, V., Vapalahti, O., Vaheri, A., and Plyusnin, A. (2001) *J Gen Virol* **82**(Pt 8), 1845-1853
262. Tulip, W. R., Varghese, J. N., Baker, A. T., van Donkelaar, A., Laver, W. G., Webster, R. G., and Colman, P. M. (1991) *J Mol Biol* **221**(2), 487-497
263. Smith, B. J., Huyton, T., Joosten, R. P., McKimm-Breschkin, J. L., Zhang, J. G., Luo, C. S., Lou, M. Z., Labrou, N. E., and Garrett, T. P. (2006) *Acta Crystallogr D Biol Crystallogr* **62**(Pt 9), 947-952
264. Varghese, J. N., Laver, W. G., and Colman, P. M. (1983) *Nature* **303**(5912), 35-40
265. Johansson, B. E., and Brett, I. C. (2003) *J Biochem (Tokyo)* **134**(3), 345-352
266. Brett, I. C., and Johansson, B. E. (2006) *J Biochem (Tokyo)* **139**(3), 439-

447

267. Yuan, P., Thompson, T. B., Wurzburg, B. A., Paterson, R. G., Lamb, R. A., and Jardetzky, T. S. (2005) *Structure* **13**(5), 803-815
268. Lawrence, M. C., Borg, N. A., Streltsov, V. A., Pilling, P. A., Epa, V. C., Varghese, J. N., McKimm-Breschkin, J. L., and Colman, P. M. (2004) *J Mol Biol* **335**(5), 1343-1357
269. Crennell, S., Takimoto, T., Portner, A., and Taylor, G. (2000) *Nature structural biology* **7**(11), 1068-1074
270. Kamata, M., Hiraki, A., and Kim, J. (1994) *Cell structure and function* **19**(5), 315-323
271. Kim, J., and Kamata, M. (1994) *Cell structure and function* **19**(5), 325-333
272. Ono, A., and Kawakita, M. (1994) *J Biochem (Tokyo)* **116**(3), 649-656
273. Tarbouriech, N., Curran, J., Ruigrok, R. W., and Burmeister, W. P. (2000) *Nature structural biology* **7**(9), 777-781
274. Shahrabadi, M. S., and Lee, P. W. (1988) *Journal of clinical microbiology* **26**(1), 139-141
275. Saez-Cirion, A., and Nieva, J. L. (2002) *Biochim Biophys Acta* **1564**(1), 57-65
276. Yu, H., Alfsen, A., Tudor, D., and Bomsel, M. (2007) *Cell calcium*
277. Moulard, M., Montagnier, L., and Bahraoui, E. (1994) *FEBS Lett* **338**(3), 281-284
278. Anzinger, J. J., Mezo, I., Ji, X., Gabali, A. M., Thomas, L. L., and Spear, G. T. (2006) *Virus research* **122**(1-2), 183-188

279. Larsen, C. E., Nir, S., Alford, D. R., Jennings, M., Lee, K. D., and Duzgunes, N. (1993) *Biochim Biophys Acta* **1147**(2), 223-236
280. Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1996) *Structure* **4**(1), 89-96
281. Bujacz, G., Jaskolski, M., Alexandratos, J., Wlodawer, A., Merkel, G., Katz, R. A., and Skalka, A. M. (1995) *J Mol Biol* **253**(2), 333-346
282. Bondzio, A., Abraham-Podgornik, A., Blankenstein, P., and Risse, S. (2001) *Biol Chem* **382**(3), 407-416
283. Garriga, D., Querol-Audi, J., Abaitua, F., Saugar, I., Pous, J., Verdaguer, N., Caston, J. R., and Rodriguez, J. F. (2006) *Journal of virology* **80**(14), 6895-6905
284. Lee, C. C., Ko, T. P., Chou, C. C., Yoshimura, M., Doong, S. R., Wang, M. Y., and Wang, A. H. (2006) *J Struct Biol* **155**(1), 74-86
285. Mathieu, M., Petitpas, I., Navaza, J., Lepault, J., Kohli, E., Pothier, P., Prasad, B. V., Cohen, J., and Rey, F. A. (2001) *Embo J* **20**(7), 1485-1497
286. Bowman, G. D., Nodelman, I. M., Levy, O., Lin, S. L., Tian, P., Zamb, T. J., Udem, S. A., Venkataraghavan, B., and Schutt, C. E. (2000) *J Mol Biol* **304**(5), 861-871
287. Simpson, A. A., Chandrasekar, V., Hebert, B., Sullivan, G. M., Rossmann, M. G., and Parrish, C. R. (2000) *J Mol Biol* **300**(3), 597-610
288. Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W., and et al. (1991) *Science* **251**(5000), 1456-1464

- 289. Wu, W., Booth, J. L., Coggeshall, K. M., and Metcalf, J. P. (2006) *Virology* **355**(1), 18-29
- 290. Nilsson, J., Miyazaki, N., Xing, L., Wu, B., Hammar, L., Li, T. C., Takeda, N., Miyamura, T., and Cheng, R. H. (2005) *Journal of virology* **79**(9), 5337-5345
- 291. Stehle, T., Gamblin, S. J., Yan, Y., and Harrison, S. C. (1996) *Structure* **4**(2), 165-182
- 292. Li, P. P., Nguyen, A. P., Qu, Q., Jafri, Q. H., Aungsumart, S., Cheng, R. H., and Kasamatsu, H. (2007) *Journal of virology* **81**(11), 6099-6105
- 293. Rodgers, R. E., and Consigli, R. A. (1996) *Virus research* **44**(2), 123-135
- 294. Ilag, L. L., McKenna, R., Yadav, M. P., BeMiller, J. N., Incardona, N. L., and Rossmann, M. G. (1994) *J Mol Biol* **244**(3), 291-300
- 295. McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S., and Incardona, N. L. (1992) *Nature* **355**(6356), 137-143
- 296. McKenna, R., Bowman, B. R., Ilag, L. L., Rossmann, M. G., and Fane, B. A. (1996) *J Mol Biol* **256**(4), 736-750
- 297. Landry, E. F., and Zsigray, R. M. (1980) *J Gen Virol* **51**(Pt 1), 125-135
- 298. Tang, J. X., Janmey, P. A., Lyubartsev, A., and Nordenskiold, L. (2002) *Biophys J* **83**(1), 566-581
- 299. Wen, Q., and Tang, J. X. (2004) *The Journal of chemical physics* **121**(24), 12666-12670
- 300. Kim, S. S., Smith, T. J., Chapman, M. S., Rossmann, M. C., Pevear, D. C.,

- Dutko, F. J., Felock, P. J., Diana, G. D., and McKinlay, M. A. (1989) *J Mol Biol* **210**(1), 91-111
301. Kasamatsu, H., Woo, J., Nakamura, A., Muller, P., Tevethia, M. J., and Liddington, R. C. (2006) *Protein Sci* **15**(9), 2207-2213
 302. Yang, W., Lee, H. W., Hellinga, H., and Yang, J. J. (2002) *Proteins* **47**(3), 344-356
 303. Stathopoulos, P. B., Li, G. Y., Plevin, M. J., Ames, J. B., and Ikura, M. (2006) *J Biol Chem* **281**(47), 35855-35862
 304. Nakhasi, H. L., Ramanujam, M., Atreya, C. D., Hobman, T. C., Lee, N., Esmaili, A., and Duncan, R. C. (2001) *Archives of virology* **146**(1), 1-14
 305. Gelebart, P., Opas, M., and Michalak, M. (2005) *Int J Biochem Cell Biol* **37**(2), 260-266
 306. Ellgaard, L., and Frickel, E. M. (2003) *Cell biochemistry and biophysics* **39**(3), 223-247
 307. Singh, N. K., Atreya, C. D., and Nakhasi, H. L. (1994) *Proc Natl Acad Sci U S A* **91**(26), 12770-12774
 308. Atreya, C. D., Singh, N. K., and Nakhasi, H. L. (1995) *Journal of virology* **69**(6), 3848-3851
 309. Mayran, N., Parton, R. G., and Gruenberg, J. (2003) *Embo J* **22**(13), 3242-3253
 310. Hoefflich, K. P., and Ikura, M. (2002) *Cell* **108**(6), 739-742
 311. Hayashi, N., Matsubara, M., Jinbo, Y., Titani, K., Izumi, Y., and Matsushima, N. (2002) *Protein Sci* **11**(3), 529-537

312. Matsubara, M., Jing, T., Kawamura, K., Shimojo, N., Titani, K., Hashimoto, K., and Hayashi, N. (2005) *Protein Sci* **14**(2), 494-503
313. Pan, Z., Radding, W., Zhou, T., Hunter, E., Mountz, J., and McDonald, J. M. (1996) *The American journal of pathology* **149**(3), 903-910
314. Radding, W., Pan, Z. Q., Hunter, E., Johnston, P., Williams, J. P., and McDonald, J. M. (1996) *Biochem Biophys Res Commun* **218**(1), 192-197
315. Srinivas, S. K., Srinivas, R. V., Anantharamaiah, G. M., Compans, R. W., and Segrest, J. P. (1993) *J Biol Chem* **268**(30), 22895-22899
316. Yuan, T., Mietzner, T. A., Montelaro, R. C., and Vogel, H. J. (1995) *Biochemistry* **34**(33), 10690-10696
317. Radding, W., Williams, J. P., McKenna, M. A., Tummala, R., Hunter, E., Tytler, E. M., and McDonald, J. M. (2000) *AIDS research and human retroviruses* **16**(15), 1519-1525
318. Otteken, A., and Moss, B. (1996) *J Biol Chem* **271**(1), 97-103
319. Alefantis, T., Flaig, K. E., Wigdahl, B., and Jain, P. (2007) *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* **61**(4), 194-200
320. Holaska, J. M., Black, B. E., Rastinejad, F., and Paschal, B. M. (2002) *Molecular and cellular biology* **22**(17), 6286-6297
321. Ding, W., Albrecht, B., Luo, R., Zhang, W., Stanley, J. R., Newbound, G. C., and Lairmore, M. D. (2001) *Journal of virology* **75**(16), 7672-7682
322. Johnson, J. M., Mulloy, J. C., Ciminale, V., Fullen, J., Nicot, C., and Franchini, G. (2000) *AIDS research and human retroviruses* **16**(16), 1777-1781

- 323. Mantovani, F., and Banks, L. (2001) *Oncogene* **20**(54), 7874-7887
- 324. Du, M., Fan, X., Hong, E., and Chen, J. J. (2002) *Biochem Biophys Res Commun* **296**(4), 962-969
- 325. Chen, J. J., Reid, C. E., Band, V., and Androphy, E. J. (1995) *Science* **269**(5223), 529-531
- 326. Chen, J. J., Hong, Y., Rustamzadeh, E., Baleja, J. D., and Androphy, E. J. (1998) *J Biol Chem* **273**(22), 13537-13544
- 327. Mason, W. S. (1993) *Journal of hepatology* **17 Suppl 3**, S137-142
- 328. Choi, J., Chang, J. S., Song, M. S., Ahn, B. Y., Park, Y., Lim, D. S., and Han, Y. S. (2003) *Biochem Biophys Res Commun* **305**(4), 1049-1056
- 329. De Meyer, S., Gong, Z., Depla, E., Maertens, G., and Yap, S. H. (1999) *Journal of hepatology* **31**(5), 783-790
- 330. Chen, M. H., Tian, G. W., Gafni, Y., and Citovsky, V. (2005) *Plant physiology* **138**(4), 1866-1876
- 331. Tomita, Y., Yamashita, T., Sato, H., and Taira, H. (1999) *J Biochem (Tokyo)* **126**(6), 1090-1100
- 332. Creutz, C. E. (1992) *Science* **258**(5084), 924-931
- 333. Pietropaolo, R. L., and Compton, T. (1997) *Journal of virology* **71**(12), 9803-9807
- 334. Raynor, C. M., Wright, J. F., Waisman, D. M., and Prydzial, E. L. (1999) *Biochemistry* **38**(16), 5089-5095
- 335. Ma, G., Greenwell-Wild, T., Lei, K., Jin, W., Swisher, J., Hardegen, N., Wild, C. T., and Wahl, S. M. (2004) *The Journal of experimental medicine*

- 200**(10), 1337-1346
336. Depla, E. (2000) *Curr Opin Investig Drugs* **1**(4), 415-420
 337. Yuan, T., Tencza, S., Mietzner, T. A., Montelaro, R. C., and Vogel, H. J. (2001) *Biopolymers* **58**(1), 50-62
 338. Beaton, A. R., Rodriguez, J., Reddy, Y. K., and Roy, P. (2002) *Proc Natl Acad Sci U S A* **99**(20), 13154-13159
 339. Frey, T. K. (1994) *Adv Virus Res* **44**, 69-160
 340. Garbutt, M., Law, L. M., Chan, H., and Hobman, T. C. (1999) *Journal of virology* **73**(5), 3524-3533
 341. Allen, M. D., Buckle, A. M., Cordell, S. C., Lowe, J., and Bycroft, M. (2003) *J Mol Biol* **330**(3), 503-511
 342. Gorbalenya, A. E., Koonin, E. V., and Lai, M. M. (1991) *FEBS Lett* **288**(1-2), 201-205
 343. Chen, J. P., Strauss, J. H., Strauss, E. G., and Frey, T. K. (1996) *Journal of virology* **70**(7), 4707-4713
 344. Liang, Y., Yao, J., and Gillam, S. (2000) *Journal of virology* **74**(12), 5412-5423
 345. Liu, X., Ropp, S. L., Jackson, R. J., and Frey, T. K. (1998) *Journal of virology* **72**(5), 4463-4466
 346. Liu, X., Yang, J., Ghazi, A. M., and Frey, T. K. (2000) *Journal of virology* **74**(13), 5949-5956
 347. Huang, M. C., Pan, P. K., Zheng, Z. F., Chen, N. C., Peng, J. Y., and Huang, P. C. (1998) *Gene* **211**(1), 49-55

- 348. Schito, M. L., Goel, A., Song, Y., Inman, J. K., Fattah, R. J., Rice, W. G., Turpin, J. A., Sher, A., and Appella, E. (2003) *AIDS research and human retroviruses* **19**(2), 91-101
- 349. Schito, M. L., Soloff, A. C., Slovit, D., Trichel, A., Inman, J. K., Appella, E., Turpin, J. A., and Barratt-Boyes, S. M. (2006) *Current HIV research* **4**(3), 379-386
- 350. Blumenschein, T. M., and Reinach, F. C. (2000) *Biochemistry* **39**(13), 3603-3610
- 351. Werber, M. M., and Oplatka, A. (1974) *Biochem Biophys Res Commun* **57**(3), 823-830
- 352. Cox, J. A., and Stein, E. A. (1981) *Biochemistry* **20**(19), 5430-5436
- 353. Roberts, P. J., and Belsham, G. J. (1995) *Virology* **213**(1), 140-146
- 354. Guarne, A., Tormo, J., Kirchweger, R., Pfistermueller, D., Fita, I., and Skern, T. (1998) *Embo J* **17**(24), 7469-7479
- 355. Wang, C. L., Aquaron, R. R., Leavis, P. C., and Gergely, J. (1982) *Eur J Biochem* **124**(1), 7-12
- 356. Drake, S. K., Lee, K. L., and Falke, J. J. (1996) *Biochemistry* **35**(21), 6697-6705
- 357. Ogawa, Y., and Tanokura, M. (1984) *J Biochem (Tokyo)* **95**(1), 19-28
- 358. Reid, R. E., and Procyshyn, R. M. (1995) *Arch Biochem Biophys* **323**(1), 115-119
- 359. Gifford, J. L., Walsh, M. P., and Vogel, H. J. (2007) *Biochem J* **405**(2), 199-221

- 360. Kay, L. E. (1995) *Curr Opin Struct Biol* **5**(5), 674-681
- 361. Yang, J. J., Gawthrop, A., and Ye, Y. (2003) *Protein Pept Lett* **10**(4), 331-345
- 362. Sharma, Y., Chandani, S., Sukhaswami, M. B., Uma, L., Balasubramanian, D., and Fairwell, T. (1997) *Eur J Biochem* **243**(1-2), 42-48
- 363. Finn, B. E., Evenas, J., Drakenberg, T., Waltho, J. P., Thulin, E., and Forsen, S. (1995) *Nature structural biology* **2**(9), 777-783
- 364. Maune, J. F., Beckingham, K., Martin, S. R., and Bayley, P. M. (1992) *Biochemistry* **31**(34), 7779-7786
- 365. Beckingham, K. (1991) *J Biol Chem* **266**(10), 6027-6030
- 366. Zhu, T., Beckingham, K., and Ikebe, M. (1998) *J Biol Chem* **273**(32), 20481-20486
- 367. Linse, S., and Forsen, S. (1995) *Adv Second Messenger Phosphoprotein Res* **30**, 89-151
- 368. Smith, C. A., Toogood, H. S., Baker, H. M., Daniel, R. M., and Baker, E. N. (1999) *J Mol Biol* **294**(4), 1027-1040
- 369. Tajima, M., Urabe, I., Yutani, K., and Okada, H. (1976) *Eur J Biochem* **64**(1), 243-247
- 370. Liang, Y., and Gillam, S. (2001) *Virology* **282**(2), 307-319
- 371. Christensen, K. A., Myers, J. T., and Swanson, J. A. (2002) *J Cell Sci* **115**(Pt 3), 599-607
- 372. Gerasimenko, J. V., Tepikin, A. V., Petersen, O. H., and Gerasimenko, O. V. (1998) *Curr Biol* **8**(24), 1335-1338

373. Pando, V., Isa, P., Arias, C. F., and Lopez, S. (2002) *Virology* **295**(1), 190-200
374. Chen, Y., Carrington-Lawrence, S. D., Bai, P., and Weller, S. K. (2005) *Journal of virology* **79**(14), 9088-9096
375. Gorelick, R. J., Chabot, D. J., Ott, D. E., Gagliardi, T. D., Rein, A., Henderson, L. E., and Arthur, L. O. (1996) *Journal of virology* **70**(4), 2593-2597
376. Mark-Danieli, M., Laham, N., Kenan-Eichler, M., Castiel, A., Melamed, D., Landau, M., Bouvier, N. M., Evans, M. J., and Bacharach, E. (2005) *Journal of virology* **79**(12), 7756-7767
377. Fukuhara, N., Huang, C., Kiyotani, K., Yoshida, T., and Sakaguchi, T. (2002) *Virology* **299**(2), 172-181
378. Joseph, J. S., Saikatendu, K. S., Subramanian, V., Neuman, B. W., Brooun, A., Griffith, M., Moy, K., Yadav, M. K., Velasquez, J., Buchmeier, M. J., Stevens, R. C., and Kuhn, P. (2006) *Journal of virology* **80**(16), 7894-7901
379. Seybert, A., Posthuma, C. C., van Dinten, L. C., Snijder, E. J., Gorbalenya, A. E., and Ziebuhr, J. (2005) *Journal of virology* **79**(2), 696-704
380. Tellinghuisen, T. L., Marcotrigiano, J., Gorbalenya, A. E., and Rice, C. M. (2004) *J Biol Chem* **279**(47), 48576-48587
381. Tellinghuisen, T. L., Marcotrigiano, J., and Rice, C. M. (2005) *Nature* **435**(7040), 374-379
382. Qin, W., Luo, H., Nomura, T., Hayashi, N., Yamashita, T., and Murakami, S.

- (2002) *J Biol Chem* **277**(3), 2132-2137
383. Han, Y. S., Chang, G. G., Juo, C. G., Lee, H. J., Yeh, S. H., Hsu, J. T., and Chen, X. (2005) *Biochemistry* **44**(30), 10349-10359
 384. Tijms, M. A., van Dinten, L. C., Gorbalenya, A. E., and Snijder, E. J. (2001) *Proc Natl Acad Sci U S A* **98**(4), 1889-1894
 385. Dvorak, C. M., Hall, D. J., Hill, M., Riddle, M., Pranter, A., Dillman, J., Deibel, M., and Palmenberg, A. C. (2001) *Virology* **290**(2), 261-271
 386. Barbato, G., Cicero, D. O., Nardi, M. C., Steinkuhler, C., Cortese, R., De Francesco, R., and Bazzo, R. (1999) *J Mol Biol* **289**(2), 371-384
 387. McCall, K. A., and Fierke, C. A. (2000) *Anal Biochem* **284**(2), 307-315
 388. Orrenius, S., Zhivotovsky, B., and Nicotera, P. (2003) *Nat Rev Mol Cell Biol* **4**(7), 552-565
 389. Jurado, L. A., Chockalingam, P. S., and Jarrett, H. W. (1999) *Physiological reviews* **79**(3), 661-682
 390. Ikura, M. (1996) *Trends Biochem Sci* **21**(1), 14-17
 391. Peracchia, C. (2004) *Biochim Biophys Acta* **1662**(1-2), 61-80
 392. Tomlinson, S., MacNeil, S., Walker, S. W., Ollis, C. A., Merritt, J. E., and Brown, B. L. (1984) *Clin Sci (Lond)* **66**(5), 497-507
 393. Alm, P. E. (1984) *Int Arch Allergy Appl Immunol* **75**(4), 375-378
 394. Zhu, M. X. (2005) *Pflugers Arch* **451**(1), 105-115
 395. Grabarek, Z. (2006) *J Mol Biol* **359**(3), 509-525
 396. Wu, X., and Reid, R. E. (1997) *Biochemistry* **36**(28), 8649-8656
 397. Reid, R. E. (1990) *J Biol Chem* **265**(11), 5971-5976

398. Browne, J. P., Strom, M., Martin, S. R., and Bayley, P. M. (1997) *Biochemistry* **36**(31), 9550-9561
399. Grabarek, Z. (2005) *J Mol Biol* **346**(5), 1351-1366
400. Peersen, O. B., Madsen, T. S., and Falke, J. J. (1997) *Protein Sci* **6**(4), 794-807
401. Martin, S. R., Bayley, P. M., Brown, S. E., Porumb, T., Zhang, M., and Ikura, M. (1996) *Biochemistry* **35**(11), 3508-3517
402. Forsen, S., Linse, S., Drakenberg, T., Kordel, J., Akke, M., Sellers, P., Johansson, C., Thulin, E., Andersson, I., Brodin, P., and et al. (1991) *Ciba Found Symp* **161**, 222-236
403. Finn, B. E., Drakenberg, T., and Forsen, S. (1993) *FEBS Lett* **336**(2), 368-374
404. Delville, A., Laszlo, P., and Nelson, D. J. (1985) *J Theor Biol* **112**(1), 157-175
405. Ohki, S., Ikura, M., and Zhang, M. (1997) *Biochemistry* **36**(14), 4309-4316
406. Tsai, M. D., Drakenberg, T., Thulin, E., and Forsen, S. (1987) *Biochemistry* **26**(12), 3635-3643
407. Malmendal, A., Linse, S., Evenas, J., Forsen, S., and Drakenberg, T. (1999) *Biochemistry* **38**(36), 11844-11850
408. Malmendal, A., Evenas, J., Thulin, E., Gippert, G. P., Drakenberg, T., and Forsen, S. (1998) *J Biol Chem* **273**(44), 28994-29001
409. Zhang, M., Tanaka, T., and Ikura, M. (1995) *Nature structural biology* **2**(9), 758-767

- 410. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C. B., and Bax, A. (1995) *Nature structural biology* **2**(9), 768-776
- 411. Wilson, M. A., and Brunger, A. T. (2000) *J Mol Biol* **301**(5), 1237-1256
- 412. Nelson, M. R., and Chazin, W. J. (1998) *Biometals* **11**(4), 297-318
- 413. Zhang, M., Li, M., Wang, J. H., and Vogel, H. J. (1994) *J Biol Chem* **269**(22), 15546-15552
- 414. Balog, E. M., Norton, L. E., Bloomquist, R. A., Cornea, R. L., Black, D. J., Louis, C. F., Thomas, D. D., and Fruen, B. R. (2003) *J Biol Chem* **278**(18), 15615-15621
- 415. Yuan, T., and Vogel, H. J. (1999) *Protein Sci* **8**(1), 113-121
- 416. Nelson, M. R., and Chazin, W. J. (1998) *Protein Sci* **7**(2), 270-282
- 417. Chattopadhyaya, R., Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) *J Mol Biol* **228**(4), 1177-1192
- 418. Babu, Y. S., Bugg, C. E., and Cook, W. J. (1988) *J Mol Biol* **204**(1), 191-204
- 419. Ikura, M., Barbato, G., Klee, C. B., and Bax, A. (1992) *Cell calcium* **13**(6-7), 391-400
- 420. Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science* **256**(5057), 632-638
- 421. Osawa, M., Tokumitsu, H., Swindells, M. B., Kurihara, H., Orita, M., Shibamura, T., Furuya, T., and Ikura, M. (1999) *Nature structural biology* **6**(9), 819-824
- 422. Elshorst, B., Hennig, M., Forsterling, H., Diener, A., Maurer, M., Schulte, P.,

- Schwalbe, H., Griesinger, C., Krebs, J., Schmid, H., Vorherr, T., and Carafoli, E. (1999) *Biochemistry* **38**(38), 12320-12332
423. Shen, Y., Zhukovskaya, N. L., Guo, Q., Florian, J., and Tang, W. J. (2005) *Embo J* **24**(5), 929-941
424. Martin, S. R., and Bayley, P. M. (2004) *FEBS Lett* **567**(2-3), 166-170
425. Schumacher, M. A., Crum, M., and Miller, M. C. (2004) *Structure* **12**(5), 849-860
426. Yuan, T., and Vogel, H. J. (1998) *J Biol Chem* **273**(46), 30328-30335
427. O'Day, D. H. (2003) *Cell Signal* **15**(4), 347-354
428. Shen, X., Valencia, C. A., Gao, W., Cotten, S. W., Dong, B., Huang, B. C., and Liu, R. (2007) *Cell calcium*
429. Huang, B. C., and Liu, R. (2007) *Biochemistry* **46**(35), 10102-10112
430. Shen, X., Valencia, C. A., Szostak, J. W., Dong, B., and Liu, R. (2005) *Proc Natl Acad Sci U S A* **102**(17), 5969-5974
431. Mathias, R. T., Rae, J. L., and Baldo, G. J. (1997) *Physiological reviews* **77**(1), 21-50
432. Musil, L. S., Beyer, E. C., and Goodenough, D. A. (1990) *J Membr Biol* **116**(2), 163-175
433. Noma, A., and Tsuboi, N. (1987) *J Physiol* **382**, 193-211
434. Peracchia, C., Sotkis, A., Wang, X. G., Peracchia, L. L., and Persechini, A. (2000) *J Biol Chem* **275**(34), 26220-26224
435. Torok, K., Stauffer, K., and Evans, W. H. (1997) *Biochem J* **326** (Pt 2), 479-483

- 436. Gandolfi, S. A., Duncan, G., Tomlinson, J., and Maraini, G. (1990) *Curr Eye Res* **9**(6), 533-541
- 437. Zhang, X., and Qi, Y. (2005) *Arch Biochem Biophys* **440**(2), 111-117
- 438. Duffy, H. S., Sorgen, P. L., Girvin, M. E., O'Donnell, P., Coombs, W., Taffet, S. M., Delmar, M., and Spray, D. C. (2002) *J Biol Chem* **277**(39), 36706-36714
- 439. Rhoads, A. R., and Friedberg, F. (1997) *Faseb J* **11**(5), 331-340
- 440. Clapperton, J. A., Martin, S. R., Smerdon, S. J., Gamblin, S. J., and Bayley, P. M. (2002) *Biochemistry* **41**(50), 14669-14679
- 441. Rosenberg, O. S., Deindl, S., Sung, R. J., Nairn, A. C., and Kuriyan, J. (2005) *Cell* **123**(5), 849-860
- 442. Yamauchi, E., Nakatsu, T., Matsubara, M., Kato, H., and Taniguchi, H. (2003) *Nature structural biology* **10**(3), 226-231
- 443. Nicol, S., Rahman, D., and Baines, A. J. (1997) *Biochemistry* **36**(38), 11487-11495
- 444. Yang, J. J., Buck, M., Pitkeathly, M., Kotik, M., Haynie, D. T., Dobson, C. M., and Radford, S. E. (1995) *J Mol Biol* **252**(4), 483-491
- 445. Lehrman, S. R., Tuls, J. L., and Lund, M. (1990) *Biochemistry* **29**(23), 5590-5596
- 446. Brokx, R. D., Scheek, R. M., Weljie, A. M., and Vogel, H. J. (2004) *J Struct Biol* **146**(3), 272-280
- 447. Zhang, M., Yuan, T., and Vogel, H. J. (1993) *Protein Sci* **2**(11), 1931-1937
- 448. Zhang, M., and Vogel, H. J. (1994) *Biochemistry* **33**(5), 1163-1171

- 449. Zhan, Q. Q., Wong, S. S., and Wang, C. L. (1991) *J Biol Chem* **266**(32), 21810-21814
- 450. Martin, S. R., and Bayley, P. M. (1986) *Biochem J* **238**(2), 485-490
- 451. Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M., and Bax, A. (1991) *Biochemistry* **30**(38), 9216-9228
- 452. Chou, J. J., Li, S., Klee, C. B., and Bax, A. (2001) *Nature structural biology* **8**(11), 990-997
- 453. Black, D. J., Tran, Q. K., and Persechini, A. (2004) *Cell calcium* **35**(5), 415-425
- 454. Johnson, J. D., Snyder, C., Walsh, M., and Flynn, M. (1996) *J Biol Chem* **271**(2), 761-767
- 455. Kasturi, R., Vasulka, C., and Johnson, J. D. (1993) *J Biol Chem* **268**(11), 7958-7964
- 456. Brown, S. E., Martin, S. R., and Bayley, P. M. (1997) *J Biol Chem* **272**(6), 3389-3397
- 457. Sorgen, P. L., Duffy, H. S., Spray, D. C., and Delmar, M. (2004) *Biophys J* **87**(1), 574-581
- 458. Sheng, M., and Sala, C. (2001) *Annu Rev Neurosci* **24**, 1-29
- 459. Urbauer, J. L., Short, J. H., Dow, L. K., and Wand, A. J. (1995) *Biochemistry* **34**(25), 8099-8109
- 460. Xiong, L. W., Newman, R. A., Rodney, G. G., Thomas, O., Zhang, J. Z., Persechini, A., Shea, M. A., and Hamilton, S. L. (2002) *J Biol Chem* **277**(43), 40862-40870
- 461. Kranz, J. K., Lee, E. K., Nairn, A. C., and Wand, A. J. (2002) *J Biol Chem* **277**(19), 16351-16354
- 462. Ishida, H., and Vogel, H. J. (2006) *Protein Pept Lett* **13**(5), 455-465
- 463. Meador, W. E., Means, A. R., and Quirocho, F. A. (1993) *Science* **262**(5140), 1718-1721
- 464. Yamniuk, A. P., and Vogel, H. J. (2004) *J Biol Chem* **279**(9), 7698-7707
- 465. Gomes, A. V., Barnes, J. A., and Vogel, H. J. (2000) *Arch Biochem Biophys* **379**(1), 28-36
- 466. Trost, C., Bergs, C., Himmerkus, N., and Flockerzi, V. (2001) *Biochem J* **355**(Pt 3), 663-

670

- 467. Turner, J. H., Gelasco, A. K., and Raymond, J. R. (2004) *J Biol Chem* **279**(17), 17027-17037
- 468. Mal, T. K., Skrynnikov, N. R., Yap, K. L., Kay, L. E., and Ikura, M. (2002) *Biochemistry* **41**(43), 12899-12906
- 469. Bhattacharya, S., Bunick, C. G., and Chazin, W. J. (2004) *Biochim Biophys Acta* **1742**(1-3), 69-79