Dissecting Key Determinants for Calcium and Calmodulin Regulation of GAP Junction and Viral Protein

Yanyi Chen

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

Recommended Citation
Chen, Yanyi, "Dissecting Key Determinants for Calcium and Calmodulin Regulation of GAP Junction and Viral Protein." Dissertation, Georgia State University, 2012. https://scholarworks.gsu.edu/chemistry_diss/64

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.
ABSTRACT

Calcium and calmodulin are implicated in mediating the Ca^{2+}-dependent regulation of gap junctions that are essential for the intercellular transmission of molecules such as nutrients, metabolites, metal ions and signal messengers (< 1000 Da) through its specialized cell membrane channels and communication to extracellular environment. To understand the key determinants for calcium and calmodulin regulation of gap junction, in this study, we identified a calmodulin binding domain in the second half of the intracellular loop of Cxonnexin50 (the major gap junction protein found in an eye lens) using peptide fragments that encompass predicted CaM binding sites and various biophysical methods. Our study provides the first direct evidence that CaM binds to a specific region of the ubiquitous gap junction protein Cx50 in a Ca^{2+}-dependent manner. Furthermore, two novel CaM binding regions in cytosolic loop and C-termini of Connexin43 (the most ubiquitous connexin) have been shown to interact with CaM with
different binding modes in the presence of Ca\(^{2+}\) using high resolution NMR. Our results also elucidate the molecular determinants of regulation of gap junction by multiple CaM targeting regions and provide insight into the molecular basis of gap junction gating mechanism and the binding of CaM to the cytosolic region Cx43-3p as the major regulation site. Upon response to the cytosolic calcium increase, CaM binds to the cytosolic loop to result in the conformational change of gap junction and close the channel. It is possible for CaM to use an adjacent region as an anchor close to the regulation site to allow for fast response. Since a large number of residues in the Cxs mutated in human diseases reside at the highly identified CaM binding sites in Cxs, our studies provide insights into define the critical cellular changes and molecular mechanisms contributing to human disease pathogenesis as part of an integrated molecular model for the calcium regulation of GJs. In addition, we have applied the grafting approach to probe the metal binding capability of predicted EF-hand motifs within the streptococcal hemoprotein receptor (Shr) of Streptococcus pyogenes as well as the nonstructural protein 1 (nsP1) of Sindbis virus and Poxvirus. This fast and robust method allows us to analyze putative EF-hand proteins at genome-wide scale and to further visualize the evolutionary scenario of the EF-hand protein family. Further, mass spectrometry has also been applied to probe modification of proteins such as CaM labeling by florescence dye and 7E15 by PEG.

INDEX WORDS: Calcium, EF-hand, Calcium binding proteins, Calmodulin, CalbindinD9k, CD2, Calmodulin target protein, Protein-protein interaction, Protein grafted approach, Gap junction, Connexin26, Connexin43, Connexin50
DISSECTING KEY DETERMINANTS FOR CALCIUM AND CALMODULIN REGULATION OF 
GAP JUNCTION AND VIRAL PROTEIN

by

YANYI CHEN

A Dissertation Submitted in Partial Fulfillment of the Requirements of the Degree of 
Doctor of Philosophy 
in the College of Arts and Sciences 
Georgia State University
2012
DISSECTING KEY DETERMINANTS FOR CALCIUM AND CALMODULIN REGULATION OF GAP JUNCTION AND VIRAL PROTEIN

by

YANYI CHEN

Committee Chair: Jenny J. Yang

Committee: Charles F. Louis

Richard D. Veenstra

Aimin Liu

Siming Wang

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2012
ACKNOWLEDGEMENTS

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. The completion of my dissertation and subsequent Ph.D. has been a long journey. Her patience, flexibility, genuine caring and concern and faith in me during the dissertation process enabled me to attend to life while also earning my Ph.D. I would like to express my deepest gratitude to Prof. Richard D. Veenstra, Prof. Charles F. Louis, and Prof. Aimin Liu who gave me their full support and guidance in performing this research. I would like to express my sincere gratitude to Dr. Siming Wang for the immeasurable amount of support and guidance she has provided throughout mass spec. study. This dissertation would not have been possible without the expert guidance of these 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. Yubin Zhou for initiating the project and teaching me almost everything about science and research, to Dr. Hing-Cheung Wong 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. Jin Zou for his help in CD and fluoremeter, to Dr. Monica Lurtz for the cell-based assays, to Xu Qin for the majority of the gap junction experiments along with Dr. Xianming Lin on the Cx43 and Cx50 projects, to You Zhuo and Juan Zou for working with me in gap junction project. I would also extend my gratitude to Natalie White and Matthew Cameron for carefully proof-reading all the chapters of this dissertation and provided many stylistic suggestions. This work is supported the grants from National Institutes of Health.

Finally, I am especially grateful to my parents Wansheng Chen and Mingxiang Liu for giving the opportunity to follow my dreams, to my wife Yangqian Shao and my daughter Amber Chen, for the patient love to make all of these a reality.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................................................................................. iv

LIST OF TABLES .............................................................................................................................................. ix

LIST OF FIGURES ........................................................................................................................................... x

LIST OF ABBREVIATIONS ........................................................................................................................ xiv

1. INTRODUCTION......................................................................................................................................... 1
   1.1. Roles of Calcium (Ca\(^{2+}\)) in biological systems ............................................................................. 1
   1.2. The properties of EF-hand motif in Ca\(^{2+}\)-binding proteins .......................................................... 3
   1.3. A grafting approach: exploring the metal binding properties of continuous Ca\(^{2+}\) binding loops. .... 4
   1.4. Roles of Ca\(^{2+}/\)Calmodulin in regulating intracellular signaling ................................................... 6
   1.5. Gap junction in cell communication ................................................................................................... 9
   1.6. Calmodulin-mediated regulation of gap junctions. ............................................................................. 11
   1.7. The objectives of this dissertation .................................................................................................... 11
   1.8. Overview of this dissertation ............................................................................................................. 12

2. MATERIALS AND METHODS .................................................................................................................. 25
   2.1. Molecular cloning and plasmid construction .................................................................................... 25
   2.2. Baculovirus vector construction ....................................................................................................... 25
   2.3. Expression and purification of proteins ............................................................................................. 26
   2.4. Peptide design and synthesize .......................................................................................................... 27
   2.5. Nuclear magnetic resonance spectroscopy ....................................................................................... 28
   2.6. Fluorescence spectroscopy ............................................................................................................... 30
   2.7. Circular dichroism spectroscopy ....................................................................................................... 33
   2.8. Mass spectroscopy ............................................................................................................................ 33
   2.9. Electrophysiological measurement of junctional conductance ....................................................... 34
   2.10. Prediction of protein transmembrane topology and calmodulin binding domains ....................... 34
3. MOLECULAR INTERACTION AND FUNCTIONAL REGULATION OF CONNEXIN50 BY CALMODULIN

3.1. Introduction
3.2. Results
  3.2.1. A putative CaMBD in the single cytoplasmic loop of Cx50
  3.2.2. Ca\(^{2+}\)-dependent uncoupling of Cx50 gap junctions requires CaM
  3.2.3. Ca\(^{2+}\)-dependent direct interaction between Cx50 CaMBD and CaM observed by MALDI-MS and SDS-PAGE with cross-linking
  3.2.4. Ca\(^{2+}\)-dependent specific interaction between Cx50 CaMBD and CaM revealed by NMR
  3.2.5. Probing the CaM-Peptide complex state by diffusion NMR
  3.2.6. Revealing CaM-Peptide interactions with CD spectroscopy
  3.2.7. Determination of the CaM-Peptide binding affinity by fluorescence spectroscopy
  3.2.8. Binding of Cx50p expands the \([Ca^{2+}]_i\) sensing range of CaM
3.3. Discussion
  3.3.1. Regulation of Cxs by intracellular Ca\(^{2+}\) and CaM
  3.3.2. Comparison of CaM binding affinity with the alpha family Cxs
  3.3.3. Implications for the factors contributing to the CaM binding affinity to Cxs
  3.3.4. Structural and disease implications
3.4. Summary

4. EXPLORE MULTIPLE CALMODULIN BINDING SITES IN CONNEXIN43

4.1. Introduction
4.2. Results
  4.2.1. Design of peptide fragments to probe interactions of Cx43 by CaM
  4.2.2. Monitoring formation of the peptide-CaM complex using Mass spectrometry
  4.2.3. Monitoring the interaction between Cx43 peptides and CaM by NMR (1H, 15N)-HSQC spectroscopy
  4.2.4. Determining CaM binding mode of holo-CaM-Cx43-2p complex by residue dipolar coupling
  4.2.5. Determining the binding affinity of Cx43-2p/4p-CaM complex by modified CaM with fluorescence dye
4.3. Discussion
4.3.1. The binding stoichiometry between CaM and Cx43 peptides ............................................ 87
4.3.2. The CaM binding modes between CaM and Cx43 peptides .............................................. 89
4.3.3. Comparison of CaM binding domains with α-, β-, γ- and δ-family ................................... 91
4.3.4. Disease implications in CaM binding domains ................................................................. 92
4.3.5. The role of CaM in regulation of Cx43 .............................................................................. 93
4.4. Future plan .......................................................................................................................... 94
4.5. Summary and conclusion .................................................................................................. 95

5. IDENTIFICATION OF EF-HAND CA²⁺ BINDING PROTEINS IN VARIOUS BIOLOGICAL SYSTEMS .................................................................................................................................................... 112
5.1. Introduction .......................................................................................................................... 112
5.1.1. EF-hand Ca²⁺ binding proteins and protein grafted approach ........................................... 112
5.1.2. Animalia: Calbindin D9k ................................................................................................ .. 114
5.1.3. Bacteria: streptococcal hemoprotein receptor of Streprococcus pyogenes ...................... 115
5.1.4. Virus: nsP1 of Sindbis and Pox ......................................................................................... 116
5.2. Results .................................................................................................................................. 117
5.2.1. Probing metal binding of grafted EF-hand by FRET pair of Trp and Tb³⁺. ...................... 117
5.2.2. Revealing Ca²⁺ dependent conformation change by CD spectroscopy. ......................... 119
5.2.3. Observing oligomeric stage of CD2 with grafted EF-hand by PFG-NMR measurement and fluorescence anisotropy. ................................................................. 120
5.3. Discussion and summary. ................................................................................................ 122

6. MASS SPECTROMETRY OF PROTEIN STUDY ..................................................................... 139
6.1. Introduction .......................................................................................................................... 139
6.2. Dansylation of CaM and its variants .................................................................................... 141
6.3. IAEDANS dye labeling of CaM cysteine variants ................................................................. 143
6.4. Protein-protein interaction study of CaM with gap junction peptide .................................. 144
6.5. Dimerization study of engineered proteins with grafted EF-hand motif. ......................... 145
6.6. Determination of chemical PEGylation sites on protein based MRI contrast agent. ........ 146
6.7. Summary ........................................................................................................................... 147

7. MAJOR FINDING AND THE SIGNIFICANCE OF THIS DISSERTATION ........................ 165
7.1. Investigating CaM regulation of gap junction proteins ................................................................. 165
7.2. Identification of continuous calcium binding sites by grafting approach. ................................. 167
7.3. The study of protein-protein interactions by mass spectrometry............................................. 168

PUBLICATIONS AND MANUSCRIPTS ............................................................................................ 169

REFERENCES ........................................................................................................................................ 172
LIST OF TABLES

Chapter 1

Table 1. 1 The connexin family................................................................. 24

Chapter 2

Table 2. 1 Summary of GST fusion protein of CD2 variants....................... 35
Table 2. 2 Common salt for buffer control in different pH............................ 36

Chapter 3

Tabel 3. 1 Electrophysiological measurement........................................... 75
Tabel 3. 2 Binding affinities to CaM and α-helicity of the Cx peptides............. 76
Tabel 3. 3 Summary of Change of CaM helicity in different binding modes........ 77
Tabel 3. 4 Effects of Cx peptides binding on the metal-binding properties of CaM........ 78

Chapter 4

Table 4. 1 Sequence of CaM binding sites in Cx43 aligned with sequence of peptides from different CaM binding mode in collapsed subclass......................... 107
Table 4. 2 CaM binding modes and Q factors of the fitting of experimentally measured RDCs of the Ca^{2+}-CaM-Cx43-2p complex.................................................. 108
Table 4. 3 The binding affinities (µM) of Cx43 peptides to CaM (NMR), D-CaM and I-CaM. ................................................................. 109
Table 4. 4 Comparison of NMR chemical shift of CaM-Cx43p in identified CaM binding pockets of SK channel................................................................. 110
Table 4. 5 Oculodentodigital dysplasia (ODDD) disease mutation overlapping with Cx43 CaMBD. ................................................................. 111

Chapter 5

Table 5. 1 The secondary structure prediction by DichroWeb...................... 136
Table 5. 2 Hydro dynamic radii (R_{hydro}) and fluorescence anisotropy (r) of calbindin D_{9k} grafted proteins................................................................. 137
Table 5. 3 The summary of Tb^{3+} dissociation constants (K_d) in grafted proteins... 138

Chapter 6

Table 6. 1 UV MALDI Matrix list.............................................................. 164
## LIST OF FIGURES

### Chapter 1

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Calcium signaling in biological systems.</td>
<td>14</td>
</tr>
<tr>
<td>1.2</td>
<td>Calcium signaling in bacteria.</td>
<td>15</td>
</tr>
<tr>
<td>1.3</td>
<td>Ca^{2+}-binding affinities and structures of different classes of Ca^{2+}-binding proteins (CaBPs).</td>
<td>16</td>
</tr>
<tr>
<td>1.4</td>
<td>Experimental approaches to validate predicted continuous Ca^{2+}-binding motifs.</td>
<td>17</td>
</tr>
<tr>
<td>1.5</td>
<td>Calcium dependent Ca^{2+}-CaM-protein interaction.</td>
<td>18</td>
</tr>
<tr>
<td>1.6</td>
<td>Binding modes between CaM and its targets.</td>
<td>19</td>
</tr>
<tr>
<td>1.7</td>
<td>Classical CaM binding modes.</td>
<td>20</td>
</tr>
<tr>
<td>1.8</td>
<td>Assembly of gap junction proteins.</td>
<td>21</td>
</tr>
<tr>
<td>1.9</td>
<td>Schematic representation of the steps that lead to synthesis, assembly, and degradation of gap junction membrane channels.</td>
<td>22</td>
</tr>
<tr>
<td>1.10</td>
<td>Possible domains of Cx43 involved in interactions with partner proteins.</td>
<td>23</td>
</tr>
</tbody>
</table>

### Chapter 3

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Connexins membrane topology and the putative CaM-binding sites.</td>
<td>53</td>
</tr>
<tr>
<td>3.2</td>
<td>Ca^{2+} and CaM-dependence of Cx50 gap junction uncoupling.</td>
<td>54</td>
</tr>
<tr>
<td>3.3</td>
<td>Mechanistic basis for Ca^{2+}/CaM-dependent Cx50 gap junction uncoupling.</td>
<td>55</td>
</tr>
<tr>
<td>3.4</td>
<td>Monitoring CaM-Cx50_{141-166} complex formation by mass spectroscopy.</td>
<td>56</td>
</tr>
<tr>
<td>3.5</td>
<td>Cross-linking SDS-PAGE electrophoresis of CaM-Cx50p complex.</td>
<td>57</td>
</tr>
<tr>
<td>3.6</td>
<td>Monitoring the interaction between CaM and Cx50 peptides by (1H, 15N)-HSQC spectroscopy.</td>
<td>58</td>
</tr>
<tr>
<td>3.7</td>
<td>Cx50 Scramble peptide eliminates the interaction with holo-CaM by by (1H, 15N)-HSQC.</td>
<td>59</td>
</tr>
<tr>
<td>3.8</td>
<td>Chemical shift perturbation in CaM induced by addition of 2-fold molar excess of Cx50p_{141-166}.</td>
<td>60</td>
</tr>
<tr>
<td>3.9</td>
<td>The chemical shift change of Gly33 during titration of holo-CaM with Cx50p_{141-166}.</td>
<td>61</td>
</tr>
<tr>
<td>3.10</td>
<td>Hydrodynamic radii of CaM-Cx50p_{141-166} complex determination by PFG-NMR.</td>
<td>62</td>
</tr>
</tbody>
</table>
Figure 3.11 Monitoring the interaction between Apo-CaM and Cx50 peptides by \(^{1}H, {^{15}}N\)-HSQC spectroscopy.................................................................................................................. 63

Figure 3.12 CD studies of the interaction between CaM and the Cx50\textsubscript{141-166} peptide.............................. 64

Figure 3.13 CD Studies of the helicity of Cx43, Cx44 and Cx50 peptide............................................................... 65

Figure 3.14 CD Studies of Cx50\textsubscript{p141-166} in free and complex form.......................................................... 66

Figure 3.15 Structural basis of the difference in directionality of chemical shift change............................. 67

Figure 3.16 The dansyl fluorescence anisotropy................................................................................................. 68

Figure 3.17 The interaction of Cx50\textsubscript{141-166} peptide with holo Dansyl-CaM (D-CaM) monitored by steady-state fluorescence........................................................................................................... 69

Figure 3.18 The interaction of Cx50\textsubscript{141-166} peptide with apo Dansyl-CaM (D-CaM) monitored by steady-state fluorescence......................................................................................................................... 70

Figure 3.19 pH dependence of Cx50\textsubscript{p141-166} and Cx44\textsubscript{p129-150} binding to CaM............................ 71

Figure 3.20 N-Domain specific equilibrium \(Ca^{2+}\) titration of CaM (○), and CaM in complex with Cx50\textsubscript{p141-166} (●). .............................................................................................................................. 72

Figure 3.21 C-Domain specific equilibrium \(Ca^{2+}\) titration of CaM (○), and CaM in complex with Cx50\textsubscript{p141-166} (●).......................................................................................................................... 73

Figure 3.22 Comparison of the α family of connexins. ............................................................................ 74

Chapter 4

Figure 4.1 Connexins membrane topology putative CaM-binding sites of Cx43................................. 96

Figure 4.2 The predicted CaM binding sequences in Cx43........................................................................... 97

Figure 4.3 Monitoring CaM-Cx43-2p and CaM-Cx43-4p complex formation by MALDI mass spectrometry................. 98

Figure 4.4 Monitoring the interaction between holo-CaM and Cx43-2p peptide by \(^{1}H, {^{15}}N\)-HSQC spectroscopy............................................................................................................................... 99

Figure 4.5 Monitoring the interaction between holo-CaM and Cx43-4p peptide by \(^{1}H, {^{15}}N\)-HSQC spectroscopy........................................................................................................................................ 100

Figure 4.6 Comparison of the three CaM binding sites of connexin43 in \(^{1}H, {^{15}}N\)-HSQC spectrum. ........................................................................................................................................... 101

Figure 4.7 Comparison of the three CaM binding sites of connexin43 in chemical shift...................... 102
Figure 4. 8 CaM binding mode determination of holo-CaM-Cx43-2p complex by residue dipolar coupling (RDC)........................................................................................................................ 103

Figure 4. 9 The interaction of Cx43-2p and Cx43-4p peptide with IAEDANS-CaM (I-CaM) monitored by steady-state fluorescence. .................................................................................................................. 104

Figure 4. 10 The interaction of Cx43-2p peptide with Dansyl labeled C lobe of CaM (D-C-CaM) monitored by steady-state fluorescence (A) and anisotropy fluorescence (B)......................... 105

Figure 4. 11 Proposed model of Ca\textsuperscript{2+}/CaM regulation of assembling, trafficking, permeability and degradation in gap junction. ...................................................................................................... 106

Chapter 5

Figure 5. 1. Calcium is involved in multiple steps of virus life cycle.............................................. 124

Figure 5. 2. The Predicted EF-hand loops in various biological systems. ........................................ 125

Figure 5. 3. Characterization of CD2 grated protein (CD2.Pox.EF) by intrinsic.............................. 126

Figure 5. 4. FRET based Tb\textsuperscript{3+} titration of the engineered protein from the nsP1 of poxvirus (CD2.Pox.EF). .............................................................................................................................. 127

Figure 5. 5. Characterization of CD2 grated protein (CD2.Cal1.EF and CD2.Cal2.EF) by intrinsic Trp fluorescence excitation spectrum (A) and emission spectrum (B). .................................... 128

Figure 5. 6. FRET based Tb\textsuperscript{3+} titration of the engineered proteins from Calbindin D\textsubscript{9k}. ........................................ 129

Figure 5. 7. Probing metal binding properties of predicted EF-hand Ca\textsuperscript{2+} binding motifs from Shr of *Streptococcus pyrogenes* (CD2.Shr.EF) and from the nsP1 of Sindbis virus (CD2.Sin.EF).... 130

Figure 5. 8. Monitoring Ca\textsuperscript{2+} dependent secondary structure changes of CD2.Cal1.EF and CD2.Cal2.EF by far ultra-violet circular dichroism spectra. .................................................... 131

Figure 5. 9. Monitoring Ca\textsuperscript{2+} dependent secondary structure changes of CD2.Pox.EF and CD2.Shr.EF by far ultra-violet circular dichroism spectra. ...................................................... 132

Figure 5. 10. Determination of the oligomeric state of CD2.Cal1.EF and CD2.Cal2.EF by PFG-NMR. ........................................................................................................................................ 133

Figure 5. 11. Determination of the oligomeric state of CD2.Cal1.EF by fluorescence anisotropy. . 134

Figure 5. 12. Determination of the oligomeric state of CD2.Cal2.EF by fluorescence anisotropy. . 135

Chapter 6

Figure 6. 1 Voyager-DE\textsuperscript{TM PRO} workstation. ................................................................................. 148

Figure 6. 2 Desorption and Ionization processed in MALDI-MS........................................................... 149
Figure 6.3 MALDI-MS spectrum of CaM dansylation in various concentrations......................... 150
Figure 6.4 MALDI-MS spectrum of CaM and variants dansylation at various pH....................... 151
Figure 6.5 MALDI-MS spectrum of CaM variant IAEDANS labeling with cysteine mutation or insertion................................................................................................................................. 152
Figure 6.6 The effect of urea on CaM modification with IAEDANS dye...................................... 153
Figure 6.7 MALDI-MS spectrum of CaM and C45 peptide complex......................................... 154
Figure 6.8 Monitoring the interaction between CaM and Cx45 peptides (Cx45p) by (1H, 15N)-
HSQC spectroscopy ................................................................................................................. 155
Figure 6.9 Chemical shift perturbation in CaM induced by addition of 2-fold molar excess of
Cx45p........................................................................................................................................ 156
Figure 6.10 MALDI-MS spectrum of CD2.STIM1.EF in the presence of CaCl₂.......................... 157
Figure 6.11 MALDI-MS spectrum of CD2.STIM1.EF in the presence of EGTA. ....................... 158
Figure 6.12 MALDI-MS spectrum of CD2.D1 in the presence of CaCl₂...................................... 159
Figure 6.13 MALDI-MS spectrum of CD2.D1 in the presence of EGTA...................................... 160
Figure 6.14 Analysis of oligomer status of CD2.STIM1.EF by both MALDI-MS and crosslinking
SDS-PAGE. ................................................................................................................................ 161
Figure 6.15 MALDI-MS spectrum of MRI contrast agent 7E15 labeling with PEG.................... 162
Figure 6.16 The possible PEGylation sites determined by Nano spray ESI-MS......................... 163
LIST OF ABBREVIATIONS

APS Ammonium persulfate
CaBP Calcium binding protein
CaM Calmodulin
CaMBD Calmodulin binding domain
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
GA Golgi apparatus
HSQC Heteronuclear single quantum coherence
ICP Inductively coupled plasma
IPTG Isopropyl-D-thiogalactopyranoside
Kd Dissociation constant
koff Off rate
LB Luria-Bertani medium
MALDI Matrix-assisted laser desorption/ionization
MS Mass spectrometry
NMR Nuclear magnetic resonance
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSP</td>
<td>Nonstructural protein</td>
</tr>
<tr>
<td>NOESY</td>
<td>Nuclear Overhauser enhancement spectroscopy</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>RDC</td>
<td>Residue dipolar couplings</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SHR</td>
<td>Streptococcal hemoprotein receptor</td>
</tr>
<tr>
<td>SIN</td>
<td>Sindbis virus</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>TFE</td>
<td>Trifluoroethanol</td>
</tr>
<tr>
<td>TOCSY</td>
<td>Total correlation spectroscopy</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1. Roles of Calcium (Ca\textsuperscript{2+}) in biological systems

Calcium is an essential element in biological systems. Ca\textsuperscript{2+}, one ion with two charges, is the fifth most abundant dissolved ion in seawater by both mass and molarity. It is chosen by nature through evolution as one of the most universal and versatile carriers in cellular signal transduction. It functions as a pivotal regulator of the cell life cycle including cell division, differentiation, apoptosis and death [1-5]. Calcium is also the fifth most abundant element by the total mass in the Earth’s crust (approximately 3.6%); meanwhile its co-crystal form with phosphate and hydroxide ions forms the hardest substance in human body such as the matrix of tooth enamel and the deposition of the matrix in bone formed by calcium phosphate. Calcium is very tightly controlled in biological processes [6], Calcium plays an important role in both the immobilized form for the support of living organisms and mobilized ion in nearly every aspect of cellular life.

Ca\textsuperscript{2+} has a flexible coordination number of 4-8, an irregular coordination geometry and rapid binding kinetics, which maximizes its potential to interact with various ligands. Ca\textsuperscript{2+} tends to precipitate both inorganic and organic anions (at ~mM range) because high Ca\textsuperscript{2+} concentrations are incompatible with life and need to be extruded, sequestered or compartmentalized [7]. These unique physio-chemical properties of Ca\textsuperscript{2+} make it the only available candidate in the environment to perform both micro (as a messenger) and macro (e.g., biomineralization) functions in biological systems.

During the evolution from prokaryotes to eukaryotes that began over two billions years ago, the Fe\textsuperscript{2+}/H\textsubscript{2}S buffering system in the sea was progressively exhausted, along with the increase of oxygen supply and prevalence of oxidative reactions [7]. Ca\textsuperscript{2+} played an important role in facilitating the adaptive changes of prokaryotes to an oxidative environment, and contributed to the evolution of more advanced organisms and land colonization [8]. For example, Ca\textsuperscript{2+} was closely implicated in pivotal evolutionary
events such as the evolution of cytoskeleton, cell motility, the formation of skeletons and hard-shelled eggs (biomineralization), and the ability to supply milk.

By mass, Ca^{2+} represents the most abundant mineral (~1 kg) in the human body, with 99% in the form of calcium phosphate in bones. In mammals, a specific Ca^{2+} homeostatic system evolved to set both the extracellular and intracellular Ca^{2+} concentrations at proper levels, and produces a steep Ca^{2+} gradient (>10,000 times) across the plasma membrane [9, 10]. The external Ca^{2+}, sensed by the extracellular Ca^{2+} sensing receptor, is maintained at 1.1~1.4 mM through the coordinated actions of three major hormones (i.e., parathyroid hormone, calcitonin and 1,25-dihydroxyvitamin D₃) [11]. Compared to the extracellular Ca^{2+} concentration of about 10⁻³ M, the internal Ca^{2+} concentration is maintained within the narrow range of 10⁻⁷ to 10⁻⁸ M in nearly all living cells. In multi-cellular eukaryotes, Ca^{2+} levels in different cellular compartments or organelles are tightly controlled through the molecular choreography of a repertoire of Ca^{2+} signaling molecules, including Ca^{2+} channels, Ca^{2+} transporters, receptors, Ca^{2+} buffers, Ca^{2+}-responsive proteins (e.g., Ca^{2+}-sensitive enzymes and transcription factors) that are distributed in the cytoplasm, endoplasmic/sarcoplasmic reticulum (ER/SR), Golgi complex, mitochondria, nucleus or the extracellular matrix (Fig. 1.1) [12].

In eukaryotic cells, Ca^{2+} functions as a universal and versatile signal by interacting with hundreds of proteins over a 10⁶-fold range of affinities (nM to mM) (Fig.1.3A). Inside cells, Ca^{2+}-binding proteins, such as calmodulin or parvalbumin, have Ca^{2+}-binding affinities in the submicromolar range corresponding to the resting cytosolic Ca^{2+} concentrations of ~100 nM. Ca^{2+} concentrations in organelles such as the ER, Golgi, and endocytic vacuoles are in the hundreds of μM. By modulating the activities of these proteins, Ca^{2+} participates in various biological processes (e.g., hormonal secretion, muscle contraction, neurotransmission and memory formation) over a wide range of timescale, ranging from as swift as milliseconds to up to days (Fig. 1.1).
Viruses also utilize the universal Ca$^{2+}$ signal to create a specific cellular environment to achieve their own purposes. Ca$^{2+}$ plays important roles in viral gene expression, post-translational processing of viral proteins, virion structure formation, virus entry, and virion maturation and release. The interplay between viruses and Ca$^{2+}$ in the infected cell falls generally into three major categories: 1) viral proteins directly or indirectly disturb Ca$^{2+}$ homeostasis by altering membrane permeability and/or manipulating key components of the Ca$^{2+}$-signaling apparatus; 2) viral proteins directly bind to Ca$^{2+}$ for structural integrity or functionality; and 3) critical virus-host interactions depend on cellular Ca$^{2+}$-regulated proteins or pathways. Until now, very limited cases of viral calcium-binding proteins were reported in the literature, possibly due to challenges in the prediction and investigation of Ca$^{2+}$-binding sites.

In eukaryotes, the importance of calcium is well established. However, the role of calcium in prokaryotes remains unclear until the interest and evidence for calcium as a regulator in bacteria increased in recent years [13]. Even in the most primitive prokaryotes, the free cytosolic Ca$^{2+}$ is maintained at $\sim 10^{-7}$ M by using a variety of prokaryotic ion transporters (e.g., Ca$^{2+}$/H$^+$ antiporter, Ca$^{2+}$/Na$^+$ antiporter, and Ca$^{2+}$ ATPase) or channels (Fig. 1.2) [14-19]. Although bacteria do not possess complex intracellular organelles and Ca$^{2+}$ is not extensively used as an intracellular messenger, accumulating evidence indicate that Ca$^{2+}$ plays an essential role in the regulation of bacterial movements, chemotaxis, survival reactions and sporulation [10, 16, 17, 20].

Overall, the Ca$^{2+}$ signaling system exhibits great versatility in speed, amplitude and spatial-temporal patterns. In chapter 8, we focus on how integrative studies can be carried out to probe the sophisticated Ca$^{2+}$ signaling network in various biological systems.

1.2. The properties of EF-hand motif in Ca$^{2+}$-binding proteins.

According to their structural features, Ca$^{2+}$-binding sites in proteins are classified as either non-continuous or continuous. In non-continuous sites the Ca$^{2+}$ ligand residues are located remotely from one another in the protein sequence. Most of the Ca$^{2+}$ binding proteins, such as cadherins, C$_2$ domains, Site I
of thermitase, phospholipase A₂, and D-galactose binding protein (GBP) belong to this family. Continuous Ca²⁺-binding sites (e.g., EF-hands) have binding pockets formed by a stretch of contiguous amino acids in the primary sequence.

Based on the arrangement of the Ca²⁺-binding ligand residues in space, continuous Ca²⁺ binding proteins can be divided as EF-hand and non-EF-hand proteins. EF-hand Ca²⁺-binding proteins with a continuous helix-loop-helix topology were originally described in the structure of parvalbumin by Kretsinger (Fig. 1.3) [21]. EF-hand proteins have a conserved Ca²⁺ binding loop flanked by two helices [22, 23]. Based on the conserved features of the Ca²⁺-binding loop, EF-hand proteins have been divided into two major groups: the canonical EF-hands as seen in CaM and the pseudo EF-hands exclusively found in the N-termini of S100 and S100-like proteins [22]. Their major difference lies in the Ca²⁺ binding loop: the 12-residue canonical EF-hand loop binds Ca²⁺ mainly via side chains (loop positions 1, 3, 5, 12), whereas the 14-residue pseudo EF-hand loop chelates Ca²⁺ mostly via backbone carbonyls (positions 1, 4, 6, 9) (Fig. 1.3B). Each type of EF-hand loop has a bidentate Ca²⁺ ligand (Glu or Asp) that functions as an anchor at the C-terminal of the binding loop. Among all the structures reported to date, the majority of EF-hand sites have been found to be paired either within multiple canonical EF-hand motifs or through the interaction of one pseudo EF-hand motif with one canonical motif [22] (Fig. 1.3). For proteins with odd numbers of EF-hands, such as the penta-EF-hand calpain, EF-hand motifs are coupled through homodimerization or heterodimerization [24-26]. The calcium binding proteins with EF-hand motif play an essential role in cellular signaling such as trigger or sensor proteins calmodulin and troponin C (Fig. 1.3C), buffering protein parvalbumin and Calbindin D9k (Fig. 1.3D), and stabilizing protein thermolysin (Fig. 1.3E) [27-31].

1.3. A grafting approach: exploring the metal binding properties of continuous Ca²⁺ binding loops.

Due to the spectroscopically-silent nature of calcium and its physiological abundance, determination of the calcium binding capability of proteins is challenging. First, most experimental methods such as
dialysis are only sensitive to the total calcium content. In addition, overcoming the persistent background contamination of calcium during the preparation of calcium-free sample for proteins with strong calcium binding affinities is a non-trivial task. Further, since most calcium binding proteins contain multiple calcium binding sites that cooperatively bind calcium resulting in induced conformational change (e.g., CaM) (Fig. 1.3C), obtaining site-specific calcium binding affinity is limited by complication from contributions from cooperativity and conformational entropy [32]. Hence, understanding the molecular mechanism of biological function related to calcium is largely hampered by the lack of site specific information about the calcium-binding properties, especially for the ubiquitous EF-hand calcium-binding motif. Progress in understanding the molecular mechanism of calcium modulated biological process requires us to answer several important questions. First, we lack robust prediction algorithms and effective validating methods. Second, given its physiological abundance and spectroscopic silence, identification and analyzing Ca$^{2+}$-binding sites in proteins is not an easy task. Most experimental methods to determine Ca$^{2+}$-binding are sensitive to the total Ca$^{2+}$ concentration. The background Ca$^{2+}$ concentration, however, is usually hard to control and evaluate. Finally, the binding of Ca$^{2+}$ to EF-hands usually induces conformational change of the EF-hands. Therefore, evaluating site-specific Ca$^{2+}$-binding affinity is often complicated by the conformational cooperativity in the case of proteins possessing multiple Ca$^{2+}$-binding sites (e.g., calmodulin and troponin C) [33].

To overcome the first limitation, we have generated a handful of motif signatures for prediction of continuous EF-hand or EF-hand-like motifs [34]. Recently, we also developed MUG program to identify non-EF hand continuous calcium binding pockets [35]. More importantly, we have developed a grafting approach coupled with luminescence resonance energy transfer (LRET) to evaluate isolated Ca$^{2+}$-binding sites from EF-hand Ca$^{2+}$-binding proteins [36-38]. To circumvent the spectroscopically silent nature of Ca$^{2+}$, we use lanthanides to probe Ca$^{2+}$-binding because of their similar ionic radii (Ca$^{2+}$, 1.00 Å; Tb$^{3+}$, 0.92 Å) and metal coordination chemistry [39, 40]. As shown in Fig.1.4A, a putative Ca$^{2+}$-binding sequence can be inserted in a flexible loop in the domain 1 of CD2 (CD2.D1) with minimal perturbations
to the overall structure. The Trp residue (W32) within CD2.D1 is capable of transferring energy to Tb$^{3+}$ that binds to the inserted Ca$^{2+}$-binding loop. We optimized the insertion position and the length of the glycine linkers to optimize LERT efficiency, and meanwhile, to minimize the perturbation to the host structure [38]. Insertion of the putative Ca$^{2+}$-binding sequence between residue Ser52 and Gly53 flanked by triple glycine linkers proves to be most effective.

A flow chart describing a general approach to probe the Ca$^{2+}$ binding capability of EF-hand proteins is outlined in Fig.1.4A. First, putative calcium-binding proteins can be predicted using our developed calcium-binding sites prediction algorithm [34, 41-44] or using the pattern search server available at ScanProsite [45, 46]. A total of 93 putative EF-hands or EF-hand like motifs have been found in the viral genomes, including 80 different viruses that spread in the majority of virus families [47]. The native structure of host protein CD2.D1 after calcium binding loop insertion is essential for the accuracy of the metal binding assay. Next, the correct protein folding of engineered proteins is confirmed by both intrinsic tryptophan fluorescence and circular dichroism spectroscopy. Finally, the site specific Tb$^{3+}$ binding affinity is obtained by titrating Tb$^{3+}$ into the sample containing engineered proteins; whereas the site specific Ca$^{2+}$-binding affinity is subsequently obtained with competition assay after obtaining site-specific Tb$^{3+}$-binding affinity. The calcium binding affinity can be calculated by the apparent dissociation constant obtained by metal competition assay, in which excess Ca$^{2+}$ competes with Tb$^{3+}$ in the binding pockets. Alternatively, 1D NMR spectrometry also can be used as a way to identify the metal binding pockets and confirm the native structure of host protein CD2.D1 (Fig. 1.4B).

1.4. **Roles of Ca$^{2+}$/Calmodulin in regulating intracellular signaling.**

As the prototypical EF-hand calcium binding protein, calmodulin (CaM) plays key roles on eukaryotic cell calcium signal transduction and homeostasis. Calcium is selected by nature because survival requires the rapid and accurate signaling response by calcium as the appropriate second messenger molecules. Calcium uses CaM as a calcium sensor and transducer. CaM activity was first discovered in the brain and heart as an activator of cyclic nucleotide phosphodiesterase in the early 1970s [48, 49]. To reflect the
multifunctional nature of CaM as a calcium regulatory protein, the name of calmodulin was derived from calcium modulator protein [50]. Watterson et al. purified CaM in 1976 and reported the first complete amino acid sequence of CaM in 1980 [51, 52]. Only a single isoform of CaM encoded by three separate genes in mammalian cells regulates all the metabolic pathways in calcium signaling [53]. In contrast, plant cells have diverse CaM isoforms to regulate different cellular events in virtually every tissue type [54, 55]. CaM in yeast has only approximate 60% sequence identity to CaM in mammalian cells and the fourth EF-hand motifs in yeast CaM lost its ability to binding with calcium [56, 57]. In this study, we focus on CaM in mammalian cells.

CaM is a small (MW: 16706 Da) and acidic protein (pI = ~4.6) which contains 148 amino acids. N- and C-terminal globular domains in CaM are connected by a flexible central linker and as a result, CaM binds a total of four calcium ions. Each domain contains two helix-loop-helix Ca\(^{2+}\) binding EF-hand motifs. Due to 46% sequence identity between its N and C-domains, each of the two domains folds similarly in both secondary and tertiary structures. Calcium binding to CaM occurs sequentially in a positively cooperative manner. When calcium concentration increase, two sites in C-domain of CaM are filled first with a $K_d \sim 10^{-6}$ M, followed by two sites in N-domain with a $K_d \sim 10^{-5}$ M [58, 59]. The binding of Ca\(^{2+}\) to each site of CaM stabilizes its flexible structure and causes the helical conformation rearrangement of four EF-hands motifs while the orientation of helix pairs A/D, B/C, E/H and F/G remains unchanged, as they move from nearly antiparallel orientations in apo-CaM to almost perpendicular orientation in Ca-CaM [60]. This structure change is often described as moving from a “closed” form of apo-CaM into an “open” form of Ca-CaM. However, it has been proven by NMR solution structure that, in the presence of Ca\(^{2+}\), the N-domain is less open than the C-domain [61]. The central link region of CaM forms a longer helix loop in Ca-CaM and hydrophobic residues in this loop are exposed to the solvent, which allows other calcium signaling proteins (e.g. smooth muscle myosin light chain kinase) to bind to Ca-CaM (Fig. 1.5).
Through its reversible binding to Ca\(^{2+}\), CaM undergoes conformational changes and interacts with or releases from more than 150 target proteins [62]. The intracellular calcium concentration range is from \(10^{-7}\) (resting stage) to \(10^{-5}\) M (upon signal or mechanical stimulation) and the calcium binding affinities of CaM in each domain are \(10^{-5}\) M and \(10^{-6}\) M, respectively, which makes CaM capable of transducing the intracellular Ca\(^{2+}\) signal changes into a myriad of divergent cellular events, such as cell proliferation, cell differentiation and apoptosis [1]. The high number of Met residues in CaM is not common in proteins [63]. CaM forms nearly half of the hydrophobic binding interface area with its target protein by four Mets in each domain when calcium is bound [64]. The sulfur atom in Met is ideal for forming Van der Waals interaction with hydrophobic residues of target proteins [65]. In addition, the other amino acids with long aliphatic side chains, such as Val, Ile and Leu, contribute to the rest of the hydrophobic binding pockets which are “sticky” to CaM target proteins [66].

CaM binds to its targets either in a collapsed or an extended confirmation. In the collapsed conformation of CaM, the sequences in CaM binding domain (CaMBD) are divided into different motifs based on the distance between key hydrophobic residues such as Phe, Trp, Ile, Leu and Val with bulky hydrophobic side chains. The target protein with a CaMBD provides at least two hydrophobic anchor residues that interact with previously mentioned hydrophobic pockets of the N- and C-domains of CaM. To date, in this class of CaM binding proteins, the structures of five different motifs, 1 – 7, 1 – 10, 1 – 14, 1 – 16 and 1 – 17, have been discovered (Fig. 1.6). Peptide fragments in CaMBD usually have the propensity to form amphipathic alpha helices with both basic and hydrophobic faces [63, 67]. As shown in Fig. 1.7, the compactness or tightness of the complex is determined by the spacing between these two anchor residues in the CaMBD with helical conformation [68]. However, the relationship between the regulation and the tightness of CaM in the complex remains unclear.

CaM binds to its target proteins either in a Ca\(^{2+}\) dependent or independent manner. Most of CaM dependent enzymes and channel proteins are inhibited in the absence of CaM and calcium, but are activated in a Ca\(^{2+}\) dependent manner in the presence of CaM. However, increasing evidence
demonstrates that calcium-free CaM (Apo-CaM) also binds a number of highly diverse target proteins, including neuromodulin [69], neurogranin [70] and PEP-19 [71], inducible nitric oxide synthase (iNOS) [72], cyclic nucleotide PDE [73] and glycogen phosphorylase b Kinase (GbK) [74] as well as many unconventional myosins [75]. Neuromodulin may function as intracellular CaM storage proteins because it only binds CaM in the absence of CaM [76]. In some CaM target channel proteins, CaM binds a target with multiple CaMBDs. The binding between these CaMBDs of certain channel protein (e.g. Ryanodine receptor [77]) and CaM could be either Ca\textsuperscript{2+} dependent or independent. Apo-CaM target interaction may provide a way to concentrate CaM to specific regions of channel proteins with multiple CaMBDs for rapid response to the Ca\textsuperscript{2+} stimulus. Sequence analysis of CaMBD of apo-CaM reveals that several share a common motif known as IQ motif ((FILV)Qxxx(RK)Gxxx(RK)xx(FILVWY)) or IQ like motif ((FILV)Qxxx(RK)xxxxxxxx). Proteins that contain an IQ motif typically bind apo-CaM only but exceptions exist [75, 78]. Binding of incomplete IQ motifs which lack the second basic residue are Ca\textsuperscript{2+} dependent [79, 80]. Besides the common CaM binding classes discussed above, novel modes of CaM binding and activation of target protein have been revealed by structural studies. When CaM binds with exotoxin edema factor, the N-domain of CaM exists in a calcium free stage like an apo-form, although saturating calcium levels are used in crystallization condition [81]. In contrast, CaM in the complex with SK channels is found to have Ca\textsuperscript{2+} bound only to the N- but not to the C-domain [82, 83]. The structure and CaM binding modes of a large number of CaM complex remains undetermined. In this dissertation, CaM binding modes of gap junction proteins have been explored in chapters 3 and 4.

1.5. Gap junction in cell communication.

Gap junctions (GJs) comprise the intercellular channels that mediate a direct cell-to-cell transfer of small molecules (< 1 kDa) including metabolites, second messengers and ions between neighboring mammalian cells [84]. The hydrophilic bridge of GJ channels are formed by head-head docking of two hemichannels (connexon) each comprised of six connexin (Cx) subunits (Fig. 1.8). Hundreds of gap junction channels assemble together to form a huge plaque in the barrier between two cells. All the Cxs share a similar
topology with four highly-conserved transmembrane regions, a short N-terminal cytoplasmic region, one intracellular and two extracellular loops, and a C-terminal intracellular tail that exhibits the greatest sequence variation among the Cxs (Fig. 1.8) [85]. Each of two extracellular loops contains three conserved Cys residues which forms disulfide bonds intramolecularly as shown in the Fig. 1.8C [86]. It is well known that gap junction proteins are regulated by Ca\(^{2+}\) [87-90]. The increase of cytosolic calcium concentration triggers gap junction from open to close form (Fig. 1.8D). At least 20 Cx genes have been identified in the human genome with Cx43 being the most ubiquitous Cxs [91, 92]. Based on their sequence similarities, Cx can be further grouped into at least five classes, α (e.g. Cxs43, 44, and 50), β (e.g. Cxs26, 30, and 32), γ, δ and ε (Table 1.1) [93-95].

Assembly of gap junction channels and the delivery of the connexon to the plasma membrane is a very complicated process. As shown in Fig. 1.8A, gap junction channel can be assembled by different subunits in various combination, such as homomeric homotypic GJ (same subunits for all), homomeric heterotypic GJ (same subunits for each connexon, different connexons for GJ), heteromeric homotypic GJ (different subunits for each connexon, same connexons for GJ) and heteromeric heterotypic GJ (different subunits for each connexon, different connexons for GJ). Different pathways of gap junction channel assembly in the β family have been proposed by Ahmad et al. (Fig. 1.8B) [96]. Segretain et al. summarize all the work about the life of gap junction channel in cell, including synthesis, assemble, trafficking, insertion into plasma membrane (PM), diffusion to form plaque and degradation (Fig. 1.9) [97]. In brief, the connexons of Cx26 and Cx32 in β family are assembled in the endoplasmic reticulum (ER); in contrast, Cx43 in the α family was assembled in the Glogi apparatus (GA) after passing from ER (steps 1-3); Assembled hemichannels traffic into the plasma membrane with the aid from the microtubules (steps 4-5); gap junction plaques are formed after insertion of hemichannels into the PM and docking of entire gap junction channels with adjacent cell (steps 8-9); after a 1-5 hours short life time, whole gap junction plaques are removal from PM and move to the degradation process in either the lysosomes or proteasomes (steps 10-11).
The distribution of gap junction proteins depends on the tissue type except Cx43 is the most ubiquitous Cx protein. Gap junction proteins expressed in different tissues are associated with a large number of developmental diseases such as deafness (Cx26 and Cx30 in cochlea), Charot-Maire-Tooth disease (Cx32 in liver and Schwann cells), oculodentodigital dysplasia (Cx43) and congenital cataracts (Cx43 and Cx50 in lens). However, the specific molecular mechanism of these gap junctions contribute to these diseases or developmental events is still lacking. Our study provides important insight into the molecular mechanism of the intracellular calcium regulation of Cxs and the molecular basis for these diseases.

1.6. Calmodulin-mediated regulation of gap junctions.
CaM, the Ca\(^{2+}\)-sensing subunit of a broad variety of ion channels, has been found in a wide range of species from Homo to paramecium [98]. Gap junction proteins associate with a large number of signaling proteins as parts of multi-protein complexes and the plaque of gap junction channels function as an “island” in the plasma membrane (Fig. 1.10) [99]. While there is an increasing evidence that CaM participates in the operation of GJs [100-102], we have shown that the Ca\(^{2+}\)-dependent regulation of GJs Cxs is effected via the association of Ca\(^{2+}\)-CaM with the intracellular loop domain of Cx43, and Cx44, the sheep ortholog of human Cx46. [103, 104]. CaM also been found to co-localize with β-family connexin Cx32 [105] and directly gates Cx32-containing GJs [106]. In this study, a new CaMBD in Cx50 was identified and two additional CaMBDs were also explored in Cx43.

1.7. The objectives of this dissertation.
The objective of this research is to understand how intracellular Ca\(^{2+}\) signaling modulates gap junction proteins via the calcium sensor protein, calmodulin, and to investigate EF-hand Ca\(^{2+}\)-binding proteins using computational tools and protein engineering. This research focuses on four key objectives:

A. Elucidating the molecular basis underlying the CaM-connexin50 interaction in eukaryotic cells.
We predicted and identified a CaM-binding region in the second half of intracellular loop of connexin50 (Cx50). The Ca\(^{2+}\)-dependent interaction between the Cx50 and CaM was verified in vitro by various biophysical experiments. In vivo, the CaM-mediated regulatory effect on the conductance of gap junction formed by Cx50 is investigated in N2a cell through the collaboration with Dr. Richard Veenstra.

B. Investigating the multiple CaM binding sites of connexin43 at the molecular level in mammalian cells.

We predicted and identified two additional CaM binding sites in Connexin43 (Cx43) using various biophysical methods, including high resolution NMR and fluorescence. We proposed a model in which the regulation of Cx43 is mediated by multiple CaM binding sites.

C. Elucidating the metal-binding properties of connexin26 in extracellular loops.

The biological role of a potential calcium binding pocket in extracellular loop of connexin26 (Cx26) will be investigated by three approaches. 1) Calcium binding prediction algorithm Mug\(^{SR}\) developed in our lab was used to predict calcium binding pockets in the available Cx26 crystal structure. 2) A grafting approach was applied to investigate the metal binding properties of the isolated calcium binding loop and disease causing mutant from Cx26. 3) The metal binding properties of Cx26 was investigated by the whole gap junction channel isolated from rat liver.

D. Investigating the metal-binding properties of EF-hand motifs in various biological systems.

The metal-binding properties of the single isolated EF-hand in virus, bacteria and mammalian proteins was probed using a grafting approach. In addition, the oligomerization status of the EF-hand II in Calbindin D9k was examined by a combination of pulse-field gradient NMR and fluorescence anisotropy.

1.8. Overview of this dissertation.

The CaM-target sequences in the second half of the intracellular loop in Cx50 were synthesized and their ability to bind CaM was determined using a range of biophysical and electrophysiological approaches in. We also developed a model of the CaM-mediated regulation of gap junction formed by the alpha family connexins via a strategy with multiple CaM binding domains.
Chapter 2 in this dissertation summarizes all the methods and materials used in this study, including molecular cloning, protein engineering, protein expression and purification, spectroscopic techniques (UV, circular dichroism, fluorescence, NMR), bio-computational analyses of genomes, homology structure modeling, and biochemical assays. All mathematical equations used in this study are also listed in this chapter.

Chapter 3 in this dissertation is focused on the elucidation of the molecular basis underlying the interaction between Cx50 and CaM. Predicted CaM-target sequences in the second half of intracellular loop in Cx50 are synthesized and their ability to bind CaM is determined using a range of biophysical approaches in vitro and electrophysiology in vitro.

Chapter 4 in this dissertation focuses on exploring a model of CaM-mediated regulation of gap junction protein Cx43 in alpha family via a strategy with multiple CaM binding domains. Heteronuclear single-quantum coherence (HSQC) NMR, fluorescence and mass spectrometry were used to confirm the interaction between the peptides from the first half of intracellular loop and C-tail of Cx43. Residual dipolar coupling (RDC) NMR was used to determine the Cx43-CaM binding mode.

Chapter 5 is devoted to the prediction and identification of EF-hand motifs from various biological systems by a grafting approach. The calcium binding sites were identified in nonstructural protein 1 (nsP1) of Sindbis virus and poxviruses, streptococcal hemoprotein receptor (Shr) of *Streptococcus pyogenes*. The metal binding properties of two EF-hand motifs isolated from Calbindin D9k has also been discussed in this chapter.

Chapter 6 in this dissertation focuses on the mass spectrometry application on protein study. Protein modification such as dansylation, PEGylation and IAEDANS labeling on protein has been improved by mass spectrometry. Protein-protein interaction and protein oligomerization also have also been studied by MALDI-MS in this chapter.

Chapter 7 is a succinct summary of major discoveries of this dissertation.
Calciomics study identifies calcium-binding proteins (CaBPs) and the calcium-binding locations in proteins, which aids in the understanding of protein structures and biological functions.
Figure 1. 2 Calcium signaling in bacteria.

Intracellular calcium concentration in prokaryotes is tightly regulated in a range from 100 to 300 nM. In bacteria organisms, calcium ions are involved in the maintenance of cell structure, motility, transport and cell differentiation processes. Although a number of calcium binding proteins (CaBPs) have been isolated in several prokaryotic organisms, the role of calcium in prokaryotes is still elusive. Our study provides a novo and rapid approach to predict and identify CaBPs in prokaryotic system.
Figure 1. 3 Ca^{2+}-binding affinities and structures of different classes of Ca^{2+}-
binding proteins (CaBPs).
A, Ca^{2+} binds to different proteins with affinities varying from nM to mM. B, cartoon representation of EF-hand Ca^{2+}-binding motif. C, trigger proteins such as calmodulin (PDB code: 3cln) with four EF-hand Ca^{2+}-binding motifs undergo Ca^{2+}-
dependent global conformational changes upon association with target proteins. D, buffer proteins, such as calbindinD9K (PDB code: 1b1g) and parvalbumin, have been shown to buffer local Ca^{2+} concentration without significant conformational changes. E, Ca^{2+} binding to thermolysin (PDB code: 1clx) stabilizes the protein structure and confers thermal stability. http://etd.gsu.edu/theses/available/etd-11272007-155719/. This figure is from Dr.Yubin Zhou’s dissertation in our lab.
Figure 1. 4 Experimental approaches to validate predicted continuous Ca\textsuperscript{2+}-binding motifs.

Ca\textsuperscript{2+}-binding sites can be predicted by using the program CaPS (continuous primary protein sequence) and Mug (crystal structure). A, the predicted Ca\textsuperscript{2+}-binding sequence is subsequently grafted into a scaffold protein CD2.D1. B, the flow chart of calciomics. After expression and purification of the engineered protein, Tb\textsuperscript{3+} titration assay (Tb\textsuperscript{3+}-LRET) is used to confirm the metal binding capability. Ca\textsuperscript{2+} competition assay can be further performed based on Tb\textsuperscript{3+}-LRET to obtain Ca\textsuperscript{2+}-binding affinity. NMR and CD can be used to check whether host protein remain native after loops insertion. If metal binds to the engineered protein, mutations on the proposed ligand residues can be subsequently introduced to double confirm the binding event. If the mutagenesis results in a decrease in binding, the predicted site has a high chance of binding Ca\textsuperscript{2+}. Possible Ca\textsuperscript{2+}-induced conformational changes will be examined by expressing the protein of interest.
Figure 1.5 Calcium dependent Ca\textsuperscript{2+}-CaM-protein interaction.

Calcium activates calmodulin by calcium induced conformational change, which is an essential step to regulate the activity of smooth muscle myosin light chain kinase (M13).
Figure 1. 6 Binding modes between CaM and its targets.

Calmodulin can bind to its targets either in a collapsed or an extended conformation. In the collapsed conformation, the targets will provide 2 hydrophobic anchor residues which interact with the hydrophobic pockets of the N- and C-terminal domains of calmodulin. These hydrophobic residues usually occur in the 1-7, 1-10, 1-14, 1-16, or 1-17 positions in the targets sequences. The target usually binds in an antiparallel fashion, that is, residue 1 interacts with the pocket in the C-lobe, while the second residue interacts with the pocket in the N-lobe. Calmodulin can also bind in the extended conformation. It can bind the target only using the C-lobe, both in the holo and apo forms such as with calcium pump or IQ motif, respectively. It can also bind using both lobes such as with anthrax exotoxin and potassium channel. It can also induce formation of dimer. It can bind with 1:1, 2:2, and 1:2 stoichiometries. This figure is derived from Dr. Hing Wong’s drawing.
In the collapsed conformation, usually, the compactness or tightness of the complex is determined by the spacing of the two anchor residues. The farther away the two residues, the less compact of the complex obtained. Here we superimposed the first 2 helices of different complexes, and removing some helices for clearness, we can see as we go from 1-7 to 1-17 (indicated by the color code), helix 7 move farther away. In the superimposed figure, 1-16 structure is special because the peptide binds in the parallel fashion. The figure on the right shows the positions of residues on the peptide helix. The most common positions are indicated as red color. All possible positions are clustered together in a region. This figure is from Dr. Hing Wong based on the report of Mal et al [68].
Homomeric, Cxs in hemichannel are identical. Homotypic, hemichannels in GJ are identical.
I: Monomer addition pathway.
II: Trimer-trimer pathway.
III: dimer-dimer pathway.

Figure 1. Assembly of gap junction proteins.
A, the assemble pattern of different gap junction channels. B, the pathway of gap junction protein assembly. C, the tropology map of connexin monomer. D, the side view of hemi-channel in open and close stage. This figure is drawed based on the previous work [100, 132, 158]
Figure 1. Schematic representation of the steps that lead to synthesis, assembly, and degradation of gap junction membrane channels.

Gap junction biosynthesis and degradation involves (1) synthesis of connexin polypeptides at endoplasmic reticulum membranes, (2) oligomerization into homo- and heteromeric gap junction connexons (hemi-channels), (3) passage through the Golgi stacks, (4) intracellular storage within trans Golgi membranes, (5) trafficking along microtubules, (6) insertion of connexons into the plasma membrane, (7) lateral diffusion of connexons in the plasma membrane, (8) aggregation of individual gap junction channels into plaques, (9) stabilization of peripheral microtubule plus-ends by binding to Cx43-based gap junctions, (10) internalization of the channel plaque leading to cytoplasmic annular junctions, and (11) complete degradation via lysosomal and proteasomal pathways. This figure is taken from the review paper of Segretain and Falk [97].
Figure 1. 10 Possible domains of Cx43 involved in interactions with partner proteins.

This figure was modified based on the review paper of Herze, et al [99].
Table 1. The connexin family.

<table>
<thead>
<tr>
<th>Family</th>
<th>Human</th>
<th>Mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Protein name</td>
<td>Gene name</td>
</tr>
<tr>
<td>α</td>
<td>Cx43</td>
<td>GJA1</td>
</tr>
<tr>
<td></td>
<td>Cx46</td>
<td>GJA3</td>
</tr>
<tr>
<td></td>
<td>Cx37</td>
<td>GJA4</td>
</tr>
<tr>
<td></td>
<td>Cx40</td>
<td>GJA5</td>
</tr>
<tr>
<td></td>
<td>Cx50</td>
<td>GJA8</td>
</tr>
<tr>
<td></td>
<td>Cx59</td>
<td>GJA9 (GJA10)</td>
</tr>
<tr>
<td></td>
<td>Cx62</td>
<td>GJA10</td>
</tr>
<tr>
<td>β</td>
<td>Cx32</td>
<td>GJB1</td>
</tr>
<tr>
<td></td>
<td>Cx26</td>
<td>GJB2</td>
</tr>
<tr>
<td></td>
<td>Cx31</td>
<td>GJB3</td>
</tr>
<tr>
<td></td>
<td>Cx30.3</td>
<td>GJB4</td>
</tr>
<tr>
<td></td>
<td>Cx31.1</td>
<td>GJB5</td>
</tr>
<tr>
<td></td>
<td>Cx30</td>
<td>GJB6</td>
</tr>
<tr>
<td></td>
<td>Cx25</td>
<td>GJB7</td>
</tr>
<tr>
<td>γ</td>
<td>Cx45</td>
<td>GJC1 (GJA7)</td>
</tr>
<tr>
<td></td>
<td>Cx47</td>
<td>GJC2 (GJA12)</td>
</tr>
<tr>
<td></td>
<td>Cx30.2/Cx31.3</td>
<td>GJC3 (GJE1)</td>
</tr>
<tr>
<td>δ</td>
<td>Cx36</td>
<td>GJD2 (GJA9)</td>
</tr>
<tr>
<td></td>
<td>Cx31.9</td>
<td>GJD3 (GJC1)</td>
</tr>
<tr>
<td></td>
<td>Cx40.1</td>
<td>GJD4</td>
</tr>
<tr>
<td>ε</td>
<td>Cx23</td>
<td>GJE1</td>
</tr>
</tbody>
</table>

Former names are indicated in parentheses.
2. MATERIALS AND METHODS

2.1. Molecular cloning and plasmid construction

Recombinant Rat CaM gene was cloned into pET7 vector between Ndel (5’ end of gene) and BamHI (3” end of gene) restriction enzyme sites. Recombinant Rat domain 1 of CD2 (CD2.D1) gene was cloned into PGEX-2T vector between BamHI (5’ end of gene) and EcoRI (3” end of gene) restriction enzyme sites. The sequence encoding the putative Ca$^{2+}$-binding domain from extracellular loop I of Cx26 was inserted between Ser-52 and Gly-53 by using standard PCR methods. In order to maintain the native structure of inserted Ca$^{2+}$-binding loop, Gly linker was added at both end of the binding loop peptide (three Gly at N-temini and two Gly at C-termini). Similarly, predicted and known Ca$^{2+}$-binding EF-hand domain from nsP1 of Sindbis virus, nsP1 of POX virus, streptococcal hemoprotein receptor of Streptococcus pyogenes, pseudo EF-hand of Calbindin D9k loop 1 and classical EF-hand of Calbindin D9k loop 2 were inserted into CD2.D1. Disease related mutants and Ca$^{2+}$-binding elimination mutants were generated by primer design. Nucleotides replacement was introduced in the middle of forward primer. The antisense sequence of the reverse primer at 5’ end is next to the forward primer with DNA polymerase running in the opposite direction. A mutation generated linear product was ligated by T4 ligase at 16 °C for at least 4 hours. All the 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.

2.2. Baculovirus vector construction

Mouse-Cx26 and human-Cx26 gene was inserted into pFastBac™ 1 donor plasmid via BamHI and HindIII restriction sites. The recombinant was transferred into MAX efficiency DH10BacTM competent E.coli cells (containing bacmid and helper plasmid). Recombinant bacmid DNA was isolated from transferred E.coli after overnight growth. Sf9 insect cells were transfected by recombinant Bacmid DNA using Cellfectin II reagent. Baculovirus contains Cx26 gene was isolated from transfected Sf9 insect cell.
2.3. Expression and purification of proteins

**CD2 Grafted proteins and their mutants.** The engineered CD2 grafted proteins and their mutants were expressed as GST fusion proteins in Escherichia coli BL21 (DE3) transformed with the plasmid pGEX2T and grown in LB medium with 100 mg/L of ampicillin and grown at 37 °C. 100 μM 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 a 0.45 μM filter and injected into a Superdex 75 gel filtration column. The eluted proteins were pooled together and further purified by Histrip SP cation exchange chromatography. The calculated molecular weight of CD2 grafted protein and their mutants were also confirmed by MALDI-TOF-MS with sinapinic acid as matrix in linear fly mode. Their concentration was determined by using the absorption at 280 nm with an adjusted extinction coefficient according to sequence calculation (r62) in Table 2.1. Isoelectric point (pI) was also calculated for each CD2 engineering protein with Protein Calculator v3.3 (Table 2.1).

**CaM.** Recombinant rat CaM in the plasmid pET-7 was expressed in E. coli strain BL21(DE3)pLysS (r63). pET-7 transformed bacteria were grown in LB medium at 37 °C. All the expression methods followed the same protocol as grafted CD2 GST fusion protein. Isotope labeled CaM for NMR study was expressed in SV minimal medium using 0.5 g/L 15NH₄Cl as the only nitrogen source for sole isotopic labeling. For 15N and 13C double labeling CaM was expressed with an additional 5g/L glucose (13C₆H₁₂O₆) as only carbon source for protein synthesis. Isotopic 15NH₄Cl and glucose were purchased from
Cambridge Isotope Laboratories, MA, USA. Recombinant rat CaM was purified by phenyl-Sepharose (Sigma, MO, USA) chromatography as described previously [104].

**CaM modifications.** In order to monitor the conformational change of the entire CaM and CaM domain, several modifications was made in Rat CaM. Five mutants, T26W and Y99W for intrinsic aromatic group, T34C, T110C and additional an Cys at the C-termini for IAEDANS fluorescence dye labeling, were introduced into CaM, respectively. Dansyl CaM was prepared according to the method of Johnson et al with slight modifications [107]. In Brief, rat CaM was dansylated in the dark by mixing 1 ml protein (~500 µM) with a two to five-fold molar excess of dansyl chloride (dissolved in 1:1 acetone/ethanol) in 10 mM Mops, 100 mM KCl, 1 mM CaCl2, pH 6.5 to 7.5 overnight at 4 °C. IAEDANS labeled CaM (I-CaM) was prepared by following the condition as previously described [108]. Briefly, three CaM mutants with Cys (~ 300 µM) in 10 mM Mops, 100 mM KCl were degased by using a stirrer and vacuum for 20 mins. High puriry nitrogen gas filled the reaction bottle after adding IAEDANS (dissolved in DMF) with a two to three-fold molar excess at 4 °C overnight. The reaction bottle was enclosed with parafilm and aluminum foil to keep reaction in a reduced environment in the dark. For both D-CaM and I-CaM, the reaction mixture was then extensively dialyzed against 10 mM Tris, 100 mM KCl at pH 7.4 to remove the free fluorescence dye. The modification conditions of CaM by dansyl chloride and IAEDANS was improved by checking the reaction product based on the result of MALDI spectrum. The amount of bound dye was determined using ε335 of 3980 M⁻¹cm⁻¹ for D-CaM [107] and ε336 of 5700 M⁻¹cm⁻¹ for I-CaM [108]. The concentration of modified CaM was determined using the BCA assay (Pierce) with CaM as standard. An average of ~0.8 mol of the dansyl and ~0.6 mol of IAEDANS chromophore was incorporated into per mol of CaM.

2.4. **Peptide design and synthesize**

The Cx50 peptide Cx50p₁₄₁₋₁₆₆ (Ac⁻₁₄₁SSKGTKKKFRLEGTLRTYVCHIIFKT₁₆₆-NH₂) was synthesized by EZ Biolab Inc. and purified by preparative reversed-phase high pressure liquid chromatography with a purity of >87%. A scrambled peptide SCx50p (Ac-FKLYKC
ISFGGTEITTRSHVLTKR-NH₂), with the same composition of amino acids which shows no predicted CaM binding capacity, was synthesized with a purity of >90%, and served as a negative control. The molecular weight of these two peptides were determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-MS) in reflectron mode. To mimic the protein environment and remove extra charges, peptides were acylated at their N termini and aminated at their C termini.

2.5. Nuclear magnetic resonance spectroscopy

NMR experiments were performed with a Varian INOVA 600 MHz NMR spectrometer at 37 °C. ¹H-¹⁵N HSQC data were collected with 2048 complex data points and a spectral width of about 14 ppm in the ¹H dimension; 128 increments and a spectral width of 33 ppm in the ¹⁵N dimension. A sample of 0.26 mM uniformly ¹⁵N-labeled CaM in buffer consisting of 5 mM MES, 10 mM Bis-Tris (pH 6.5), 5 mM DTT, 0.1 mM NaN₃, 10 mM CaCl₂ or 10 mM EGTA, and 10% D₂O was titrated with peptide stock solutions (~5 mM) of the CaM binding peptides (CaMBP) such as Cx50p₁₄₁₋₁₆₆ or scramble peptides such as SCx50p in the same buffer. NMR data were processed using NMRPipe [109] and analyzed using Sparky [110].

The pulse-field-gradient diffusion NMR spectra were collected with a modified PG-SLED pulse sequence [111] on a 600 MHz Varian INOVA spectrometer. 8k complex data points with a spectral width of ~13 ppm were contained in each FID. The pulse field gradient level (G) was arrayed from ~0.2 to ~31.0 Gauss/cm with a pulse gradient time (δ) of 5 ms and a diffusion time (Δ) of 112.5 ms. The data were processed using FELIX (Accelrys). The relationship between the NMR signal intensity (A) and the diffusion constant (D) follows the equation
Equation 2.1.
\[ A = A_0 \exp\left[-(\gamma \delta G_2)(\Delta-\delta/3)D\right] \]
where \( A_0 \) is the signal intensity when pulse gradient is not used and \( \gamma \) is the gyromagnetic ratio of the proton. In the data processing, the signal intensity as a function of pulse field gradient level was fitted with Equation 2 using KaleidaGraph 3.5 (Synergy). The values discussed in this article were all determined by fitting with a linear correlation coefficient \( \geq 0.999 \).

Equation 2.2.
\[ A = A_0 \exp(-CG_2) \]
where \( C \) is a constant (equals to \( \gamma \delta (\Delta-\delta/3)D \)), the diffusion constant and the hydrodynamic radii \( R \) were obtained by comparing the \( C \) values of different molecules measured under identical conditions using the equations

Equation 2.3.
\[ D/D_0 = C/C_0 \]

Equation 2.4.
\[ R/R_0 = C_0/C \]
where \( D_0 \), \( R_0 \) and \( C_0 \) are the diffusion constant, hydrodynamic radius, and the measured \( C \) value of lysozyme. The hydrodynamic radius and diffusion constant of lysozyme have been reported previously [112]. The gradients were calibrated at 25 °C on the residual 1H signal in a sample of 99.9% D2O, using the published value of \( 1.902 \pm 0.002 \times 10^{-9} \) m² s⁻¹ for the self-diffusion coefficient of HDO at 25 °C [113].

NMR samples contained 0.23 mM CaM in 10% D2O, 100 mM KCl, and 10 mM imidazole at pH 6.5, with 10 mM Ca²⁺ for the Ca²⁺-bound form. The intensities of the protein signals were integrated from the
methylene and methyl region of ~2 ppm spectral width (-0.2–1.8 ppm). The integrated regions for each species were carefully selected to avoid or reduce interferences from buffer signals. The integrated intensities were further normalized to minimize the experimental errors from phase adjustment and baseline correction during the processing.

Chemical shift perturbations (Δ) of the 1H-15N HSQC spectra with and without peptide were calculated using both 1H and 15N chemical shifts (δ) as shown in Equation 5:

\[ Δδ = \sqrt{\left(\frac{Δδ^1H}{2}\right)^2 + \left(\frac{Δδ^{15N}}{5}\right)^2} \]

2.6. Fluorescence spectroscopy

Steady-state fluorescence spectra were recorded using a QM1 fluorescence spectrophotometer (PTI) in a 1 cm path length cell with a xenon short arc lamp at 25 °C. Intrinsic tryptophan (Trp) emission spectra were recorded using 2 to 6 µM protein samples in 50 mM Tris-100 mM KCl at pH 7.4. The Trp fluorescence spectra were recorded from 300 to 400 nm with the excitation wavelength at 282 nm. The slit widths were set at 4 and 8 nm for excitation and emission, respectively. For Tyr/Trp-sensitized Tb\textsuperscript{3+} luminescence energy transfer (Tb\textsuperscript{3+}-LRET) experiments, emission spectra were collected from 500 to 600 nm with the excitation at 282 nm and the slit widths were set at 8 and 12 nm for excitation and emission, respectively. To circumvent secondary Raleigh scattering, a glass filter with a cutoff of 320 nm was used. The Tb\textsuperscript{3+} titration experiments were performed by gradually adding 5-10 µl aliquots of Tb\textsuperscript{3+} stock solutions (10 mM) into the protein samples (6 µM) in 20 mM PIPES, 100 mM KCl at pH 6.8 to prevent precipitation. For the Ca\textsuperscript{2+} competition studies, the solution containing 30 µM of Tb\textsuperscript{3+} and 1.5 µM of protein was set as the starting point. The stock solution of 10-100 mM CaCl\textsubscript{2} with the same concentration of Tb\textsuperscript{3+} and protein was gradually added in 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 the equation

\[
\text{Equation 2.6.}
\]

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

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 was first analyzed to derive the apparent dissociation constant by equation 1. By assuming that the sample is saturated with Tb$^{3+}$ at the starting point of the competition, the Ca$^{2+}$-binding affinity is further obtained by using the equation

\[
\text{Equation 2.7.}
\]

\[
K_{d,Ca} = \frac{K_{d,Tb}}{K_{d,Tb} + [Tb]}
\]

where \( K_{d,Ca} \) and \( K_{d,Tb} \) are the dissociation constants of Ca$^{2+}$ and Tb$^{3+}$, respectively. \( K_{app} \) is the apparent dissociation constant.

The fluorescence emission spectra were acquired between 400 and 600 nm with an excitation wavelength at 335 nm for D-CaM. The slit widths were set at 4 nm for excitation and 8 nm for emission. A solution (0.8 ml) containing 0.25-2 µM D-CaM in 10 mM Tris-HCl, 100 mM KCl, pH 7.4, with 5 mM Ca$^{2+}$ or 5 mM EGTA was titrated by gradually adding 5-10 µl aliquots of the peptide stock solution (0.25-0.8 mM) in the same buffer. The binding constant of the synthetic peptide to modified CaM was obtained with a 1:1 binding model (Equation 2.6) by fitting normalized fluorescence data as described previously [114, 115]. pH profile of binding affinity between peptide and D-CaM was determined by various buffer system listed in Table 2.2.

The fluorescence anisotropy of D-CaM before and after addition of CaMBP was measured with excitation at 335 nm and emission at 500 nm. An integration time of at least 20 seconds (20 points per
second) was used to record the fluorescence signal, and all measurements were repeated at least three times by following established protocols with Equation 2.8 [116, 117].

Equation 2.8.

\[
r = \frac{I_H - I_V}{I_H + 2I_V}
\]

where the \( r \) stands for fluorescence anisotropy and the \( I \) terms indicate fluorescence intensity measurement parallel (\( I_H \)) and perpendicular (\( I_V \)) to the incident polarization.

Due to domain specific distribution of phenylalanine and tyrosine in CaM, the equilibrium Ca\(^{2+}\) binding constants for CaM were determined by monitoring fluorescence of intrinsic phenylalanine (\( \lambda_{ex} = 250 \) nm, \( \lambda_{em} = 280 \) nm) for N-domain or tyrosine (\( \lambda_{ex} = 277 \) nm, \( \lambda_{em} = 320 \) nm) for C-domain at 25 °C as described [118]. For N or C-domain Ca\(^{2+}\) binding constants, 5-8 µM CaM was titrated with 5-10 µl aliquots of 50 or 15 mM Ca\(^{2+}\) stock solution in Ca\(^{2+}\) equilibrium buffer with 50 mM HEPES, 100 mM KCl, 5 mM NTA, 0.05 mM EGTA, pH 7.4. The same method was applied for the 1:1 CaM-peptide mixture. Free Ca\(^{2+}\) at each titration point was determined with the Ca\(^{2+}\) dye Oregon Green 488 BAPTA-5N (0.2 µM; \( K_d = 21.7 \pm 2.5 \) µM) at emission of 520 nm with excitation of 495 nm [115]. The free concentration was obtained with the equation:

Equation 2.9.

\[
[Ca^{2+}]_{\text{free}} = K_d \left( \frac{F - F_{\text{min}}}{F_{\text{max}} - F} \right)
\]

Where \( F \) is the fluorescence intensity of the dye at each titration point; \([Ca^{2+}]_{\text{free}}\) is the concentration of free ionized calcium in solution; \( F_{\text{min}} \) and \( F_{\text{max}} \) represent the dye fluorescence intensities of Ca\(^{2+}\) -free form and Ca\(^{2+}\) -saturated dye, respectively. The Ca\(^{2+}\) binding affinity was obtained by fitting to the nonlinear Hill equation:
Equation 2. 10.

\[ f = \frac{[Ca^{2+}]_{\text{free}}^n}{K_d + [Ca^{2+}]_{\text{free}}^n} \]

Where \( f \) is the fractional change of intrinsic fluorescence intensity; \([Ca^{2+}]_{\text{free}}\) is the concentration of free ionized calcium in solution; \( K_d \) represents \( Ca^{2+} \) dissociation constants and \( n \) is the Hill coefficient.

2.7. Circular dichroism spectroscopy

Circular Dichroism (CD) spectra were recorded in the far ultraviolet (UV) (190-260 nm) range on a Jasco-810 spectropolarimeter at ambient temperature using a 0.1-cm path length quartz cuvette. The measurements of CaM and CaM-CaMBP complex (10 µM) were made in 10 mM Tris, 100 mM KCl, at pH 7.5 with 5 mM CaCl\(_2\) or 5 mM EGTA. All spectra presented were averaged for at least 15 scans and the background signal from the buffer was removed from the sample signals. The far UV CD spectra of the peptide in different percentages of trifluoroethanol (TFE) were obtained using a concentration of 20 µM CaMBP in the same buffer. The secondary structure contents of the peptides were calculated with the online secondary structure prediction server DICHROWEB [119] The \( \alpha \)-helical content of peptides was predicted using Agadir algorithm [120-122].

2.8. Mass spectroscopy

The MALDI mass spectrometry analysis was performed on an Applied Biosystems 4800 plus MALDI TOF/TOF analyzer mass spectrometer (Framingham, MA). The data were acquired in a linear positive mode with sinapinic acid (SA) as matrix for CaM (50 µM) and CaM- CaMBP complex. The molecular weight of the CaMBP was also confirmed by MALDI in reflectron positive mode with \( \alpha \)-Cyano-4-hydroxycinnamic acid (CHCA) as matrix. 50 µM CaM and 50 µM CaMBP were mixed in 100mM KCl, 50mM Tris-HCl, at pH 7.5. The reaction was equilibrium at least half an hour. A 1 µl mixture was added with 10 µl saturated SA solution and then dried on the MALDI plate for the measurement.
2.9. Electrophysiological measurement of junctional conductance

Parental N2a cells were grown to 80% confluency in a 12 well format and were transiently transfected with 1 µg of the pTracer-Cx50 cDNA plasmid for four hours, lightly trypsinized, and placed in 35 mm culture dishes overnight. Dual whole cell patch clamp experiments were performed the next day on GFP-expressing cell pairs to measure gap junction conductance (g_j) and the sensitivity to [Ca^{2+}]_i was determined by permeabilizing the cells with 1 µM ionomycin and perfusing with ±1.8 mM external CaCl_2 saline solutions. A +20 mV transjunctional voltage (V_j) pulse was applied every 15 sec to monitor g_j during bath perfusion with nominally zero or 1.8 mM external Ca^{2+} ([Ca^{2+}]_o) saline solutions. The g_j was normalized to the initial value (G_j = g_j/g_j, initial) for each Cx50 cell pair and the results were pooled for each data set. To examine the effect of CaM inhibitor or the candidate CaM binding peptide on g_j, Cx50-N2a cells were pretreated with 1 µM calmidazolium (CDZ; Calbiochem) for 10-15 min or 1 µM of the synthetic Cx50 mimetic peptides (Cx50p_{141-166} or SCx50p) were added to both whole cell patch pipettes.

2.10. Prediction of protein transmembrane topology and calmodulin binding domains

Based on a hidden Markov model, the topology and orientation of the transmembrane helices of the mouse Cx50 were predicted using four different programs including TMHMM, MEMSAT, SOSUI and HMMTOP [123-126]. ClustalW2 was used to align multiple mammalian gap junction polypeptide sequences in the alpha-family, including Cx43 (accession ID: NP_000156; Gja1 human), Cx44 (accession ID: AAD56220; Gja3 sheep), Cx46 (accession ID: NP_068773; Gja3 human) and Cx50 (accession ID: AAF32309; Gja8 human). The candidate CaM binding region in Cx50 was predicted using the CaM Target Database developed on the basis of over 100 known CaM target sequences [62].
Table 2.1 Summary of GST fusion protein of CD2 variants.

<table>
<thead>
<tr>
<th>CD2 Grafted Protein</th>
<th>Molecular Weight (Da)</th>
<th>pI</th>
<th>$\varepsilon_{280\text{nm}} (\text{M}^{-1} \cdot \text{cm}^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD2.Cal1$^a$</td>
<td>14832</td>
<td>6.16</td>
<td>12980</td>
</tr>
<tr>
<td>CD2.Cal2$^a$</td>
<td>14640</td>
<td>5.05</td>
<td>11700</td>
</tr>
<tr>
<td>CD2.Pox$^b$</td>
<td>14616</td>
<td>8.36</td>
<td>14260</td>
</tr>
<tr>
<td>CD2.Shr$^c$</td>
<td>14480</td>
<td>7.19</td>
<td>11700</td>
</tr>
<tr>
<td>CD2.Sin$^d$</td>
<td>14753</td>
<td>8.99</td>
<td>18670</td>
</tr>
<tr>
<td>CD2.Sin II$^d$</td>
<td>14721</td>
<td>9.41</td>
<td>18670</td>
</tr>
<tr>
<td>CD2.Cx26.EL1$^e$</td>
<td>15763</td>
<td>5.34</td>
<td>19950</td>
</tr>
<tr>
<td>CD2.Cx26.EL2$^e$</td>
<td>15763</td>
<td>5.31</td>
<td>19950</td>
</tr>
<tr>
<td>CD2.Cx26.EL1 G45E$^e$</td>
<td>15835</td>
<td>5.17</td>
<td>19950</td>
</tr>
<tr>
<td>CD2.Cx26.EL1 D46E$^e$</td>
<td>15777</td>
<td>5.34</td>
<td>19950</td>
</tr>
<tr>
<td>CD2.Cx26.EL1 E47K$^e$</td>
<td>15762</td>
<td>5.97</td>
<td>19950</td>
</tr>
<tr>
<td>CD2.Cx26.EL1 D50N$^e$</td>
<td>15762</td>
<td>5.57</td>
<td>19950</td>
</tr>
</tbody>
</table>

a) The grafted EF-hand loops are from loop1 (Cal1) and loop2 (Cal2) of calbindin D9k, respectively.
b) The predicted EF-hand loop from nsP1 of POX virus was grated into CD2 host protein.
c) The predicted EF-hand loop from streptococcal hemoprotein receptor of *Streptococcus pyogenes* was grated into CD2 host protein.
d) The predicted EF-hand loop and calcium binding elimination mutant from Sindbis virus was grated into CD2 host protein.
e) The predicted calcium binding pockets in extracellular loop1 (EL1) and loop2 (EL2) from connexin26 was grated into CD2 host protein. Disease mutations, G45E, D46E, E47K and D50N were introduced as well.
Table 2.2 Common salt for buffer control in different pH.

<table>
<thead>
<tr>
<th>Common Name</th>
<th>$pK_a$ at 25°C</th>
<th>Buffer Range</th>
<th>Temp Effect $d\ pH/d\ T$ in (1/K)</th>
<th>Mol. Weight</th>
<th>Full Compound Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAPS</td>
<td>8.43</td>
<td>7.7–9.1</td>
<td>−0.018</td>
<td>243.3</td>
<td>$3{[\text{tris(hydroxymethyl)methyl}]\text{amino}}$propanesulfonic acid</td>
</tr>
<tr>
<td>Bicine</td>
<td>8.35</td>
<td>7.6–9.0</td>
<td>−0.018</td>
<td>163.2</td>
<td>N,N-bis(2-hydroxyethyl)glycine</td>
</tr>
<tr>
<td>Tris</td>
<td>8.06</td>
<td>7.5–9.0</td>
<td>−0.028</td>
<td>121.14</td>
<td>tris(hydroxymethyl)methylamine</td>
</tr>
<tr>
<td>Tricine</td>
<td>8.05</td>
<td>7.4–8.8</td>
<td>−0.021</td>
<td>179.2</td>
<td>N-tris(hydroxymethyl)methylglycine</td>
</tr>
<tr>
<td>HEPES</td>
<td>7.48</td>
<td>6.8–8.2</td>
<td>−0.014</td>
<td>238.3</td>
<td>4-2-hydroxyethyl-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>TES</td>
<td>7.40</td>
<td>6.8–8.2</td>
<td>−0.020</td>
<td>229.20</td>
<td>2-{$[\text{tris(hydroxymethyl)methyl}]$amino}$ethanesulfonic acid</td>
</tr>
<tr>
<td>MOPS</td>
<td>7.20</td>
<td>6.5–7.9</td>
<td>−0.015</td>
<td>209.3</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>PIPES</td>
<td>6.76</td>
<td>6.1–7.5</td>
<td>−0.008</td>
<td>302.4</td>
<td>piperazine-N,N’-bis(2-ethanesulfonic acid)</td>
</tr>
<tr>
<td>MES</td>
<td>6.15</td>
<td>5.5–6.7</td>
<td>−0.011</td>
<td>195.2</td>
<td>2-(N-morpholino)ethanesulfonic acid</td>
</tr>
</tbody>
</table>
3. MOLECULAR INTERACTION AND FUNCTIONAL REGULATION OF CONNEXIN50 BY CALMODULIN

3.1. Introduction

The circulation of current fluid flow through the close packed cells of the lens of the eye is essential for maintaining lens homeostasis, and is highly dependent on functional cell-to-cell coupling. This is accomplished by gap junctions, relatively nonselective channels that are permeable to low molecular weight (<1 kDa) molecules that directly link the cytoplasm of adjacent cells [127-129]. The gap junction channel formed between two adjacent cells is comprised of two connexon hemichannels, one from each cell, joined in mirror symmetry. Each hemichannel is composed of six transmembrane connexin (Cx) proteins embedded in the plasma membrane of the cell. The mammalian lens is comprised of three different Cx proteins namely Cx43, Cx46 and Cx50 that like all other Cxs have four transmembrane segments, a short N-terminal cytoplasmic region, one intracellular loop, and two extracellular loops with a set of highly conserved cysteine residues; the major difference between all connexins is in the sequences of their cytoplasmic loops and C-terminal tails [129].

Inhibition of cell-to-cell communication in the lens by elevated intracellular Ca$^{2+}$ concentration ([Ca$^{2+}$]$_i$) was first demonstrated as an increased internal electrical resistance [130] that was prevented by pre-incubation with calmodulin (CaM) antagonists. Elevated [Ca$^{2+}$]$_i$ also inhibits cell-to-cell communication in both bovine [131] and sheep [132] lens primary cell cultures. We have demonstrated that lens cell-to-cell communication was half-maximally inhibited at ~ 300 nM [Ca$^{2+}$]$_i$ [132] and was prevented by pre-incubation of lens cultures with CaM antagonists [133]. The rapid onset of this inhibition (within seconds) suggests that this is mediated by a direct interaction of CaM with one or more of the lens Cxs rather than being mediated via the action of a CaM-dependent protein kinase. Indeed Peracchia et al. have demonstrated that CaM directly gates Cx32-containing gap junctions [100], and Török et al [134] have identified two distinct CaM binding amino acid sequences in Cx32 with the N- and
C-lobes of CaM showing separate functions, suggesting trans-domain or trans-subunit bridging by CaM as a possible mechanism of gap junction gating [135]; notably this CaM binding motif in Cx32 is absent from the three lens Cxs.

To better understand the Ca\(^{2+}\)-dependent inhibition of lens gap junctions we identified high affinity CaM binding sites in the lens connexins Cx43, and Cx44 (the sheep homolog of human Cx46). Specifically, these domains encompass residues 138-157 in Cx43 [114] and residues 129-148 in Cx44 (corresponding to residues 138-157 in Cx46) [115]. These sequences have conserved hydrophobic residues at positions 1, 5 and 10 and 14 that matched well with two major CaM binding classes, 1-10 and 1-14, which are similar to the CaM binding motifs found in calcineurin, nitric oxide synthase, adenylyl cyclase and skeletal myosin light chain kinase [62, 136]. These CaM binding domains are located at the C-terminal end of the loop region between the predicted Cx transmembrane regions 2 and 3 [124, 128, 129, 137-139] and are highly conserved in all three lens Cxs.

The study reported here was conducted to determine whether Cx50, like Cx43 and Cx44, was also regulated by [Ca\(^{2+}\)], in the μM range, and if so whether this Ca\(^{2+}\) regulation was CaM-mediated. A bioinformatic analysis [62] of Cx50 resulted in the identification of a candidate CaM binding domain with high predictive score that was located in the same C-terminal end of the loop region between the predicted Cx transmembrane regions 2 and 3 of Cx50. This region is highly conserved in all three lens Cxs. A peptide model, proven to be successful in our studies with Cx43 and Cx44 [114, 115], was then adopted to demonstrate the physical interaction between the predicted Cx50 CaM binding domain and CaM. Studies using high-resolution NMR and fluorescence spectroscopy clearly established a high affinity, Ca\(^{2+}\)-dependent interaction between a domain in the intracellular loop of Cx50 and CaM. Our results provide important new insight into the molecular mechanism by which altered regulation of intracellular Ca\(^{2+}\) concentration in the lens effects the regulation of lens Cx50 gap junctions.
3.2. Results

3.2.1. A putative CaMBD in the single cytoplasmic loop of Cx50

The CaM binding site predication server, Calmodulin Target Database, was used to identify potential CaM binding domains within Cx50, by taking into account the hydropathy, alpha-helical propensity, residue weight, residue charge, hydrophobic residue content and helical class [62]. The highest predictive score was assigned to a stretch of sequences (residues 141-166; Fig. 3.1) in the C-terminal portion of the intracellular loop of Cx50. Sequence alignments with previously characterized CaMBDs in rodent Cx43 [114] and sheep Cx44 (human Cx46) [115], and secondary structure prediction, identified common features among the three alpha-family connexins that included: i) the spacing of hydrophobic residues with bulky side chains, often used as anchor in CaM-target complexes [140-142], followed a 1-5-10 pattern; ii) all of the domains exhibited strict conservation in positions with positively-charged residues, which help drive the formation of CaM-target complexes via electrostatic interactions [143]; and iii) the CaMBDs were predicted to have a predominantly alpha-helical structure. Helical wheel analysis further suggests the clustering of hydrophobic residues on one side, and the frequent occurrence of charged residues on the opposite side. A peptide corresponding to the predicted CaMBD was subsequently synthesized for the functional and biophysical studies described below.

3.2.2. Ca^{2+}-dependent uncoupling of Cx50 gap junctions requires CaM

To determine whether Cx50 gap junctions like those comprised of Cx43 [114, 144] and Cx46 [115] are regulated by [Ca^{2+}], we first performed dual whole cell patch clamp experiments in Cx50-N2a cells. Perfusion with saline solution containing 1 µM ionomycin + 1.8 mM [Ca^{2+}]_o induced a time-dependent reduction in G_j (Fig. 3.2A) (Table 3.1); we have shown previously that such treatment effects an elevation of [Ca^{2+}]_i in HeLa cells [145]. Omission of 1.8 mM CaCl_2 from the bath saline prevented the Cx50 gap junction uncoupling response in the presence of 1 µM ionomycin.
To test whether the Ca\textsuperscript{2+}-dependent decrease in Cx50 G\textsubscript{j} produced by addition of 1 μM ionomycin + 1.8 mM [Ca\textsuperscript{2+}]	extsubscript{o} required CaM, Cx50-N2a cell cultures were acutely (≤ 15 min) treated with calmidazolium (CDZ) (2 μM), a CaM inhibitor with a dissociation constant (\(K_d\)) of ≈1 μM [146]. Pretreatment with CDZ, prevented the Ca\textsuperscript{2+}-dependent decrease in G\textsubscript{j} (Fig. 3.2B).

A more specific test for a role of Cx50 CaMBD in mediating the Ca\textsuperscript{2+}-dependent inhibition of this Cx was to add the 26-mer Cx50 mimetic peptide (Cx50p) corresponding to residues 141-166 of the cytoplasmic loop to the Cx50-transfected N2A cells. A scrambled peptide containing a randomized sequence of the same amino acids was also utilized as a control. The addition of the Cx50p, but not the scrambled peptide, prevented the decline in G\textsubscript{j} (Fig. 3.2C). Overall, these electrophysiological data support the hypothesis that an elevation of [Ca\textsuperscript{2+}]\textsubscript{i} can uncouple Cx50 gap junctions in a CaM-dependent manner mediated by the Cx50 141-166 domain.

To examine the molecular basis for the Ca\textsuperscript{2+}-dependent decrease in Cx50 G\textsubscript{j}, gap junction channel recordings were analyzed in two poorly coupled cell pairs. Both experiments exhibited a maximum of three open channels (N) and their cumulative open probability (N\cdot P_o) declined from 1.21 open channels per unit time to zero upon exposure to 1 μM ionomycin and 1.8 mM [Ca\textsuperscript{2+}]	extsubscript{o} saline (Fig. 3.3A). Unitary Cx50 gap junction channel currents remained constant (314 ± 18 pS, std, n = 14) during the decline in P_o, indicative of a conformational gating type behavior (Fig. 3.3B).

These experiments in this section were performed by Qin Xu and Dr. Xianming Lin in Dr. Veenstra's laboratory.

### 3.2.3. Ca\textsuperscript{2+}-dependent direct interaction between Cx50 CaMBD and CaM observed by MALDI-MS and SDS-PAGE with cross-linking.

After confirming the requirement for CaM in the Ca\textsuperscript{2+}-dependent regulation of the Cx50 gap junctions, we further asked the question whether or not there is a direct interaction between the predicted CaMBD of Cx50 (or Cx50p) and CaM. The formation of the CaM-Cx50p complex was first confirmed by MALDI
mass spectrometry (Fig. 3.4). In the presence of Ca\(^{2+}\), a distinct peak representing the CaM-Cx50p complex (Theoretical MW: 19.93 kDa; Experimental MW: 20.02 kDa.) was detected. The peak intensity in MALDI spectrum cannot be used as a quantitative method to determine the percentage of complex because the complex was formed in solution face with hydrophobic interaction and salt bridge while all the condition was changed when the complex sample was co-crystalized on MALDI plate. The complex peak in the spectrum only can confirm the direct interaction.

The Ca\(^{2+}\)-dependent direct interaction between the peptide including the CaMBD of Cx50 and CaM was also observed by SDS-PAGE electrophoresis with 0.2 % glutaradehyde as cross-linking reagent (Fig. 3.5). In calcium loaded solution, the increase of molecular weight of the protein sample demonstrated that glutaradehyde is able to cross-link Cx50p and Ca\(^{2+}\)-CaM. When the peptide concentration increased, more of the complex was formed in the 1:2 and 1:3 CaM to Cx50p ratio. There was no complex observed in SDS-PAGE gel in EGTA condition. The dimer form of Ca\(^{2+}\)-CaM and complex also can be observed by cross-linking which shows a band (> 29 kDa). Glutaradehyde is a frequently used amine-reactive homobifunctional crosslinker reagent with the shortest space arm which only can link the strong binding with close interaction. After 10 mins reaction at 37°C, the amine groups of lysine residues between CaM and Cx50p were cross-linked by glutaradehyde. Glutaradehyde contains almost shortest space arm (~1.8 Å) between two reaction groups. It provides brief estimation about the distance between CaM and the gap junction peptide in the complex. The mechanism in chemistry of glutaradehyde remains unclear [147].

### 3.2.4. Ca\(^{2+}\)-dependent specific interaction between Cx50 CaMBD and CaM revealed by NMR

To identify residue-specific changes in the conformation of CaM during complex formation with Cx50p, we carried out \(^1\)H, \(^{15}\)N) HSQC NMR experiments. Unlabeled Cx50p peptide stock solution was gradually titrated into \(^{15}\)N-labeled CaM. Under Ca\(^{2+}\)-loaded condition, the HSQC spectrum of Ca\(^{2+}\)-saturated CaM (holo-CaM) underwent significant changes on addition of 1 molar eq of Cx50p (Fig. 3.6). The scrambled Cx50 peptide, which is predicted to have no binding affinity to CaM, was similarly added.
to $^{15}\text{N}$-labeled Holo-CaM. This scrambled peptide failed to induce chemical shift changes in CaM (Fig. 3.7), therefore ruling out the possibility of non-specific interactions between peptide and CaM.

By analyzing the NMR chemical shifts of CaM-Cx50p complexes deposited in the Biological Magnetic Resonance Bank (BMRB codes: 1634, 4270, 5480, 5770, 5893, 5896, 15470, and 16465), we identified some significant changes in chemical shifts ($\delta$(CaM) - $\delta$(complex) > 1 S.D.) that are common to all complexes such as the $^1$H shifts of A57 and M71, and $^{15}$N shift of A128. In the classical CaM-CaMKII$\alpha$ peptide structure with 1-10 binding mode, M71 and A128 are in the hydrophobic pockets of CaM and are within 5 Å to the Cx50p. A57 occupies the second position of the second Ca$^{2+}$ binding loop, suggesting that peptide binding leads to a conformational change in this loop. In the CaM-Cx50p complex, A57 and M71 have $\delta_H$ of 0.29 and 0.05 ppm, respectively, while A128 has $\delta_N$ of 0.21 ppm, consistent with the Cx50p bound to Ca$^{2+}$-CaM.

Upon formation of a Ca$^{2+}$/CaM-Cx50p complex, the most significant perturbations, with weighted average chemical shift changes greater than 0.1 ppm, were observed in both the N-lobe (e.g., S17, T26, T29, E31, L39, A57) the C-lobe (e.g., E82, D93, L105, R106, A147, K148) and the linker region (e.g., D64, F68), indicating that Cx50p causes a global conformation changes in CaM. A detailed residue-to-residue plot of chemical shift perturbation is shown in Fig. 3.8. The overall change of the N-lobe residues (2.74 ppm) was greater than in the C-lobe residues (1.78 ppm). During the titration, we observed the simultaneous appearance of two peaks representing the same residue. For example, a progressive disappearance of the amide signal of residues G33 representing unbound CaM was accompanied by the concomitant emergence of a new set of peaks from the Cx50p-CaM complex (Fig. 3.9). Such a slow exchange process, which occurs when the exchange rate of the bound and unbound states is smaller than the amide frequency difference, often heralds a high affinity protein-protein or protein-peptide association (i.e. interactions with submicromolar affinities) [115, 148].
3.2.5. Probing the CaM-Peptide complex state by diffusion NMR

Pulse-field gradient NMR was performed to determine the diffusion constant of the CaM-Cx50p complex, enabling us to assess the overall size of this complex in solution. The diffusion constant was determined to be \((12.4 \pm 0.4) \times 10^7\) cm/s (Fig. 3.10). By assuming a spherical shape of the complex, the calculated size of the Ca\(^{2+}\)/CaM-Cx50p complex was determined to be 17.3 ± 0.6 Å, a value comparable to the hydrodynamic radii of CaM in complex with well-known targets such as smMLCK (17.9-21.8 Å), PDE (18.8-22.3 Å), and CaMKI (21.2 Å) [149]. The hydrodynamic radius of CaM (22.6 ± 0.6 Å) was decreased by 23% on formation of the CaM-Cx50p complex. This size conforms to a collapsed structure of the complex that involves the unwinding of the central helix within CaM to embrace the target peptide. Further exploring the interaction of apo-CaM with Cx50p using NMR was not possible due to limited solubility of the peptide in 5 mM EGTA (Fig. 3.11).

3.2.6. Revealing CaM-Peptide interactions with CD spectroscopy

CD spectra also confirmed the structural and conformational changes during the formation of the CaM-peptide complex. As indicated in Fig. 3.12, the addition of the Cx50p to holo-CaM at a 1:1 molar ratio resulted in a ~6% increase in absolute value of the negative ellipticity at 222 nm. The hydrophobic environment in the peptide binding pocket of holo-CaM could be the reason for the increase in the absolute value of negative ellipticity attributed to the peptide itself. In contrast, a 5% decrease at 222 nm and a 20% decrease at 208 nm in the CD signal intensity were observed after the addition of the Cx50p peptide to apo-CaM. The data suggested that Cx50p may bind with apo-CaM in different region which is not the same as holo-CaM. It is not clear why the helical content of the complex decreased in the presence of EGTA.

We further examined the secondary structure of the Cx50p peptide under different solvent conditions. TFE is known to induce \(\alpha\)-helical formation of peptides and the helical content in the presence of TFE usually reflects its helical propensity [150]. The Cx50p peptide becomes predominantly helical (> 90%) at TFE ≥ 30% (v/v) (Fig. 3.13), suggesting that the Cx50p CaM binding sequence readily adopts an \(\alpha\)-
helical configuration. Cx50, Cx44 and Cx43 exhibit 94, 55 and 33% helical conformation, respectively, in the presence of 30% TFE (Table 3.2). As shown in Fig. 3.1, Cx50p contains both the positive charge residue R (R for Cx44p and K for Cx43p) and negative charge E (A for Cx44p and R for Cx43p) before and after hydrophobic residue I at position 1 which may facilitate and stabilize its helical conformation.

The peptide in the absence of CaM has only 7% alpha helix and 72% random coil. The addition of peptide into CaM resulted in the addition of helicity as shown in Fig. 3.12. Fig. 3.14 shows that our difference CD spectrum of the CaM in the presence and absence of the peptide exhibits two negative minima at 208 and 222 nm. To understand the origin of the additional helicity due to the formation of the CaM-peptide complex, we analyzed the helical contents in the X-ray structure of Ca\(^{2+}\)-loaded calmodulin alone (pdb: 3CLN) and the complex structure of CaM-kinase II peptide (pdb:1CDM) since our chemical shift studies suggest that CaM binds to Cx50p in a similar binding mode as to the peptide of CaM-kinase II (Fig. 3.15). CaM secondary structure almost remains the same in its helical content upon formation of the complex with a helicity about 62% (Table 3.3). Therefore, the complex induced helicity (about 20% helical calculated by DICHROWEB based on peptide spectrum after subtraction of CaM signal) is attributed to the conversion of peptide after binding with Ca\(^{2+}\)-CaM. Table 3.3 summaries the helicity change of CaM upon formation of the complex in different CaM binding modes. The maximum helical content change of CaM complex is equal or less than 4% compare to CaM itself (62%) despite the increase or decrease. Thus, about 20% helical change introduced by adding Cx50p is mainly from the peptide itself even considering other different CaM binding Mode.

### 3.2.7. Determination of the CaM-Peptide binding affinity by fluorescence spectroscopy

Dansyl-labeled calmodulin (D-CaM) has been frequently used to determine the binding affinity of CaM-peptide interaction due to its sensitivity to the changes in the surrounding chemical environment and its ease of preparation [151]. The dansyl moiety has an emission maximum at \(~500-510\) nm, a range in which the signals from intrinsic aromatic residues (Phe, Tyr and Trp) are negligible. We first examined the dansyl fluorescence anisotropy changes to confirm the formation of complex in the low micromolar
range. Fluorescence anisotropy is dependent on the rotational correlation time of fluorophore in the sample, and is often used to reflect the hydrodynamic properties of biomolecules [152]. The increase in anisotropy often arises from slower tumbling of the fluorophore and thus reports the formation of a larger complex. Indeed, upon addition of 1 molar eq of Cx50p, the anisotropy of the dansyl moiety within D-CaM increased from 0.083 to 0.125, suggesting the association of Cx50p with Ca\(^{2+}\)-CaM (Fig. 3.16).

We then carried out the titration of peptide with D-CaM by monitoring the fluorescence emission between 400 nm and 600 nm. In the presence of Ca\(^{2+}\), D-CaM showed a fluorescence maximum at 500 nm. With the addition of Cx50p, the dansyl fluorescence emission maximum blue-shifted to 474 nm with a concomitant enhancement of its fluorescence intensity (Fig. 3.17) implying that the dansyl group entered a more hydrophobic environment upon complex formation with peptide. In contrast, in the presence of EGTA (Fig. 3.18), although the dansyl fluorescence intensity also increased, the fluorescence maxima changed less than in the presence of Ca\(^{2+}\). The data confirms the CD signal change on the addition of peptide to apo-CaM. By fitting the titration curve with a 1:1 binding mode, we obtained a dissociation constant \((K_d)\) of 4.9 ± 0.6 nM for the interaction between Ca\(^{2+}\)-CaM and Cx50p and thousands times lower \(K_d\) for the interaction between apo-CaM and Cx50p (Table 3.2).

To further investigate the possible role of pH on CaM-peptide association, the effect of varying pH on the binding affinity of CaM for Cx50p was also examined. As shown in Fig. 3.19, the binding affinity (represented as \(-\log K_d\)) exhibited a pH-dependent change between pH 5.0 to pH 10.0. The highest binding affinity was obtained at pH 6.5 and the binding affinities from pH 6 to 9 almost remain the same level. The observed pH-dependence of the interaction between CaM and Cx50p indicates that electrostatic interactions are likely the main force driving CaM-Peptide complex formation.

3.2.8. Binding of Cx50p expands the \([Ca^{2+}]_i\) sensing range of CaM

We next performed the Ca\(^{2+}\) equilibrium titration of CaM to obtain macroscopic Ca\(^{2+}\)-binding constants for both the N- (phenylalanine) and C-lobes (tyrosine) of CaM by monitoring domain-specific
fluorescence changes as described previously [117, 118]. The decrease in phenylalanine fluorescence ($\lambda_{ex} = 250$ nm, $\lambda_{em} = 280$ nm) and the increase in tyrosine fluorescence ($\lambda_{ex} = 277$ nm, $\lambda_{em} = 320$ nm) report Ca$^{2+}$ binding to the N-lobe and the C-lobe of CaM, respectively. This allows us to examine the effect of peptide binding on the domain-specific Ca$^{2+}$ sensing capability of CaM. As shown in Fig. 3.20, Fig. 3.21 and summarized in Table 3.4, the addition of Cx50p led to a right shift of the titration curve representing the Ca$^{2+}$ binding to the N-lobe, and concomitantly caused a left shift of the titration curve reflecting the Ca$^{2+}$ binding to the C-lobe of CaM. Specifically, the Ca$^{2+}$ affinity of the N-lobe of Cx50p-bound CaM decreased by 2 fold, whereas the Ca$^{2+}$ affinity of the C-lobe increased by ~20%. As a result of Cx50p binding, CaM responses to subtle changes in [Ca$^{2+}$], over a broader range of calcium concentrations.

3.3. Discussion

3.3.1. Regulation of Cxs by intracellular Ca$^{2+}$ and CaM

The results reported here demonstrate that intracellular calcium is important in the regulation of Cx50 gap junctions [145]. Furthermore, using the CaM inhibitor CDZ as well as a peptide fragment encompassing the predicted CaM binding region, we are able to demonstrate that this Ca$^{2+}$-dependent inhibition of Cx50 gap junctions is mediated by the ubiquitous intracellular Ca$^{2+}$-receptor CaM. Thus CaM appears to mediate the Ca$^{2+}$-dependent regulation of all three major alpha connexins in the lens, i.e. Cx43, Cx46 and Cx50. Our results also demonstrate that this inhibition of Cx50 gap junction channels by Ca$^{2+}$ is due to a decrease in channel open time probability as opposed to a reduction in single channel conductance (Fig. 6.3AB). These records provide the first direct evidence for the gated closure of individual connexin gap junction channels by Ca$^{2+}$-CaM. The average observed $\gamma_j$ value of 314 pS is indicative of Cx50 since it is at least double the observed $\gamma_j$ for Cx40, Cx43, or Cx45 gap junctions in our N2a cell cultures and consistent with the reported 210-290 pS values from reconstituted lens fiber cell membranes or exogenously expressed Cx50 gap junction channels [153-158].
The possible role of CaM in the chemical gating of the beta family Cxs was first proposed by Peracchia and colleagues [159]. CaM colocalizes with Cx32 [134] and directly gates Cx32-containing GJs [160]. Török et al. [134] reported that fluorescent labeled CaM derivatives bind to synthetic peptides spanning most of the cytoplasmic sequences from two regions of Cx32: an N-terminal sequence comprising residues 1-21 (K_d = 27 nM) or residues 1-19 (K_d = 1.1 µM); and a C-terminal sequence comprising residues 216-230 (K_d = 2.1 µM) or residues 208-226 (K_d = 3.5 µM). Both exhibit Ca^{2+}-dependent CaM-binding properties [134, 135]. They also identified two distinct CaM binding amino acid sequences in Cx32 within the N- and C-lobes of CaM showing separate functions, suggesting trans-domain or trans-subunit bridging by CaM as a possible mechanism of GJ gating [160]. In addition, Ahmad et al using in vitro translation approaches have shown that oligomerization of Cx32 is CaM-dependent [161] since CaM interacts with Cxs at an early stage of GJ assembly [100, 134]. The C-terminal tail of the CaM binding site of Cx32 is likely involved in the assembly of Cxs and is Ca^{2+} dependent [100]. Furthermore, Blodow et al reported that CaM antagonists suppress gap junction coupling of Cx26 in isolated Hensen cells of the guinea pig cochlea [162]. Taken together, [Ca^{2+}]_i and CaM play vital roles in regulating gap junctions mediated by both the alpha and beta family of Cxs.

3.3.2. **Comparison of CaM binding affinity with the alpha family Cxs**

Using high resolution NMR, CD and fluorescence studies, we have shown using a peptide model that CaM interacts with the predicted cytosolic CaM binding region of Cx50. Our current studies, as well as previous studies, have clearly shown that CaM exhibits a Ca^{2+} dependent interaction of all three peptide fragments encompassing the intracellular loop regions of all three lenses Cxs. Fluorescence spectroscopy revealed conformational changes of both the peptide and CaM following formation of the CaM-Cx peptide complex. NMR studies demonstrated that the peptide binds to CaM with a 1:1 stoichiometry. It also indicated that the peptide induces structural changes in both the N- and C-terminal domains, as well as in the linker region of CaM and the binding of Cx peptides to CaM reflects a classical embracing mode of interaction. Overall for these three connexins, binding of the peptide to CaM decreases the apparent K_d.
of Ca\(^{2+}\) for CaM, and the Hill Coefficient \(n_{\text{Hill}}\) increased. In addition, for the peptide from the intracellular loop, Cx43 exhibit weaker affinity for CaM than either Cx50 or Cx44.

Consistent with the CaM-Cx peptide binding affinity studies using fluorescence spectroscopy, our NMR chemical shift analysis also supports a lower affinity for the CaM-Cx43 peptide complex. The addition of Cx43 and Cx50 peptides results in similar changes in the CaM spectra as shown in Fig. 3.22. For most signals, the directionalities of the changes are the same, reflecting the same 1-5-10 binding mode of these peptides to CaM. However, we have observed the slow exchange process for residue Gly33 in CaM shown in Fig. 3.9. A similar scenario was observed for residues T29, A57 and T117 for the CaM-Cx44 peptide complex [115]. Such a slow exchange process occurs when the exchange rate of the bound and unbound states is smaller than the amide frequency difference, and often heralds a high affinity protein-protein or protein-peptide association (i.e. interactions with submicromolar affinities) [148] that is consistent with the CaM binding affinity to Cx50p and Cx44p peptides. In contrast, the NMR signals of CaM upon addition of Cx43 peptide exhibits a gradual change in chemical shifts characteristic of the fast exchange regime, suggesting a weaker binding affinity.

### 3.3.3. Implications for the factors contributing to the CaM binding affinity to Cxs

There are several positive charged residues such as K147, R149 and R156 in Cx50 that are conserved in all alpha family members. To identify the key factors responsible for the binding affinity of CaM to Cxs, we have determined the CaM binding affinity of the Cx50 peptide to CaM as a function of pH. The peptide binding affinity is decreased at pH values lower than 5.5 and greater than 8.5. This trend is very similar to our previous results reported for Cx44 [115]. Both Cx44 peptide and Cx50 peptide behave similarly when pH is varied, but CaM binding to the Cx50 peptide is ~ 1 order of magnitude stronger than to the Cx44 peptide. Such a similar pH dependence suggests that electrostatic interactions following the protonation of Asp and Glu and deprotonation of Lys are important for this complex formation.

The difference in CaM binding affinity of Cx50 and Cx44 peptides may be originate from the intrinsic sequences encoded in the CaM binding regions of different Cxs. As shown in Fig. 3.1, there are
several differences in the predicted regions especially with variations between conserved residue positions 1 and 5. Cx43 contains two flexible Gly instead of the helical forming residues in Cx44 and Cx50. Such sequence variation reflects their ability to form α-helices that are important for their interaction with CaM [62]. As shown in Fig. 3.13 and Table 3.2, Cx43 peptide has the lowest α-helical formation capacity. The addition of TFE (> 60%) induces the highest α-helical content in the Cx50 and Cx44 peptides. In Fig. 3.13, the α-helicity of Cx50 peptide is higher than that of the Cx43 and Cx44 peptides reported previously [114, 115]. The calculated α-helical content of peptides in TFE exhibits the rank order Cx50p > Cx44p > Cx43p. Interestingly, negatively charged E151 located close to R149 in Cx50 may stabilize the α-helical conformation of this peptide. In contrast, Cx43 has a positive charged R at the same location that may disable α-helical conformation. The observed α-helical content of the CaM binding peptides are consistent with the trend for the CaM binding affinity results from our fluorescence experiments (Table 3.2). Both Cx44 and Cx50 peptides also bind CaM in the absence of Ca^{2+} as revealed by fluorescence data. The CaM binding affinities to Cx44 and Cx50 in EGTA are 100- and 1000-fold weaker than those in the presence of Ca^{2+}, respectively. The apo-form of CaM does not interact with the Cx43 peptide which has the least α-helical content. Therefore, our results reveal that the α-helicity of the CaM binding peptide is an important factor in predicting the CaM binding affinity of these Cxs.

Our detailed NMR analysis has also revealed the differential interactions of CaM with the Cx50 and Cx43 peptides. We observed that in general the N-lobe of CaM-Cx50p has larger chemical shift changes, while the C-lobe has smaller changes than those of CaM-Cx43p although such interactions belong to a similar 1-5-10 binding mode. In addition, chemical shifts of residues such as F12, K13, K94, L116, I130, and M145 of CaM-Cx50p changed in different directions compared with those of CaM-Cx43p (Fig. 3.22). Among these residues, K94 and I130 are on the third and fourth Ca^{2+} binding loops respectively, and these loops could be flexible and adopt different conformations, leading to different directions of change in chemical shifts. In the structure of CaM-CaMKIIα (pdb code 1CDM), L116 interacts with K298, which corresponds to K146 in Cx43 and R149 in Cx50. Similarly, M145 interacts with K300, which
corresponds to R148 in Cx43 and E151 in Cx50 (Fig. 3.15). Such changes in directions of the chemical shifts of CaM are likely due to the intrinsic properties such as the \( \alpha \)-helical propensity of the sequences in the CaM binding regions of Cxs.

3.3.4. Structural and disease implications

The intracellular region regulated by CaM has been reported to be involved in several disease related mutations. Oculodentodigital dysplasia (ODDD) is caused by mutations in Cx43 such as G138R, G138S, G143S, G143D, K144E, V145G, M147T, R148G, R148L, T154A and T154N [163], overlapping with CaM binding region identified previously [114]. There have been no reports of point mutations in the CaM binding domain of Cx50 associated with cataract formation. However, Beahm and Hall (57) have reported that the Cx50 H161N mutant does not form detectable hemichannels but forms gap junctions indistinguishable from wild type. Interestingly, Toluoe (58) identified three nonfunctional mutations in Cx50, T157C, H161C, and E169C, that failed to function homotypically or in heterotypic pairings with wt Cx50. The threonine at position 157 is highly conserved and this residue has been shown previously to be functionally essential for all connexins tested, including Cx26, Cx43, and Cx50 [164]. Mutation of a conserved threonine in the third transmembrane helix of alpha- and beta-connexins creates a dominant negative closed gap junction channel [164]. It is possible that such mutations may compromise the regulation of this Cx by CaM. The testing of such mutational effects on Cx50 function is currently underway in this laboratory. The difficulty of testing these mutations is they could produce "dead" gap junctions by not trafficking to the membrane or forming nonfunctional channels.

The molecular scenario of CaM-mediated regulation of the alpha-class of gap junctions seems to be quite different from the beta-class, probably due to the fact that the alpha-class connexins lack the N-terminal glycine hinge and possess longer intracellular loops and C-termini than the beta-class Cxs. Indeed, the CaM targeting sequences in the alpha-class connexins, such as rodent Cx43, sheep Cx44 (or human Cx46), and mouse Cx50 as revealed in this work, are in the cytoplasmic loop connecting TM2 and TM3, in close proximity to the CL-TM3 interface. The apparent juxtaposition of the Cx32 N-terminal
CaMBD to TM1 and of the Cx43 CL CaMBD to TM3 may facilitate the closing of the connexin channel pore since both TM1 and TM3 are suggested to be involved in forming the transmembrane pore while the N-terminus loops back to form a pore cytoplasmic vestibule.

The recently determined 3D structure of the connexin 26 gap junction channel at 3.5 Å resolution provides insights into the gating mechanism of the beta-class of gap junction at atomic resolution [165]. In this tsuzumi (Japanese drum)-shaped structure, the permeation pathway is mainly defined by the short N-terminal helix (NTH, residues 2-10) that lines the pore funnel of gap junction channel, and the first transmembrane domain (TM1) that constitutes the pore-lining helix. This structure suggests that the molecular identity of the channel “plug” observed under electron microscopy [166] is likely the N-terminal helix. However, due to the invisibility of both the cytoplasmic loop and the C terminus in the crystal structure, both of which play important roles in the chemical gating of gap junction [167], it is still too premature to explain the chemical gating mechanism with the current structure.

Given that the N-terminal CaMBD in Cx32 overlaps with the NTH that is believed to form the channel “plug” in beta-class gap junction, it is conceivable that the Ca²⁺-dependent binding of CaM to NTH might function as “cork” to block the channel by the complex per se or by inducing conformational changes in NTH. However, since the NTH is lined interior to the channel to form a funnel, its accessibility to CaM in the assembled channel still remains to be confirmed.

Despite the absence of X-ray structural information about the intracellular loop of Cxs [165], we have provided detailed structural information about the association of CaM with all three alpha connexins members. Our results provide the first direct evidence that CaM binds to a specific region of the lens gap junction connexins Cx43, Cx44 and Cx50 in a Ca²⁺-dependent manner [114, 115]. Our data suggests a common conformational gating mechanism by which the Ca²⁺-dependent inhibition of the α-class of gap junction proteins is mediated by the direct association of an intracellular loop region of these proteins with Ca²⁺-CaM [132, 133, 144, 145].
3.4. **Summary**

Connexin 50 (Cx50), a member of the α family of gap junction proteins expressed in the lens of the eye, has been shown to be essential for normal lens development. In the present study, we identified a calmodulin (CaM) binding domain (CaMBD, residues 141-166) in the intracellular loop of Cx50. Elevations in intracellular Ca\(^{2+}\) concentration effected a 95% decline in junctional conductance (g\(_j\)) of Cx50 in N2A cells that is likely mediated by CaM, because inclusion of the CaM inhibitor, calmidazolium, prevented this Ca\(^{2+}\)-dependent decrease in g\(_j\). The direct involvement of the Cx50 CaMBD in this Ca\(^{2+}\)/CaM-dependent regulation was further demonstrated by inclusion of a synthetic peptide encompassing the CaMBD in both whole cell patch pipettes, which effectively prevented the intracellular Ca\(^{2+}\)-dependent decline in g\(_j\). Biophysical studies using NMR and fluorescence spectroscopy further reveal that the peptide stoichiometrically binds to Ca\(^{2+}\)/CaM with an affinity of ~5 nM. The binding of the peptide expanded the Ca\(^{2+}\) sensing range of CaM by increasing the Ca\(^{2+}\) affinity of the C-lobe of CaM, while decreasing the Ca\(^{2+}\) affinity of the N-lobe of CaM. Overall, these results demonstrate that the binding of Ca\(^{2+}\)/CaM to the intracellular loop of Cx50 is critical for mediating the Ca\(^{2+}\)-dependent inhibition of Cx50 gap junctions in the lens of the eye.
The α-class of connexins are composed of four transmembrane (TM) segments, two extracellular loops, one cytoplasmic loop, a short N-terminus and a much longer C-terminal tail. The predicted CaM binding sites are located in the second half of the intracellular loop between TM2 and TM3. The numeric score (1-9) represents the probability of an accurate prediction of high affinity CaM binding sites. The identified CaM-binding sequences in the α-class connexins are similar to the Ca²⁺-CaM-dependent kinase II (CaMKII). All the aligned sequences fit the 1-5-10 CaM-binding mode subclass, where each number represents the presence of a hydrophobic residue. m, mouse; h, human; s, sheep; r, rat; #, hydrophobic residues (highlighted in gray); B, basic residues (underscored).
Figure 3. 2 Ca\textsuperscript{2+}- and CaM-dependence of Cx50 gap junction uncoupling.

A, Cx50-N2a cell pair superfusion with 1 µM ionomycin bath saline ± 1.8 mM CaCl\textsubscript{2} demonstrates a time-dependent uncoupling of Cx50 gap junctions that requires 1.8 mM extracellular [Ca\textsuperscript{2+}]. B, pretreatment with 2 mM calmidazolium for 10-15 min prior to patch clamp analysis completely prevented the 1.8 mM [Ca\textsuperscript{2+}]\textsubscript{o}-dependent uncoupling response. C, Inclusion of 1 µM Cx50p\textsubscript{141-166} peptide in both whole cell patch pipettes prevented 80% of the Ca\textsuperscript{2+}/CaM-dependent decline in Cx50 G\textsubscript{j}. In contrast, experiments performed with the scrambled control SCx50p peptide prevented only 20% of the 1.8 mM [Ca\textsuperscript{2+}]\textsubscript{o}-CaM-dependent reduction of Cx50 G\textsubscript{j}. The average initial g\textsubscript{j} ± SEM values for all experimental groups are displayed in Table 3.1. These experiments were performed by Qin Xu and Dr. Xianming Lin in Dr. Veenstra's laboratory.
Figure 3. 3 Mechanistic basis for Ca$^{2+}$/CaM-dependent Cx50 gap junction uncoupling.

A. The product of the number of open channels (N) and open probability ($P_o$) from two poorly coupled Cx50-N2a cell pairs is plotted as a function of time and 1 µM ionomycin + 1.8 mM [Ca$^{2+}$]o superfusion. N x $P_o$ declined from an average control value of 1.21 open channels to zero within two minutes of ionomycin/CaCl$_2$ saline superfusion.

B. Cx50 gap junction channel current recordings from one experiment are shown for the time points indicated by the circles in panel A. The number and duration of the open channels declined progressively to zero without any apparent reduction in the single channel conductance (314 pS), indicative of closure of a Ca$^{2+}$-CaM-dependent gate without a block of the ion permeation pathway. These experiments were performed by Qin Xu and Dr. Xianming Lin in Dr. Veenstra's laboratory.
Figure 3. Monitoring CaM-Cx50\textsubscript{141-166} complex formation by mass spectroscopy.
The MALDI-MS spectra of the free form of Ca\textsuperscript{2+}-CaM (~17 kDa) and the complex form of CaM-Cx50\textsubscript{141-166} (~20 kDa). The inset showed the molecular weight of the Cx50\textsubscript{141-166} peptide (~3 kDa). [Cx50p] stands for Cx50\textsubscript{141-166} peptide. Sinapic acid (10 mg/ml) was mixed with protein sample as matrix.
Figure 3. 5 Cross-linking SDS-PAGE electrophoresis of CaM-Cx50p complex. A, SDS-PAGE results of CaM-Cx50p complex with cross-linking. B, the structure of glutardehyde. 100 µM CaM (20 mM Mops, pH 7.4) in 1mM calcium chloride or 1 mM EGTA was mixed with Cx50p141-166 as ratio 1:1, 1:2 and 1:3 respectively. Cross-linking reagent, 0.2% (final concentration) glutardehyde was added and incubate at 37 °C for 10 mins. 1 µl 1 M Tris-HCl with large amount of amine group, pH 8.0 was added into each sample to stop reaction. In calcium loaded solution, the increase of molecular weight of protein sample demonstrate that glutardehyde is able to cross-link Cx50p141-166 and Ca²⁺-CaM. There is no complex observed in SDS-PAGE gel with calcium free condition. The dimer form of Ca²⁺-CaM and complex also can be stabilized by cross-link which shows a band (> 29 kDa).
Figure 3. Monitoring the interaction between CaM and Cx50 peptides by \(^{(1}\text{H},^{15}\text{N})\text{-HSQC}\) spectroscopy. An overlay of HSQC spectra of holo-CaM (red) with the spectrum of the holo-CaM-Cx50p\(_{141-166}\) (green). The black box indicates residue in CaM with dramatic change in chemical shift such as G33, T29, A57 and D64.
Figure 3. 7 Cx50 Scramble peptide eliminates the interaction with holo-CaM by by (1H, 15N)-HSQC. An overlay of the HSQC spectra of holo-CaM (red) with the spectrum of the holo-CaM-SCx50p (green). Comparing holo-CaM spectrum, no chemical shift change can be observed after adding Cx50 scramble peptide. The result indicates that the certain sequence order of Cx50 CaMBD is specific for the interaction.
Figure 3. 8 Chemical shift perturbation in CaM induced by addition of 2-fold molar excess of Cx50p_{141-166}. The weight average chemical shift change ($\Delta \delta$) was calculated using Equation 2.5. Residues with $\Delta \delta > 0.1$ ppm (highlighted in red) were mapped to the three-dimensional structure of holo-CaM (Protein Data Bank code: 3cln). The C-termini exhibits less conformational change than N-termini by calculated chemical shift. The highlighted residues in red color was considered as potential interaction face between CaM and Cx50p_{141-166}. 
Figure 3. The chemical shift change of Gly33 during titration of holo-CaM with Cx50p_{141-166}. The disappearance of the peak (free form) was accompanied by the appearance of the corresponding peak (bound form) at a new position. The observation for the disappearance and the appearance of the peaks in 1:0.4 and 1:1 indicates the slow exchange during the binding process. The strong binding between CaM and Cx50p_{141-166} was confirmed by the slow exchange of chemical shift for Gly33 in CaM.
Figure 3. Hydrodynamic radii of CaM-Cx50p\textsubscript{141-166} complex determination by PFG-NMR. The NMR signal decay of Holo-CaM (○) and Holo-CaM-Cx50p\textsubscript{141-166} complex (●) as a function of field strength. The calculated hydrodynamic radii of the CaM and complexes are indicated on the top. The decrease of hydrodynamic radii of the complex indicates the compact binding mode in which peptide could be embraced by both N-lobe and C-lobe of CaM.
Figure 3. Monitoring the interaction between Apo-CaM and Cx50 peptides by \(^{(1}H, 15N)\)-HSQC spectroscopy. An overlay of HSQC spectra of apo-CaM (red) with the spectrum of the apo-CaM-Cx50\(_{141-166}\) (green). White heavy precipitation was observed in NMR tube after adding high concentration (5 mM) Cx50 peptide. The solubility of the peptide is much better in 5 mM CaCl\(_2\) than in 5 mM EGTA. The intensity of CaM in the HSQC spectrum (Green) was decreased but no chemical shift can be observed because CaM may precipitate as well.
Figure 3. 12 CD studies of the interaction between CaM and the Cx50<sub>141-166</sub> peptide. Far UV CD spectra of CaM in the presence of 5mM EGTA (○ Apo-CaM) or CaCl<sub>2</sub> (□ Holo-CaM), and a CaM-Cx50<sub>141-166</sub> (1:1) mixture with 5mM EGTA (● Apo-CaM-Cx50<sub>141-166</sub>) or CaCl<sub>2</sub> (■ Holo-CaM-Cx50<sub>141-166</sub>).
Figure 3. 13 CD Studies of the helicity of Cx43, Cx44 and Cx50 peptide. Far UV spectra of the synthetic peptide Cx50p141-166 with addition of TFE (0\% - 80\%). The inset shows calculated \( \alpha \)-helical content as a function of TFE concentration for the peptides Cx50\_141-166 (■), Cx44\_129-150 (▲) and Cx43\_136-158 (●).
Figure 3. 14 CD Studies of Cx50p₁₄₁₋₁₆₆ in free and complex form. Far UV spectra of the synthetic peptide Cx50₁₄₁₋₁₆₆ with (●) or without (○) addition of CaM. The Cx50p became helical upon binding in the hydrophobic core of CaM.
Figure 3.15 Structural basis of the difference in directionality of chemical shift change. The structure of CaM-CaMKII complex (pdb code: 1CDM) was shown with the peptide (inside) and the CaM residues (outside). L116 and M145 of CaM are within 5 Å to K298p and K300p of the peptide, respectively. These peptide positions correspond to K146p and R148p for Cx43, and R149p and E151p for Cx50, as can be seen in the superposition of their sequences. In Cx50, E151p may be pulled away from M145 by K147p. On the contrary, in Cx43, R148p may be pushed toward M145 by K144p. These differences in peptide sequences can cause changes of chemical shifts in different directions.
Figure 3. The dansyl fluorescence anisotropy.

D-CaM (1 μM) was titrated with Cx50_{141-166} in 100mM KCl, 5mM CaCl₂, 50mM Tris-HCl at pH 7.5. The fluorescence anisotropy were measured at λ_{ex} = 335 nm and λ_{em} = 495 nm with an integration time of 20s. The inset shows the anisotropy change for D-CaM (free), and the complex of D-CaM-Cx50p (bound). All experiments were repeated in triplicate.
Figure 3. The interaction of Cx50_{141-166} peptide with holo Dansyl-CaM (D-CaM) monitored by steady-state fluorescence.

The titration curve of D-CaM (2 µM) with Cx50p_{141-166} in the presence of 5mM Ca^{2+} in a buffer consisting of 100mM KCl, 50mM Tris, pH 7.5. The inset shows the fluorescence spectrum of D-CaM in the absence (dashed line) of in the presence (solid line) of an equivalent molar concentration of Cx50p_{141-166}. X axis on the top line indicates the ratio of Cx50p_{141-166} over CaM and the bind ratio is 1:1 according to the fitting curve (at 498 nm). A clear blue shift of the fluorescence emission spectrum of holo D-CaM caused by adding Cx50p_{141-166} indicates the conformation change of D-CaM introduced by the interaction. All experiments were repeated in triplicate.
Figure 3. The interaction of Cx50_{141-166} peptide with apo Dansyl-CaM (D-CaM) monitored by steady-state fluorescence.
The titration curve of D-CaM (2 μM) with Cx50p_{141-166} in the presence of 5mM EGTA in a buffer consisting of 100mM KCl, 50mM Tris, pH 7.5. The inset showed the fluorescence spectrum of D-CaM in the absence of (dashed line) in the presence of (solid line) an equivalent molar concentration of Cx50p_{141-166}. In the Ca^{2+} free solution, although the dansyl fluorescence intensity also increased, the fluorescence maxima changed less than in the presence of Ca^{2+} and the binding is much weaker indicated by the dissociation constant in Table 3.2. The data confirms the CD signal change on the addition of peptide to apo-CaM. All experiments were repeated in triplicate.
Figure 3. pH dependence of Cx50p141-166 and Cx44p129-150 binding to CaM. The binding affinities were derived from the peptides titration curves of D-CaM at various pH. All experiments were repeated in triplicate. The binding affinity (represented as $-\log K_d$) exhibited a pH-dependent change between pH 5.0 to pH 10.0. The highest binding affinity was obtained at pH 6.5 for Cx50p141-166 and 7.5 for Cx44p129-150.
Figure 3. 20 N-Domain specific equilibrium Ca$^{2+}$ titration of CaM (○), and CaM in complex with Cx50p$_{141-166}$ (●).

The intrinsic phenylalanine fluorescence ($\lambda_{\text{ex}} = 250$ nm, $\lambda_{\text{em}} = 280$ nm) was monitored to report the equilibrium Ca$^{2+}$-binding constants of the N-lobe of CaM. The free ionized Ca$^{2+}$ concentration was measured using the Ca$^{2+}$ indicator dye Oregon Green 488 BAPTA-5N. All experiments were repeated in triplicate in 100 mM KCl, 5 mM NTA, 0.5 mM EGTA, 50 mM HEPES, pH 7.5. The Ca$^{2+}$ affinity of the N-lobe of Cx50p-bound CaM decreased by 2 fold.
Figure 3. 21 C-Domain specific equilibrium Ca$^{2+}$ titration of CaM (○), and CaM in complex with Cx50p_{141-166} (●).

The intrinsic tyrosine fluorescence ($\lambda_{\text{ex}} = 277$ nm, $\lambda_{\text{em}} = 320$ nm) was monitored to report the equilibrium Ca$^{2+}$-binding constants of the C-lobe of CaM. The free ionized Ca$^{2+}$ concentration was measured using the Ca$^{2+}$ indicator dye Oregon Green 488 BAPTA-5N. All experiments were repeated in triplicate in 100 mM KCl, 5 mM NTA, 0.5 mM EGTA, 50 mM HEPES, pH 7.5. Upon binding with Cx50p_{141-166}, the Ca$^{2+}$ affinity of the C-lobe of CaM increased by ~20%.
Figure 3. Comparison of the $\alpha$ family of connexins.
An overlay of the HSQC spectra of holo-CaM (yellow), holo-CaM-Cx50p141-166 (green), and holo-CaM-Cx43p136-158 (violet). Arrows indicated that most signals shifted to the same direction due to identical binding mode, while rectangles highlighted some exceptions.
### Table 3.1 Electrophysiological measurement.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Initial ( g_j ) (ns)</th>
<th>SEM (nS)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.8 mM CaCl₂</td>
<td>9.90</td>
<td>2.90</td>
<td>5</td>
</tr>
<tr>
<td>0 mM CaCl₂</td>
<td>14.74</td>
<td>6.51</td>
<td>4</td>
</tr>
<tr>
<td>2 mM CDZ</td>
<td>16.79</td>
<td>2.95</td>
<td>4</td>
</tr>
<tr>
<td>Cx50p₁₄₁₋₁₆₆</td>
<td>7.05</td>
<td>2.53</td>
<td>7</td>
</tr>
<tr>
<td>SCx50p</td>
<td>3.36</td>
<td>1.80</td>
<td>6</td>
</tr>
</tbody>
</table>

\( g_j = I_j/V_j \) (Ohm’s Law, junctional current divided by transjunctional voltage).

nS = nanoSiemens, S = inverse Ohm (g = 1/R).

SEM = standard error of the mean.

N = number of cell pairs recorded from each data set.

Data collected by Qin Xu and Xianming Lin in Dr. Veenstra’s laboratory.
Tabel 3. 2 Binding affinities to CaM and α-helicity of the Cx peptides.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Dissociation constants (K_d, nM)</th>
<th>α-helicity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5mM Ca²⁺</td>
<td>5mM EGTA</td>
</tr>
<tr>
<td>Cx50p141-166</td>
<td>4.9 ± 0.6</td>
<td>&gt; 8000</td>
</tr>
<tr>
<td>Cx44p129-150</td>
<td>49 ± 3.0</td>
<td>&gt; 5000</td>
</tr>
<tr>
<td>Cx43p136-158</td>
<td>860 ± 20</td>
<td>nd</td>
</tr>
</tbody>
</table>

All the experiments were repeated at n = 3.

nd, not detectable.

*Dansyl-CaM fluorescence in 50 mM Tris-HCl, 100 mM KCl, pH 7.5.

*Prediction of the helical content of peptides using Agadir algorithm.

§Far UV CD measurement with 30% TFE in 10 mM Tris, 100 mM KCl, at pH 7.5.
### Tabel 3.3 Summary of Change of CaM helicity in different binding modes.

<table>
<thead>
<tr>
<th>PDB. ID</th>
<th>Peptide in Complex</th>
<th>Binding Mode</th>
<th>Number of Helical Residue in CaM</th>
<th>CaM Helicity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3CLN</td>
<td>nd</td>
<td>nd</td>
<td>96 ± 4</td>
<td>62 ± 3</td>
</tr>
<tr>
<td>2HQW</td>
<td>NR1C1</td>
<td>1-7</td>
<td>88 ± 8</td>
<td>59 ± 5</td>
</tr>
<tr>
<td>2F3Y</td>
<td>cardiac Ca(v)1.2</td>
<td>1-10</td>
<td>94 ± 7</td>
<td>58 ± 5</td>
</tr>
<tr>
<td>1CDM</td>
<td>CaM Kinase II</td>
<td>1-10</td>
<td>84 ± 7</td>
<td>62 ± 5</td>
</tr>
<tr>
<td>1MXE</td>
<td>CaM Kinase I</td>
<td>1-14</td>
<td>96 ± 8</td>
<td>64 ± 5</td>
</tr>
<tr>
<td>1QS7</td>
<td>RS20</td>
<td>1-8-14</td>
<td>95 ± 8</td>
<td>65 ± 6</td>
</tr>
<tr>
<td>1QTX</td>
<td>RS20</td>
<td>1-8-14</td>
<td>89 ± 8</td>
<td>60 ± 5</td>
</tr>
<tr>
<td>1IQ5</td>
<td>CaM Kinase Kinase</td>
<td>1-16</td>
<td>90 ± 8</td>
<td>60 ± 6</td>
</tr>
<tr>
<td>2BCX</td>
<td>RyR</td>
<td>1-17</td>
<td>86 ± 7</td>
<td>58 ± 5</td>
</tr>
</tbody>
</table>
### Table 3.4 Effects of Cx peptides binding on the metal-binding properties of CaM.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>N-domain (sites I and II)*</th>
<th>C-domain (sites III and IV)#</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_d$ (µM) $n_{Hill}$</td>
<td>$K_d$ (µM) $n_{Hill}$</td>
</tr>
<tr>
<td>None</td>
<td>12.0 ± 0.2 1.4 ± 0.1</td>
<td>3.78 ± 0.03 2.14 ± 0.03</td>
</tr>
<tr>
<td>Cx50p_{141-166}</td>
<td>22.7 ± 0.6 0.8 ± 0.1</td>
<td>3.18 ± 0.03 2.74 ± 0.07</td>
</tr>
<tr>
<td>Cx44p_{129-150}</td>
<td>11.6 ± 0.1 1.5 ± 0.1</td>
<td>0.93 ± 0.02 2.20 ± 0.10</td>
</tr>
<tr>
<td>Cx43p_{136-158}</td>
<td>14.5 ± 0.1 1.6 ± 0.2</td>
<td>1.16 ± 0.02 2.10 ± 0.10</td>
</tr>
</tbody>
</table>

All the experiments were repeated at n= 3.

$K_d$ and the Hill coefficient ($n_{Hill}$) were obtained by fitting the titration curve with Eq.7.

*Phenylalanine fluorescence ($\lambda_{ex} = 250$ nm; $\lambda_{em} = 280$ nm) reports the Ca$^{2+}$ binding to the N-domain of CaM.

#Tyrosine fluorescence ($\lambda_{ex} = 277$ nm; $\lambda_{em} = 320$ nm) reports the Ca$^{2+}$ binding to the C-domain of CaM.
4. EXPLORE MULTIPLE CALMODULIN BINDING SITES IN CONNEXIN43

4.1. Introduction

Gap junctions (GJs) comprise the intercellular channels that mediate the cell-to-cell transfer of small molecules (< 1 kDa) including metabolites, second messengers and ions between neighboring mammalian cells [84]. GJ channels are formed by head-head docking of two hemichannels each comprised of six connexin (Cx) subunits. At least 20 Cx genes have been identified in the human genome with Cx43 being one of the most ubiquitous Cxs [91, 92]. Cx46 and Cx50 are mainly expressed in the lens [168-171]. Based on their sequence similarities, Cx can be further grouped into at least five classes, α (e.g. Cx43, 44, and 50), β (e.g. Cx26, 30, and 32), γ, δ and ε [93] (Table 1.1). All the Cxs share a similar topology with four highly-conserved transmembrane regions, a short N-terminal cytoplasmic region, one intracellular and two extracellular loops, and a C-terminal intracellular tail that exhibits the greatest sequence variation among the Cxs (Fig. 4.1) [85].

It is well known that gap junction proteins are regulated by Ca\(^{2+}\) [87-90]. Modulation of the cytoplasmic Ca\(^{2+}\) concentration is a ubiquitous mechanism by which cells transduce external signals into biological responses. The signaling cascade initiated by the rise in intracellular Ca\(^{2+}\) concentration ([Ca\(^{2+}\)]) is often mediated via Ca\(^{2+}\)-binding proteins such as CaM [172, 173]. Upon Ca\(^{2+}\) binding, CaM undergoes a large conformational change, exposing hydrophobic patches that are important in its binding to more than 500 target proteins in various pathways [174]. CaM uses either its N- or C-domain to differentiate between local and global [Ca\(^{2+}\)], changes and to regulate a diverse group of membrane channels/pumps [175-177] to adjust [Ca\(^{2+}\)]. Depending on the properties of protein targets, CaM uses different binding modes to regulate different biological processes. For example, CaM prefers to interact with 12-30 continuous amino acid with amphipathic helical regions containing hydrophobic residues spaced at 1-5-10, 1-8-14 or 1-16 and interspersed with positively charged residues (≥+3) [62, 178].
While there is increasing evidence that CaM participates in the operation of GJs [100-102], we have shown that the Ca$^{2+}$-dependent regulation of GJs Cxs is effected via the association of Ca$^{2+}$-CaM with the intracellular loop domain of the Cx43, Cx44, the sheep ortholog of rodent Cx46, and Cx50 [103, 104, 179]. Similar to the previously reported CaM-binding sequences [62, 178], we have reported that CaM binds to the second half of the intracellular loop region of Cx43 via a 1-5-10 hydrophobic residue motif [104]. The same binding mode was also reported in other well-characterized CaM targeting proteins such as CaM Kinase I [180], CaM Kinase II [181], MARCKS [141] and synapsin [140]. CaM binds to C-terminal domain of Cx36 in δ-family [182] and binds to N and C-terminal regions of Cx32 in β-family [183-185]. However, the detailed structural mechanism for CaM to regulate Cxs in all families is not clear. To date, multiple regions of CaM interaction for alpha family has not been reported.

In this chapter, we report our identification of two additional potential CaM binding sites on the first half of the intracellular loop (CL) and the C-terminal loop region (CT) in Cx43 (Fig. 4.1). As shown in Fig. 4.1, we define two CaM binding regions in the cytosolic loop as Cx43-2p and Cx43-3p, and define the one at C-terminal as Cx43-4p. The CaM biding domain like Cx43-2p has never been discovered before. Our predicted CaM-binding regions in Cxs encompass the disease-rich cytosolic regions of Cxs that harbor several disease-linked mutations in Cx43 such as oculodentodigital dysplasia (ODDD) [186]. Using high resolution NMR and peptide model, we have shown that CaM interacts with different regions using different binding modes. This study provides important insight into the molecular mechanism of intracellular calcium regulation of Cxs and provide a novo model how CaM involve in gap junction assemble, traffic and calcium dependent regulation.
4.2. Results

4.2.1. Design of peptide fragments to probe interactions of Cx43 by CaM.

The CaM binding sequences of Cx43 in Fig. 4.2A were predicted using the CaM target database [62] with the prediction for each residue scored on a scale from 0 (no binding) to 9 (high affinity binding). The predicted CaM binding region highlighted by the blue line in Fig. 4.2A has been reported in our previous paper [104]. Based on the conserved hydrophobic residue positions and spacing (Fig. 4.1), the key conserved hydrophobic residues in Cx43-2p have a spacing similar to RyR13614-3643 with a binding mode of 1-17 (box 2), while reported Cx43-3p has a 1-5-10 mode of calmodulin kinase II (box 3). However, the binding mode in Cx43-4p region cannot be aligned with peptide sequences from collapsed CaM binding mode subclass (Table 4.1). The predicted CaM binding sequences of Cx43 and Cx32 in the C-tail (box 4) share the most similarity. The CaM binding mode of Cx32 in this region has not been determined.

Two other peptide fragments of Cx43 (green and purple line in Fig. 4.2A) were designed based on the following considerations. First, peptides encompassing the predicted high affinity CaM binding sequences were selected for synthesis. Second, the direct comparison of the CaM binding capability of these sequences allowed us to precisely define the interacting regions. In addition, scrambled peptides with an identical composition of amino acids yet different sequence order and low predictive score for CaM binding were also synthesized as negative controls. Third, all chemically synthesized peptides are longer than 20 amino acid residues based on the common length of CaM binding peptides, and to ensure their ability to adopt secondary structure in solution. Fourth, all of the designed peptides were blocked with an acetyl group at their N-termini and an amide group at their C-termini to mimic their protein environment and remove extra charges [187]. The Cx43 peptide, Cx43-2p, in the first half of the intracellular loop region (Ac-86SVPTLLYLAHVFYVMRKEEKLN107-NH2) and the other in the C terminus region, Cx43-4p, (Ac-224NIIELFYVFFKGVKDRVKGRSDPY247-NH2) was synthesized by EZ Biolab Inc. and purified by preparative reversed-phase high pressure liquid chromatography with purity of >87%.
CaM binding peptides usually form a helical conformation in the hydrophobic environment provided by CaM. Three Cx43 CaM binding domains with highest scores (18 residues) are presented in the helical wheel (Fig. 4.2B). In the helical wheel of Cx43-3p, the hydrophobic residues in 1-5-10 binding mode, M147, L151 and I159 stay close in one side of the helix; while two positive charged residues R148 and K144 were inserted in between each of them. In another side of the helix, the positively charged residues such as H142, R153 and K146 stay together. In the helical wheels of Cx43-2p, hydrophobic residues (F84, M100) in 1-17 CaM binding mode are separated into both ends and hydrophilic residues such as P88 and H95 are close to F84. In the helical wheel of Cx43-4p, the hydrophobic (I225, I226, F229, F232, and F233) and charged residues (D227, K234, D238 and K241) are distributed on opposite ends of the helix.

4.2.2. Monitoring formation of the peptide-CaM complex using Mass spectrometry.

Fig. 4.3AB shows the MALDI-MS spectra of the peptide, Cx43-2p and Cx43-4p (inset) and the mixture of CaM incubated with these two peptides in the presence Ca²⁺. CaM (50 µM) was mixed with 50 µM Cx43 peptides (1:1 ratio) in 50 mM Tris-HCl, 100 mM KCl, and 1 mM CaCl₂ at pH 7.5. The reaction remained at equilibrium for at least half an hour. The data were acquired in a linear positive mode with sinapinic acid (SA) as matrix for CaM and CaM-Cx43-2p/4p complex. A 1 µL mixture was added with 10 µL saturated SA solution and then dried on the MALDI plate for the measurement. Mass over charge (m/z) of protein and peptide in MALDI-MS spectrum usually equal to their mass because they only take 1 charge with proton from matrix. Peptides Cx43-2p/Cx43-4p produced a major peak with the molecular mass of 2693.4/2933.7 that is close to the calculated mass value of 2693.2/2933.4 ([M+H]+). The molecule mass after subtracting the adduct (Na⁺, K⁺) and binding Ca²⁺ from the major peaks at 16.9 kDa is consistent with the calculated mass of CaM, 16.7 kDa. The additional peak with molar mass of ~20 kDa is consistent with the expected value of the CaM-peptide complex. The error occurred in calculation is from the slat adduct to the large molecule such as protein. Under this condition, CaM complexes with Cx43-2p and Cx43-4p in 1 to 1 binding ratio were observed.
4.2.3. Monitoring the interaction between Cx43 peptides and CaM by NMR (1H, 15N)-HSQC spectroscopy.

Using $^{15}$N labeled CaM purified as previously described [188], we have examined the possible interactions of peptides Cx43-2p and Cx43-4p with CaM in the presence of Ca$^{2+}$. In the presence of 5 mM Ca$^{2+}$ (zoom region in Fig. 4.4A), T29, I27, G33, D64 and N137 in CaM exhibit further downfield shift after adding Cx43-2p peptide, while A57 has its proton and nitrogen in further upfield shift. As shown in Fig 4.4B, addition of Cx43-2p resulted in the appearance of a new set of chemical shifts such as G33 while the intensities of the existing set of resonances decreases. The induced chemical shift changes by peptide suggesting it is in the slow exchange regime. Fig. 4.4C displays the chemical shift perturbation in CaM induced by addition of 3-fold molar excess of Cx43-2p. The weight average chemical shift change ($\Delta\delta$) was calculated using Equation 2.5. E helix and H helix of C-lobe of CaM has higher chemical shift than the entire N-lobe of CaM.

Fig. 4.5 shows that the addition of Cx43-4p also results in the change of chemical shifts of the residues of CaM. Although the 3-fold molar excess of Cx43-4p were added into CaM solution, the complex cannot reach to saturation as shown in Fig. 4.5A (spectrum with purple color) due to the limited peptide solubility in the presence of 5 mM Ca$^{2+}$. In this NMR spectrum, there are two sets of peaks correlated to the free form and complex form for most of amino acids in CaM. Interestingly, the chemical shifts of these residues G33 (Fig. 4.5C) and M145 (Fig. 4.5D) were initially small upon the addition of a small amount of peptide Cx-4p with a ratio of CaM to peptide 0-1.0. It becomes significantly greater when peptide concentration is further increased from ratio to 1.5 to 3.0. Both fast (D78 and T146) and slow exchanges (G33 and M145) of chemical shifts were observed in the enlarged spectrum (Fig. 4.5D). Fig. 4.5B depicts the chemical shift perturbation in CaM induced by addition of 3-fold molar excess of Cx43-4p. Although the average chemical shift changes in C-lobe CaM is bigger than N-lobe CaM, B helix and Ca$^{2+}$ binding loop II in N-lobe of CaM has bigger chemical shift change compare to all helixes in C-lobe of CaM.
Fig. 4.6 summarizes the overall chemical shift changes of CaM upon addition of the three peptide fragments from Cx43. Clearly, residues have significant chemical shift changes spread from both N- and C- termini domain of CaM similar to that of Cx43-3p (Fig. 4.6B). Fig. 4.6A also confirms the result of peptide sequence analysis that the different CaM binding modes exist in these CaM-peptide complexes because they exhibit different change processes and chemical exchange behaviors.

### 4.2.4. Determining CaM binding mode of holo-CaM-Cx43-2p complex by residue dipolar coupling.

Residual dipolar coupling (RDC) [189] has been used to assist in characterization of molecular binding processes, determination of relative domain orientations in proteins with multiple domains, and homology modeling [190, 191]. RDC measurements were performed on a Varian 800 MHz spectrometer using a 0.8 mM 15N-labeled holo-CaM in complex with the Cx43-2p (1:2) in 5 mM MES, 10 mM Bis-Tris, 0.1% NaN₃, 5 mM CaCl₂ and 10% D₂O at pH 6.5. First, the ¹H-¹⁵N isotropic couplings of the complex were measured by a TORSY-based J-modulation experiment [192]. The dipolar coupling between nitrogen and protein nuclei in the backbone for each amino acid depends on the distance between them, and the angle of this bond between them is relative to the external magnetic field provided by the NMR instrument. It occurs when the protein molecule in solution exhibit a partial alignment with addition of align medium such as Pf1 phage we used in this study. Upon completion of the experiment, Pf1 phage (70 mg/mL) in the same buffer was added to the complex sample such that the HDO signal was split to a doublet (no window function applied and tn = “lk”) with a coupling of 10 ~ 13Hz. The final concentration of Pf1 phage in the aligned sample is around 12.5 mg/ml. Pf1 phage is negatively charged in buffer conditions and could interact with a positively charged sample. CaM is an acidic protein and should be negatively charged at pH 6.5. However, the iso-electric point of Cx43-2p is 9.74, and it therefore takes positive net charge ~1.1 at pH 6.5. Fortunately, precipitation was not observed when Pf1 phage was added into the sample containing high concentration of Cx43-2p peptide possibly because the positively charged residues form a salt bridge with negatively charged residues of CaM in the binding interface. A second J-modulation experiment with aligned sample was performed. Then RDC values were obtained by
subtracting the isotropic couplings from the values measured in the phage alignment experiment. RDC data were analyzed with CaM complex structures in different binding modes deposited in the protein data bank (PDB). The missing proton information in the X-ray crystal structure was generated by SYBYL 6.7 (Tripos Inc.). The calculated RDC values based on known structures were fit to the experimentally measured couplings using the program REDCAT [193].

Recently, RDC has also been applied to investigate the CaM and Cx43 binding mode with limited structure information [68, 191, 194, 195], using the measured RDCs (CaM-Cx43-2p) with calculated values in the PDB (Table 4.2). Both collapsed (PDB codes 2HQW, 1CDM, 1CDL, 1SY9, 1NIW, 1IQ5, and 2BCX) and extended (PDB codes 1CFF, 1G4Y, 1K93, 2IX7 and 1NWD) CaM binding mode have been examined. In collapsed CaM binding modes (1-7, 1-10, 1-14, 1-16 and 1-17), both N lobe and C lobe of CaM embrace the peptides in CaM binding domains. More than one peptide can bind CaM in its extended binding mode (PDB codes 1G4Y, 1K93 and 1NWD). We also compared the data with the structure of holo-CaM (PDB code 3CLN). The quality of the fit between calculated RDC and experimental measured RDC is determined by the Q-factor and the agreement between domain orientations. As listed in Table 4.2, the best correlation of measured and calculated RDCs (lowest Q-factor, 0.149) were calculated when the RDCs were fitted using the structure of gating domain from small conductance K+ channel (SK channel, PDB codes, 1G4Y) [83]. Schumacher et al reported that the CaM binding domain in SK channel forms a dimer in extended conformation with a CaM molecule (Fig 4.7A inset). As shown in Fig 4.7 and Table 4.2, 33 RDCs were used in the CaM binding mode fitting. The residues used for 33 RDCs fitting are almost evenly distributed with the exception of the linker region and helix E (Fig 4.7B).

4.2.5. Determining the binding affinity of Cx43-2p/4p-CaM complex by modified CaM with fluorescence dye.

Dansyl-labeled CaM (D-CaM) and IAEDANS-labeled CaM (I-CaM) have been frequently used to determine the binding affinity of CaM-peptide interaction due to its sensitivity to the changes in the
surrounding chemical environment and its ease of preparation [108, 151, 196]. The dansyl moiety has an emission maximum at ~500-510 nm, a range in which the signals from intrinsic aromatic residues are negligible. However, as we report before, no significant fluorescence changes were observed for the titration of D-CaM with Cx43-2p and Cx43-4p peptides [104]. Another fluorescence dye, IAEDANS, was introduced to instate Cys at C-terminal of CaM to detect the CaM conformation change after binding with these two peptides. As seen in Fig 4.9, I-CaM (2 µM with 1 mM Ca\(^{2+}\)) showed a fluorescence maximum at 484 nm. With the addition of Cx43-4p (Fig. 4.8A), the IAEDANS fluorescence emission maximum blue-shifted to 473 nm with a concomitant enhancement of its fluorescence intensity (Fig. 4.8A) implying that the IAEDANS group entered a more hydrophobic environment upon complex formation with peptide. However, no significant fluorescent changes were seen upon the addition of the Cx43-2p peptide to I-CaM probably because the Cx43-2p peptide binds to N-lobe of CaM which is far away to the dye conjugated in the C-lobe of CaM. By fitting the titration curve with hill equation (modified Equation 2.6), we obtained a dissociation constant \((K_d)\) 23.3 µM for the interaction between Ca\(^{2+}\)-CaM and Cx43-4p (Fig. 4.8B and Table 4.3). This is based on the assumption of that free peptide concentration is equal to total peptide concentration when it is much higher than I-CaM (2 µM). The titration curve and hill number (2) suggests the binding stoichiometry is more than 1 peptide per CaM molecule.

We also engineered and purified the C lobe of CaM with dansyl-labeled (D-C-CaM) to capture the conformational change of CaM upon addition of Cx43-2p peptide. We first examined the dansyl fluorescence anisotropy changes to confirm the formation of complex in the low micromolar range. Fluorescence anisotropy is dependent on the rotational correlation time of fluorophore in the sample, and is often used to reflect the hydrodynamic properties of biomolecules [152]. The increase in anisotropy often arises from slower tumbling of the fluorophore and thus reports the formation of a larger complex. Indeed, upon addition of 100 µM of Cx43-2p into 1 µM D-C-CaM, the anisotropy of the dansyl moiety within D-C-CaM increased from 0.081 to 0.296, suggesting the association of Cx43-2p with Ca\(^{2+}\)-CaM (Fig. 4.9B inset). We then carried out the titration of peptide with D-C-CaM by monitoring the
fluorescence emission between 400 nm and 600 nm. In the presence of Ca\(^{2+}\), D-C-CaM showed a fluorescence maximum at 508 nm. With the addition of Cx43-2p (Fig. 4.9A inset), the dansyl fluorescence emission maximum blue-shifted to 490 nm with a concomitant enhancement of its fluorescence intensity (Fig. 4.9A) implying that the dansyl group entered a more hydrophobic environment upon complex formation with peptide. By fitting the titration curve with a 1:1 binding mode (Equation 2.6), we obtained a dissociation constant \((K_d)\) of 13 ± 4 µM (fluorescence anisotropy) and 77 ± 5 µM (static fluorescence) for the interaction between Ca\(^{2+}\)-CaM and Cx43-2p (Table 4.3).

4.3. Discussion

4.3.1. The binding stoichiometry between CaM and Cx43 peptides

Since the gap junction channel is a hexamer embedded into the membrane, to further understand the regulation of gap junction by CaM, it is important to clarify the stoichiometry between CaM and these three Cx43 peptides. As we reported previously [104], both NMR and fluorescence titration experiment suggested that Cx43-3p peptide binds with CaM in a molar ratio of 1 to 1. In the HSQC-NMR spectrum, peaks of G25, G33, A57, G61, D64, K94 and K148 in CaM were selected to indicate the global conformational change of CaM upon addition of Cx43-3p peptide from 0 to 2 in a molar ratio. The chemical shifts of these residues are clearly observed to stop moving after the peptide to CaM ratio reaches to 1 to 1. In this study, however, it seems that the binding stoichiometry between Cx43-2p or Cx43-4p peptide and CaM is greater than one. As shown in Fig. 4.4B, the bound form of G33 in CaM did not appear until adding more than 1 molar ratio peptide and the free form of G33 in CaM disappears only when Cx43-2p peptide to CaM ratio reaches to 3 to 1. Similarly, as shown in Fig. 4.5C, between the mix ratios from 0 to 1-fold molar excess of Cx43-4p, G33 in CaM went through fast exchange when interaction occurred. After adding Cx43-4p peptide to CaM solution at increasing ratios from 1.5 to 3, the decrease of the peak intensity of G33 was accompanied by an increase in the corresponding peak intensity (bound form) at a new position (N, 8.95 ppm and H, 107.0 ppm). The concentration of the peptides was calculated by weight, purity, dissolving volume and molecule weight and it was confirmed by UV
absorbance of tyrosine (two Tyr for each peptide). The NMR experiments for both peptides interacting with CaM have been repeated in both 600 MHz and 500 MHz NMR in our lab. In addition, when the RDC data was collected in 800 MHz at University of Georgia, the CaM-Cx43-2p complex HSQC spectrum could not reach saturation at a ratio of two peptides per CaM because peaks in two components were observed. We conclude that more than one Cx43-2p or Cx43-4p peptide binds with CaM based on the evidence from NMR experiment.

As shown in Fig. 4.9B, the binding curve for I-CaM titrated with Cx43-4p peptide only can be fitted by the Hill equation. The Hill number is 1.99, which indicates more than 1 Cx43-4p peptide binding with CaM and the cooperative binding between them. However, I-CaM did not show any fluorescence change upon addition of Cx43-2p peptide. In MALDI-MS spectrum, we also observed the complex molecule with two peptides and one CaM for both peptides (in the higher range of m/z and data not shown in Fig. 4.2). It could be matrix artificial effect (Sample co-crystallization with matrix and matrix may facilitate the complex formation) and ESI-MS monitoring the interaction in solution interface will be used as another control to demonstrate the complexes of multiple peptides with CaM in solution phase.

It is also important to clarify the differences of binding affinities between these three peptides and CaM. As shown in Table 4.3, the dissociation constants determined by fluorescence are Cx43-3p > Cx43-4p > Cx43-2p. However, in the HSQC-NMR spectrum (Fig. 4.4 & 4.5), the binding affinities between CaM and all the peptides are in the different manner because the chemical exchange rates from fast to slow are Cx43-3p > Cx43-4p > Cx43-2p. The possible reasons for the inconsistent results are, 1) the 1 to 1 binding equation used in the fluorescence experiments may not be optimal because more than 2 to 1 stoichiometry was observed in NMR experiment; 2) upon addition of Cx43-2p, the fluorescence of both dansyl and IAEDANS group in CaM does not change and dissociation constant is only available from dansyl labeled half-CaM which is not equal to the entire CaM molecule. Further study of the binding affinity measurement requires additional experiments such as surface plasmon resonance (SPR) which could provide both dissociation constants and kinetic information with more accuracy. Fluorescence and
anisotropy fluorescence data, somehow, at least indicate that the binding between them is confirmed and CaM should interact with all peptides in different regions.

**4.3.2. The CaM binding modes between CaM and Cx43 peptides.**

CaM was reported to bind with more than 150 target proteins in different binding modes [62]. More and more proteins, especially channel proteins and enzymes like protein kinases, have been reported to bind CaM with more than one binding sites. The multiple roles of CaM in regulation of certain proteins require multiple interaction domains. In 1988, Ladant identified two separated CaM binding domains (CaMBD) in *Bordetella pertussis* adenylate cyclase [197]. Two years later, Trewhella *et al.* studied the solution structure of CaM complexes with two peptides, PhK5 (342-366) and PhK13 (301-326) from the catalytic subunit of phosphorylase kinase [198]. They discovered that binding of PhK5 to CaM induces a compact formation of CaM while CaM remains extended upon binding of PhK13. But in the presence of both peptides (PhK5 and PhK13), CaM still keeps the extended conformation [198]. The other enzymes such as adenylyl cyclase type I and II, synapsin I, plant glutamate decarboxylase and caldesmon have also been found to have multiple regions that bind CaM [199-202]. Interaction of CaM with multiple binding sites of the olfactory and rod cyclic nucleotide-gated channels, the plasma membrane Ca$^{2+}$ pump, TRP channels, Cav1.2 Ca$^{2+}$ channels and ryanodine receptor has also been reported [77, 203-206].

According to the structure of CaM in the binding complex, CaM binding modes can be classified as two major groups, collapsed and extended conformation. In our previous report, Cx43-3p in cytosolic loop region of Cx43 has been suggested to bind with CaM in the 1-10 binding mode, in which CaM is in a collapsed conformation and embraces the peptide in its central loop region [104]. The 1-10 motif refers to a group of sequences whose key hydrophobic residues are spaced 9 residues apart as it was first reported for the CaM binding domain in CaM dependent kinase I and II [68, 207].

Based on extensive studies using various approaches such as proteinase digestion [208, 209], several CaM binding regions within the RyR1 primary sequence such as RyR11975-1999 and RyR13614-3643, have been identified by Hamilton and co-workers [210, 211]. As shown in **Fig. 4.1** and **Table 4.1**, the
sequence of Cx43-2p fits very well with RyR1_{3614-3643} as the 1-17 CaM binding mode. They both contain the key hydrophobic residues at positions 1 and 17 in the binding interface with CaM and a positively charged arginine right after position 17. As shown in Table 4.2, calculated RDC values for the CaM-Cx43-2p complex are similar to the CaM-RYR1 complex (PDB code: 2BCX) in a 1-17 CaM binding mode (31 residues with $Q = 0.198$). However, the best aligned structure with calculated RDC values is CaM with peptide fragments from the gating domain of the potassium channel (SK channels) (33 residues with $Q = 0.149$). As shown in Fig. 4.7B, 33 residues with RDC values in binding mode determination are almost evenly distributed in CaM with the exception of the linker region and helix E. Unlike gap junction proteins, small-conductance Ca$^{2+}$-activated K$^+$ channels (SK channels) are independent of voltage and gated only by intracellular Ca$^{2+}$ [212, 213]. Schumacher et al. solved a crystal structure of CaM complex which provides an evidence of both Ca$^{2+}$-dependent and -independent CaM interaction with SK channel CaMBD [83]. It this structure, a pocket (pocket 1) in N-lobe and 3 pockets in C-lobe of holo-CaM have been identified to interact with the SK channel CaMBD. A complementary hydrophobic binding pocket in Ca$^{2+}$-CaM N-lobe is created by Phe 12, Phe 19, Val 35, Met 36, Leu 39, Phe 68, Met 71 and Met 72. In CaM C-lobe, three additional pockets, pocket 2 with Met 124, Phe 141 and Met 144, pocket 3 with Phe 89, Phe 92, Leu 105 and Met 109, and pocket 4 with Val 91 and Phe 92, has also been located in this paper [83]. Table 4.4 summarizes all chemical shift changes between the Ca$^{2+}$-CaM-Cx43-2p Complex and Ca$^{2+}$-CaM in HSQC-NMR spectrum. It is clear that all the key residues in the SK channel CaMBD binding pockets have larger chemical shift changes in our complex. The changes for most of them are greater than 0.1 ppm, which is large enough to be considered as a possible binding interface. However, the incomplete assignment limits the agreement such as the unavailable Val 91 and Phe 92 in pocket 4. According to the RDC data and the comparison of chemical shift changes, we can conclude that the CaM binding mode of Ca$^{2+}$-CaM-Cx43-2p Complex is similar as the manner in SK channels. The interaction is involving multiple CaM binding regions and CaM is in an extended mode which is quite different as we proposed for Cx43-3p previously [103, 104, 179].
4.3.3. Comparison of CaM binding domains with α-, β-, γ- and δ-family.

CaM binding domains (CaMBD) of the connexin proteins in different gap junction subfamilies have been reported previously. The vertebrate connexin family has been summarized and defined as four major groups, which are group I (β-family), group II (α-family), group IIIa (δ-family) and group IIIb (γ-family) [94, 214]. Cx56 in α-family, the primary gap junction protein in chicken lens, is the first CaM binding connexin protein in a Ca^{2+}-dependent manner identified in 1982 [215]. In mammal lens fiber cells, inhibition of cell-to-cell communication by elevated [Ca^{2+}], was first demonstrated as an increased internal electrical resistance [216] that was prevented by pre-incubation with CaM antagonists[217-219]. The rapid onset of this inhibition (within seconds) suggests that this is mediated by a direct interaction of CaM with one or more of the Cxs. Our lab has reported that CaM interact with the second half of the intracellular loop of three major Cxs in α-family, Cx43, Cx44 (sheep homologue of human Cx46) and Cx50 in lens fiber cells [103, 104, 179]. The CaM binding modes for these regions has been suggested to be like CaM-CaMKII having the 1-5-10 subclass [103, 104, 179]. Our preliminary data also indicates that the same region of cytosolic loop of the γ-family Cx45 also interacts with CaM and the proposed binding mode is similar to the α-family.

CaM has also been found to co-localize with Cx32 in the β-family [105] and directly gates Cx32-containing GJs [106]. Török et al. [185] reported that fluorescently labeled CaM derivatives bind to synthetic peptides covering most of the cytoplasmic sequences of Cx32 and they identified two regions, one at the N-terminus, and another at the C-terminus, showing Ca^{2+}-independent CaM-binding properties [185]. Dodd et al. also identified two distinct CaM-binding amino acid sequences in Cx32 within the N- and C-lobes of CaM showing separate functions, suggesting trans-domain or trans-subunit bridging by CaM as a possible mechanism of GJ gating [184]. Using in vitro synthesis approaches, we have shown that oligomerization of Cx32 is CaM-dependent [161] suggesting CaM interacts with Cxs at an early stage of GJ assembly [100, 185]. Blodow et.al reported that CaM antagonists suppress GJ coupling of Cx26 in isolated Hensen cells of the guinea pig cochlea [162]. However, there is still no clear evidence for the direct interaction of CaM with Cx26. Interestingly, the same CaMBD of Cx34.7, Cx35 and Cx36
in \( \delta \)-family also has been reported [182]. The CaM interaction in these three \( \delta \)-family gap junction proteins is in concentration and \( \text{Ca}^{2+} \)-dependent manner with rapid kinetics. Two possible binding sites, with a high affinity (\( \sim 72 \text{ nM} \)) and a low affinity (\( \sim 2.4 \mu \text{M} \)), have been identified in SPR measurement with two-component model. The weaker dissociation constants of CaM with Cx43-2p or Cx43-4p by fluorescence could be due to the same reason. The average dissociation constants we observed in these experiments may contain both weaker and strong binding processes. The possibility of CaM binding with intracellular loops of these Cxs has been eliminated in this paper as well [182]. The sequence of binding region at C-terminal of Cx43 (Cx43-4p) can be aligned well with the same region in Cx32 (\( \beta \)-family) and Cx36 (\( \delta \)-family) (Fig. 4.1). However, CaM binding modes for both \( \beta \)-family and \( \delta \)-family have not been reported previously. The RDC experiment for determining the CaM binding mode of \( \text{Ca}^{2+} \)-CaM-Cx43-4p is under preparation. Cx43, the most ubiquitous connexin in mammal, should contain the unique CaM binding properties crossing all the gap junction families.

4.3.4. Disease implications in CaM binding domains.

Three CaM binding regions in Cx43 are highly conserved in the alpha family. Germ line mutations in nearly 50% of the genes that encode 21-member Cxs are linked to one or more human diseases including deafness, Oculodentodigital dysplasia (ODDD), cataract and Charcot-Marie-Tooth disease [186, 220]. Congenital cataracts are attributed to mutations in the genes encoding Cx46 and Cx50 [221, 222]. While some of these disease causing mutations are autosomal recessive, some are autosomal dominant [223]. For example, autosomal dominant mutations in GJA1 genes encoded Cx43 cause the mild developmental disorder, oculodentodigital dysplasia (ODDD). While it is difficult to establish how these genotypic changes lead to disease outcomes, given the broad and overlapping distributions of Cxs in a wide range of tissues, we have noticed that quite a few disease-linked mutations reside within the intracellular loop regions, which we have shown are important for the regulation of these Cxs. Table 4.5 summarizes all the ODDD disease mutations which overlap with the CaM binding regions we reported here and previously.
It implicates that CaM binding regions in Cx43 are the potential therapeutic target which could prevent ODDD diseases.

4.3.5. The role of CaM in regulation of Cx43.

CaM accessibility within the assembled channel is a major consideration because all the predicted and identified CaMBD in Cx43 involves transmembrane domains (TM). The Cx43-2 site is predominantly located within the TM2. Pro 88 (Pro 87 in β-family) is highly conserved and critical for gating [224, 225]. This locus is also in the middle of the TM2 domain that probably ends within the FYV sequence of this predicted sequence (the RKEE is certainly cytoplasmic). The Cx43-4 site probably also contains ~34.8% transmembrane region since the TM4 domain likely ends with the FYVFF sequence. Accessibility to cytosolic CaM is a major question with this site within the assembled channel. However, the reported Cx36 C-terminal CaMBD also involves TM4 (~ 34.8%). In addition, CaMBD located at TM in the pore lining surface may be exposed to the solvent and become accessible to CaM. Even the pore is open widely enough to provide the binding surface for CaM since we propose the CaM binds with Cx43-2p in an extended mode which only require N-lobe or C-lobe of CaM (Fig. 4.7).

Protein-protein interaction plays an essential role in channel localization and activity and gap junction protein are involved in different cellular functions by interacting with many other proteins. For example, the C-terminal tail of Cx43 has been reported to bind with SH2 and SH3 domains of sarcoma, PDZ domain of zonula occludes-1, N-cadherin, protein kinase C, CaM, CaM dependent protein kinase II and microtubules [99, 226]. Identified microtubule binding region at the C-terminal of Cx43 overlaps with our reported CaM binding region here (Cx43-4p) [227]. Shaw et al. reported that delivering Cx43 hemi gap junction channel into plasma membrane requires microtubule proteins [228]. However, both groups did not mention the role of CaM and the possibility of CaM being involving in this cellular event cannot be eliminated. The CaM and microtubule binding domains are localized to overlapping regions at N-terminus of p35 (the neuronal activator of Cdk5) [229]. Ca²⁺-CaM inhibits p35 association with microtubules, suggesting p35 may be involved in the Ca²⁺-CaM mediated inhibition of microtubule
assembly [229]. CaM has also been also reported to bind with microtubule-associate proteins 2 and 6 which facilitate microtubule assembly and stabilize tubulin structure [230, 231]. Passareiro et al. reported that two structure domains of calmodulin interact with mature and immature rat brain microtubules [232]. Ahmad et al. also suggested that CaMBD in the C-terminal tail of Cx32 likely regulates connexin oligomerization [233]. Therefore, our hypothesis is that CaM involved in the hemi-gap junctional trafficking into plasma membrane through microtubules and CaMBD at C-terminus of Cx43. In the proposed model (Fig. 4.11), CaM binding to Cx43-4p region induces connexon hexamer assembly of the hemi gap junction channel in Golgi apparatus. With help from CaM at Cx43-4p region, gap junction hemi-channels traffic through microtubule proteins into plasma membrane. After formation of functional channels with adjacent cell, CaM sticks at the Cx43-2p region in the cytosolic calcium concentration. CaM starts to bind to the additional region Cx43-3p when local Ca\(^{2+}\) concentration increases. CaM binding to Cx43-3p region in the presence of Ca\(^{2+}\) regulates the closure of the channel. The phosphorylation of Cx43 in the C-terminal tail may trigger its internalization and degradation. CaM binding to Cx43-4p region may protect gap junction from CaM kinase II (CaMKII) phosphorylation which also function as a hexamer. There is no clear evidence to support that CaM involve in gap junction protein degradation. Further investigation into the function of CaM induced gap junction trafficking is required.

### 4.4. Future plan.

a. Peptide competition experiment in CaM binding sites of Cx43 by HSQC NMR.

b. CaM binding modes determination of CaM-Cx43-3p and CaM-Cx43-4p complex by RDC NMR measurement.

c. Binding affinities measurement and stoichiometry between CaM and three Cx43 peptides by SPR and ESI-MS.

d. Proposed functional studies (Dr. Richard Veenstra)
4.5. **Summary and conclusion**

Gap junction, allowing the intercellular transmission of molecules through its specialized cell membrane channels, plays a major role in intercellular calcium signaling between adjacent cells. Connexin43 (Cx43), the most ubiquitous connexin, belongs to α family of gap junction proteins expressed in heart where are essential for normal heart development. Calmodulin (CaM) has been implicated in mediating the Ca$^{2+}$-dependent regulation of gap junctions. We have reported CaM binding site in the second half of intracellular loop of Cx43. In this study, two additional CaM binding regions in cytosol loop and C-termini of Cx43 have been identified by biophysical studies. Our results indicate that in the presence of Ca$^{2+}$, synthesized Cx peptide fragments, Cx43-2p and Cx43-4p, encompassing predicted CaM binding regions are able to bind with high affinity to CaM using NMR spectroscopy. Unlike Cx43-3p, the suggested stoichiometry of these new discovered peptides binding with CaM is greater than one. The NMR RDC result suggests that the CaM binding mode of Cx43-2p to CaM is similar as CaMBD in SK channel in an extended mode. Our results elucidate the molecular level of regulation of Cx43 by multiple CaM targeting regions and provide a proposed model about how CaM regulate gap junction assemble, trafficking and gating.
The integral α-class of connexins are composed of four transmembrane (TM) segments, two extracellular loops, one cytoplasmic loop (separated as the green box 2 and the blue box 3), a short N-terminus (the black box 1) and a much longer C terminal tail (the purple box 4). Three potential CaM binding sites of Cx43 illustrated in the intracellular region of connexins from 1 to 4 are listed in the table box colored with corresponding regions. Two of them named as Cx43-2p and Cx43-3p are located in the intracellular loop between TM2 and TM3; Cx43-4p is located in the C-terminal tail. The numeric score (1-9) represents the probability of an accurate prediction of high affinity CaM binding sites. The sequence of Cx43-2p is aligned very well with peptide (3619-3643) from Ryanodine receptor 1 (RyR1) which fits 1-17 CaM binding mode subclass. The aligned sequence of Cx43-3p fit the 1-5-10 CaM-binding mode subclass. The identified CaM-binding sequences from the α-class connexins in the second half of intracellular loop (box 3) are similar to the Ca2+-CaM-dependent kinase II (CaMKII). Each number represents the presence of a hydrophobic residue in blue color. The predicted CaM Binding sequences of Cx43 and Cx32 in the C-tail (box 4) share the most similarity. The CaM binding mode of Cx32 in this region has not been determined. h, human; m, mouse; r, rat; s, sheep; B, basic residues (red); A, acidic residues; #, hydrophobic residues (blue); ?, the potential hydrophobic residue which could determine the binding modes.

Figure 4.1 Connexins membrane topology putative CaM-binding sites of Cx43.
Figure 4.2 The predicted CaM binding sequences in Cx43.

A, the primary sequence of Cx43 (NP_000156.1). The predicted high affinity CaM binding site (136-158 highlighted above green line, box 3 in Fig.1) of highest predicted score (13) is located in the second half of the intracellular loop between TM2 and TM3. Two other regions (86-107 highlighted above blue line, box 2 in Fig.1; 224-247 highlighted above purple line, box 4 in Fig.1) with lower scores are also predicted. B, helical wheel displays of the CaM binding sites, Cx43-2p, Cx43-3p and Cx43-4p (Helical Wheel Projections created by Don Armstrong and Raphael Zidovetzki http://rzlab.ucr.edu/scripts/wheel/wheel.cgi). Circles, hydrophilic residues; diamonds, hydrophobic residues; triangles, negative charged residues; pentagons, positive charged residues. Hydrophobicity is decreasing proportionally to the amount of green color. The most hydrophobic residue is green and zero hydrophobicity coded as yellow. Hydrophobicity is based on the whole residue interface scale. Hydrophilic residues are coded red and the residues with charge are light blue. 18 residues of each CaM binding domain with highest scores are presented in the helical wheel.
Figure 4. Monitoring CaM-Cx43-2p and CaM-Cx43-4p complex formation by MALDI mass spectrometry.

The MALDI-MS spectra of the free form of CaM (~17 kDa) and the complex form of CaM-Cx43-2p (A) and CaM-Cx43-4p (B) (~20 kDa). The inset shows the molecular weight of the Cx43-2p (A) and Cx43-4p (B) peptides. 50 µM CaM was mixed with 50 µM Cx43 peptides (1:1) in 50 mM Tris-HCl, 100 mM KCl, 1 mM CaCl₂ at pH 7.5. The reaction was equilibrium for at least 30 minutes. The data were acquired in a linear positive mode with sinapinic acid (SA) as matrix for CaM and CaM-Cx43-2p/4p complex. A 1 µL mixture was added with 10 µL saturated SA solution and then dried on the MALDI plate for the measurement. The molecular weight of the Cx43 peptides was confirmed by MALDI in reflectron positive mode with α-Cyano-4-hydroxycinnamic acid (CHCA) as matrix.
Figure 4. Monitoring the interaction between holo-CaM and Cx43-2p peptide by (1H, 15N)-HSQC spectroscopy.

A, an overlay of HSQC spectra of holo-CaM (red) with the spectrum of the holo-CaM-Cx43-2p (blue). The Cx43-2p peptide is from the first half of intracellular loop (highlighted in red color in the top cartoon) in Cx43. As indicated in the zoom in graphs, T29, I27, G33, D64 and N137 in CaM exhibit further downfield shift after adding Cx43-2p peptide; while A57 has its proton and nitrogen in further upfield shift. B, the chemical shift change of Gly33 during titration of holo-CaM with Cx43-2p. The disappearance of the peak (free form) was accompanied by the appearance of the corresponding peak (bound form) at a new position. C, chemical shift perturbation in CaM induced by addition of 3-fold molar excess of Cx43-2p. The weight average chemical shift change (Δδ) was calculated using Equation 5. E helix and H helix of C-lobe CaM has higher chemical shift compare to N-lobe CaM.
Figure 4. Monitoring the interaction between holo-CaM and Cx43-4p peptide by ($^1$H, $^{15}$N)-HSQC spectroscopy.

A, an overlay of HSQC spectra of holo-CaM (red) with the spectrum of the holo-CaM-Cx43-4p (purple). The Cx43-4p peptide is from the C-terminal tail (highlighted in red color in the inset) in Cx43. B, chemical shift perturbation in CaM induced by addition of 3-fold molar excess of Cx43-4p. The weight average chemical shift change ($\Delta\delta$) was calculated using Equation 2.5. Although the average chemical shift changes in C-lobe CaM is bigger than N-lobe CaM, B helix and Ca$^{2+}$ binding loop II in N-lobe CaM has bigger chemical shift change compare to all helixes in C-lobe CaM. C, the chemical shift change of Gly33 during titration of holo-CaM with Cx43-4p. Between the mix ratios from 0 to 1-fold molar excess of Cx43-4p, G33 went through fast exchange when interaction occurred. After adding Cx43-4p peptide from to 1.5 to 1 ratio, the decrease of the peak intensity of G33 was accompanied by the increase of the corresponding peak intensity (bound form) at a new position. D, D78 and T146 in CaM-Cx43-4p interaction behaves as fast exchange; while M145 has similar binding pattern as G33.
Figure 4.6 Comparison of the three CaM binding sites of connexin43 in \((^{1}H, ^{15}N)-HSQC\) spectrum.

An overlay of the HSQC spectra of holo-CaM-Cx43-2p complex (cyan), holo-CaM-Cx43-3p complex (yellow), and holo-CaM-Cx43-4p complex (violet). Due to dramatic differences in these three spectrum, they are not the same CaM binding mode. E6 residue was selected to align all three spectrum because it did not move during the peptide titration.
Figure 4.7 Comparison of the three CaM binding sites of connexin43 in chemical shift.

An overlay of chemical shift perturbation in CaM induced by addition of Cx43-2p (cyan), Cx43-3p (yellow), and Cx43-4p (violet).
Figure 4. 8 CaM binding mode determination of holo-CaM-Cx43-2p complex by residue dipolar coupling (RDC).

A, correlations between experimentally determined $^1$H-$^{15}$N RDCs of the holo-CaM-Cx43-2p complex and the best fit RDC values calculated using the structure of holo-CaM-K$^+$ channel complex (pdb code: 1G4Y, the inset is the cartoon picture of its structure). CaM (purple) binds to the Ca$^{2+}$ activated K$^+$ channel (blue) with a 2 to 2 binding ratio in an extended mode. The quality factor (Q) of the fitting is 0.149. B, the distribution of CaM residues used in structure comparison. Top panel is CaM (14GY) cartoon structure drawing with PyMOL (red, residues in CaM used in RDC value calculation; grey, CaM; green, Ca$^{2+}$). 33 residues (red dots) from CaM were used in binding mode determination. (bottom panel). They are almost evenly distributed with exception of the linker region and helix E.
Figure 4.9 The interaction of Cx43-2p and Cx43-4p peptide with IAEDANS-CaM (I-CaM) monitored by steady-state fluorescence.

A, the fluorescence spectrum of I-CaM (2 µM) in the absence of (dash lined with hollow dots) and in the presence of (solid line with solid dots) Cx43-4p in a buffer consisting of 1 mM Ca\(^{2+}\), 100 mM KCl, 50 mM Tris-HCl, 10 mM DTT at pH 7.4. The inset shows the fluorescence spectrum of I-CaM (2 µM) with and without Cx43-2p in the same condition. B, the titration curve of I-CaM (2 µM) with Cx43-4p in the presence of 1mM Ca\(^{2+}\).
Figure 4. 10 The interaction of Cx43-2p peptide with Dansyl labeled C lobe of CaM (D-C-CaM) monitored by steady-state fluorescence (A) and anisotropy fluorescence (B).

A, the fluorescence titration curve of D-C-CaM (1 µM) in the presence of 1mM Ca^{2+} with Cx43-2p in a buffer consisting of 1 mM Ca^{2+}, 100 mM KCl, 50 mM at pH 7.4. The inset shows the fluorescence titration spectrum before (broken line with hollow dots) and after (continuous line with solid dots) adding Cx43-2p peptide into D-C-CaM. B, the anisotropy fluorescence titration curve of D-C-CaM (1 µM) with Cx43-2p in the same condition. The inset shows the anisotropy change for D-C-CaM (free, light gray bar), and the complex of D-C-CaM-Cx43-2p (bound, dark gray bar).
In the proposed model, Cx43 was synthesized at endoplasmic reticulum membrane (ER, double circle in blue color, yellow color inside indicates the high calcium concentration at $10^{-4}$ M). CaM binding to Cx43-4p (black ellipsoid) region induce the hemi-channel formation of gap junction hexamer (grey cylinder and circle) at Golgi apparatus (GA) With help of CaM in Cx43-4p region, gap junction hemi-channels traffic through microtubule proteins (green bar) into plasma membrane (PM). Purple dots stand for lipids and light yellow color indicates the increase of local Ca$^{2+}$ concentration. After formation of functional channels with adjacent cell, CaM sticks at Cx43-2p region (green ellipsoid) in the cytosolic calcium concentration. CaM starts to bind to the additional region Cx43-3p (red ellipsoid) when local Ca$^{2+}$ concentration increase. CaM binding to Cx43-3p region in the presence of Ca$^{2+}$ regulates the closure of the channel. The phosphorylation of Cx43 in the C-terminal tail may trigger its internalization and degradation. CaM binding to Cx43-4p region (black ellipsoid) may protect gap junction from CaM kinase II phosphorylation as well. However, there is no clear evidence to support that CaM involve in gap junction protein degradation. ER, endoplasmic reticulum membrane; GA, Golgi apparatus; PM, plasma membrane.
<table>
<thead>
<tr>
<th>Binding Mode</th>
<th>Cx43-2p</th>
<th>Cx43-3p</th>
<th>Cx43-4p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 – 7 NMDA receptor NR1 C1</strong></td>
<td>FSVPTLLYLAHV-VMRKEKLNI 24 - KKKATFRAI TSTLASSFKRRSSK 24</td>
<td>FKYG EEHGVKQRGGLRTYI IS 23 KKKATFRAI TSTLASSFKRRSSK 23</td>
<td>N I I ELFYVFFKGVKDRVKGRSDHY 23 KKKATFRAI TSTLASSFKRRSSK 23</td>
</tr>
<tr>
<td><strong>1 – 10 CaMKIIα</strong></td>
<td>FSVPTLLYLAHV-VMRKEKLNIK 24 LKKFNARRKLGII LTMLATRNF 24</td>
<td>FKYG EEHGVKQRGGLRTYI IS- I S 26 QQRG- GFRI ARVLVMREWYHNFR 26</td>
<td>LNI I ELFYVFFKGVKDRVKGRSDHY 25 LKKFNARRKLGII LTMLATRNF 25</td>
</tr>
<tr>
<td><strong>1 – 14 CNG channel</strong></td>
<td>FSVPTLLYLAHV-VMRKEKLNIK 24 - QQRGGFRRI ARVLVMREWYHNFR 26</td>
<td>KFKYG EEHGVKQRGGLRTYI IS- I S 26 QQRG- GFRI ARVLVMREWYHNFR 26</td>
<td>LNI I ELFYVFFKGVKDRVKGRSDHY 25 QQRGGFRRI ARVLVMREWYHNFR 25</td>
</tr>
<tr>
<td><strong>1 – 16 CaMKK</strong></td>
<td>FSVPTLLYLAHV-VMRKEKLNIK 27 RFPNFRKRRHAKVL- ILTDLRPI VRV 27</td>
<td>KFKYG EEHGVKQRGGLRTYI IS- I S 27 RFPNFRKRRHAKVL- ILTDLRPI VRV 27</td>
<td>LNI I ELFYVFFKGVKDRVKGRSDHYA 27 RFPNFRKRRHAKVL- ILTDLRPI VRV 27</td>
</tr>
<tr>
<td><strong>1 – 17 RyR1</strong></td>
<td>FSVPTLLYLAHV-VMRKEKLNIK 19 VHKLSSQRRRAVACF RMPFLYN 19</td>
<td>KFKYG EEHGVKQRGGLRTYI IS- LFK 29 KSKKAWHKLSSQRRRAVACF RMPFLYN 29</td>
<td>LALNI I ELFYVFFKGVKDR-VK - RSDPYHAT 29 KSKKAWHKLSSQRRRAVACF RMPFLYN 29</td>
</tr>
</tbody>
</table>
Table 4. 2 CaM binding modes and Q factors of the fitting of experimentally measured RDCs of the Ca\(^{2+}\)-CaM-Cx43-2p complex.

<table>
<thead>
<tr>
<th>Pdb</th>
<th>Target protein</th>
<th>Binding Mode</th>
<th>Q-factor</th>
<th>RMSD</th>
<th>No. of RDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2HQW</td>
<td>Glutamate NMDA receptor subunit zeta 1</td>
<td>1-7</td>
<td>0.184</td>
<td>2.07</td>
<td>34</td>
</tr>
<tr>
<td>1CDM</td>
<td>CaMKII</td>
<td>1-10</td>
<td>0.175</td>
<td>2.11</td>
<td>40</td>
</tr>
<tr>
<td>1CDL</td>
<td>CaMKII alpha chain</td>
<td>1-10</td>
<td>0.183</td>
<td>2.21</td>
<td>38</td>
</tr>
<tr>
<td>1NIW</td>
<td>Endothelial nitric-oxide synthase</td>
<td>1-14</td>
<td>0.172</td>
<td>2.06</td>
<td>37</td>
</tr>
<tr>
<td>1SY9</td>
<td>Cyclic-nucleotide-gated olfactory channel</td>
<td>1-14</td>
<td>0.170</td>
<td>2.02</td>
<td>38</td>
</tr>
<tr>
<td>1IQ5</td>
<td>Ca(^{2+})/calmodulin dependent kinase kinase</td>
<td>1-16</td>
<td>0.179</td>
<td>2.03</td>
<td>38</td>
</tr>
<tr>
<td>2BCX</td>
<td>Ryanodine receptor I</td>
<td>1-17</td>
<td>0.198</td>
<td>2.24</td>
<td>31</td>
</tr>
<tr>
<td>1CFF</td>
<td>Ca(^{2+}) pump</td>
<td>Extended</td>
<td>0.258</td>
<td>2.55</td>
<td>27</td>
</tr>
<tr>
<td>1G4Y</td>
<td>Gating domain of K(^+) channel</td>
<td>Extended</td>
<td>0.149</td>
<td>1.72</td>
<td>33</td>
</tr>
<tr>
<td>1K93</td>
<td>Adenylate cyclase domain of anthrax edema factor</td>
<td>Extended</td>
<td>0.215</td>
<td>1.81</td>
<td>24</td>
</tr>
<tr>
<td>2IX7</td>
<td>Myosin 5A</td>
<td>Extended</td>
<td>0.235</td>
<td>2.23</td>
<td>27</td>
</tr>
<tr>
<td>1NWD</td>
<td>Plant glutamate decarboxylase</td>
<td>Partially extended dimer</td>
<td>0.242</td>
<td>2.63</td>
<td>24</td>
</tr>
<tr>
<td>3CLN</td>
<td>Calcium-bound CaM</td>
<td>N.A</td>
<td>0.2496</td>
<td>2.05</td>
<td>21</td>
</tr>
</tbody>
</table>
Table 4.3 The binding affinities (μM) of Cx43 peptides to CaM (NMR), D-CaM and I-CaM.

<table>
<thead>
<tr>
<th>CaM-Cx43</th>
<th>Cx43-2p</th>
<th>Cx43-3p</th>
<th>Cx43-4p</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-CaM</td>
<td>77 ± 5(^a)</td>
<td>0.86 ± 0.02</td>
<td>NA</td>
</tr>
<tr>
<td>I-CaM</td>
<td>ND</td>
<td>NA</td>
<td>23.3</td>
</tr>
<tr>
<td>NMR(^b)</td>
<td>slow</td>
<td>fast</td>
<td>slow and fast</td>
</tr>
</tbody>
</table>

\(^a\) Data obtained from dansyl labeled C-lobe of half CaM
\(^b\) Chemical shift of slow exchange in NMR usually indicates strong binding (~10\(^{-8}\) M); chemical shift of fast exchange in NMR usually indicates relative weak binding (~10\(^{-6}\) M);
<table>
<thead>
<tr>
<th>CaM-Cx43-2p</th>
<th>Pocket 1</th>
<th>Pocket 2</th>
<th>Pocket 3</th>
<th>Pocket 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA, Δ NH-H (ppm)</td>
<td>Phe 12 (0.128), Phe 19 (N.A.), Val 35 (0.146), Met 36 (N.A.), Leu 39 (0.049), Phe 68 (0.062), Met 71 (0.120), Met 72 (N.A.)</td>
<td>Met 124 (N.A.), Phe 141 (0.158), Met 144 (N.A.)</td>
<td>Phe 89 (0.149), Phe 92 (N.A.), Leu 105 (0.121), Met 109 (N.A.)</td>
<td>Val 91 (N.A.), Phe 92 (N.A.)</td>
</tr>
</tbody>
</table>

AA, amino acid;
N.A. not available due to incomplete assignment.
Table 4. 5 Oculodentodigital dysplasia (ODDD) disease mutation overlapping with Cx43 CaMBD.

<table>
<thead>
<tr>
<th>Cx43 CaMBD</th>
<th>Cx43-2p</th>
<th>Cx43-3p</th>
<th>Cx43-4p</th>
</tr>
</thead>
</table>

*The Y230 Cx43 ODDD mutations alters 6 amino acids before truncating the protein after 236.*
5. IDENTIFICATION OF EF-HAND Ca^{2+} BINDING PROTEINS IN VARIOUS BIOLOGICAL SYSTEMS

5.1. Introduction

Cells maintain a Ca^{2+} concentration in a gradient of 20,000-fold between their intracellular (~100 nM free) and extracellular (mM) regions. Hundreds of cellular proteins have been identified to bind Ca^{2+} over a 10^6-fold range of affinities (nM to mM) (Fig. 1.3A). Ca^{2+} binding has been shown to be essential for stabilizing proteins as well as maintaining physiological cellular free Ca^{2+} concentrations as seen in buffer proteins such as calbindin D9k in mammalian intestinal epithelial cells (Fig. 1.3D). Depending on the nature of the Ca^{2+}-modulated event, some of these cellular proteins have been adapted to buffer Ca^{2+} concentration in low levels and others have been adapted to trigger cellular events by affecting Ca^{2+} concentration change. As shown in Fig. 1.2, bacterial cells also have a well-regulated cytosolic free Ca^{2+} concentration (approximately 0.1-2 M) that is significantly lower than that observed in the extracellular medium (mM) due to Ca^{2+} transporters and channels [234-237]. Viruses, on the other hand, utilize the universal Ca^{2+} signal to create a specific cellular environment to increase their viability (Fig. 5.1). Ca^{2+} plays important roles in viral gene expression, post-translational processing of viral proteins, virion structure formation, virus entry, and virion maturation and release.

5.1.1. EF-hand Ca^{2+} binding proteins and protein grafted approach

Ca^{2+} binding sites in proteins are classified as either non-continuous or continuous. For the protein with a continuous Ca^{2+} binding loop, a helix-loop-helix structure domain with Ca^{2+} binding capability was named as EF-hand motif because it is like the spread thumb and forefinger of the human hand, in which the Ca^{2+} ions are coordinated by the ligands within the loop. The term of EF-hand motif was
first used by Kretsinger to describe the Ca$^{2+}$ binding by parvalbumin which is also a Ca$^{2+}$ buffer protein in fast-contracting muscles [29]. Based on the conserved features of the Ca$^{2+}$ binding loop, EF-hand proteins have been divided into two major groups: the canonical EF-hands as seen in CaM, and the pseudo EF-hands exclusively found in the N-termini of S100 and S100-like proteins [22]. At present, more than 1000 protein with EF-hand motifs have been found in prokaryotes and eukaryotes and they can be classified into 77 distinct subfamilies [238-240]. However, there are likely numerous Ca$^{2+}$ binding proteins which contain an EF-hand motif in bacterial and viral genomes have not been discovered. The role of calcium in bacteria and virus is still elusive. The cooperativity and conformational changes of coupled EF-hand motifs in the well-known EF-hand protein, calbindin D9k, have to be further clarified with our grafting approach (Fig. 1.4). To understand the role of Ca$^{2+}$ in biological systems, we have predicted and analyzed potential EF-hand and EF-hand like Ca$^{2+}$ binding motifs on a genome-wide scale using our developed bioinformatic tool (http://www.chemistry.gsu.edu/faculty/Yang/Calciomics.htm). Proteins from bacteria and virus such as *Streptococcus pyogenes*, Sindbis virus and pox virus have been predicted and identified with our grafted approach as clarified in Chapter 1.3. As shown in Fig. 5.2, the single EF-hand motif grafted into CD2.D1 (CD2.Shr.EF) from Shr has a significant homology to that of CaM with all the conserved Ca$^{2+}$ binding ligand residues and two flanking helices (Fig. 5.2). Similarly, the other two predicted EF hand motifs in virus and two well-known EF-hand motifs were grafted into CD2.D1 as well. As indicated in Fig 5.2, all the loops in the canonical EF-hands family can be aligned well with the EF-hand loop III in the calcium modulated protein, calmodulin. The structure of CD2.D1 (blue) was draw by PyMOL with PDB 1hng. LRET or FRET occurs between aromatic Trp residues (orange) in CD2.D1 and Tb$^{3+}$ (green) ion in grated EF-hand loops (Magenta). The binding ligands at positions 1, 3,
5, 7, 9, 12 for canonical EF-hands and at positions 1, 3, 5, 7, 11, 14 for pseudo EF-hands are highlighted in red color.

5.1.2. Animalia: Calbindin D9k

Calbindin refers to several calcium-binding proteins including Calbindin D9k and Calbindin D28k. Both are Vitamin D-dependent calcium binding proteins present in intestinal epithelial cells (enterocytes) of mammalians or birds, respectively. Calbindin D9k can also be found in the kidney and uterus in some mammalian species. Calbindin D9k mediates the transport of calcium across the enterocytes from the apical side, where entry is regulated by the calcium channel TRPV6, to the basolateral side, where calcium pumps such as PMCA1 utilize intracellular adenosine triphosphate to pump calcium into the blood [241]. The transport of calcium across the enterocyte cytoplasm appears to be rate-limiting for calcium absorption in the intestine; the presence of calbindin increases the amount of calcium crossing the cell without raising the free concentration [242]. Calbindin D9k may also stimulate the basolateral calcium-pumping ATPases. Expression of calbindin D9k, like that of calbindin-D28k, is stimulated by the active vitamin D metabolite, calcitriol although the precise mechanisms are still controversial [243]. Calbindin D9k, a small Ca^{2+} binding protein (75 residues, MW 8,600) with two EF-hand motifs bundled together, is one of the primary model systems in the S100 protein family for the studies of Ca^{2+} binding properties of proteins [244]. It carries two distinct EF-hands: a canonical EF-hand at the C terminus and a pseudo EF-hand motif at the N-terminus. These two EF-hands were grated into CD2.D1 (CD2.Cal1.EF and CD2.Cal2.EF) in order to determine their metal binding affinity with cooperativity elimination between them. The oligomerization status of the engineered protein introducing by EF-hand insertion was studied by anisotropy fluorescence.
5.1.3. *Bacteria: streptococcal hemoprotein receptor of* *Streptococcus pyogenes*

Although bacterial cells do not have complex subcompartments or organelles, there is strong evidence that Ca$^{2+}$ plays an essential role in bacterial signaling, communication and stability similar to that observed in eukaryotic cells (Fig. 1, 2) [234-237, 245]. Bacterial cells also have a well-regulated cytosolic free Ca$^{2+}$ concentration (approximately 0.1-2 M) that is significantly lower than that observed in the extracellular medium (mM) due to Ca$^{2+}$ transporters and channels [234-237]. Similar to the eukaryotic systems, P-type ATPase Ca$^{2+}$ efflux pumps have been characterized from *Synechococcus* and *Flavobacterium*. A Ca$^{2+}$ transporter of *S. pneumoniae* is involved in Ca$^{2+}$-DNA uptake, lysis, and competence [246, 247]. Uptake of Ca$^{2+}$ and other divalent cations can also accompany uptake of phosphate by the phosphate transport system of *E. coli*. Furthermore, it has been reported that bacteria contain Ca$^{2+}$ binding proteins that are essential for cell adhesion and communication [248-251].

*Streptococcus pyogenes* is a spherical, Gram-positive bacterium that is the cause of group A streptococcal infections. Streptococci are catalase-negative. In ideal conditions, *S. pyogenes* has an incubation period of approximately 1–3 days. It is an infrequent, but usually pathogenic, part of the skin flora. It is estimated that there are more than 700 million infections each year and over 650,000 cases of severe, invasive infections that have mortality rate of 25 % [252]. The streptococcal hemoprotein receptor (Shr) from *Streptococcus pyogenes* is a surface protein with a role in iron uptake that has no significant homologues in other bacteria but shares partial homology with eukaryotic receptors such as Toll and G-protein dependent receptors (gi 15675635, GeneBank). Additional sequence analysis identified a leucine-rich repeat domain, an EF-hand Ca$^{2+}$ domain, and two NEAT domains [253].
5.1.4. *Virus: nsP1 of Sindbis and Pox*

As shown in **Fig. 5.1**, the interplay between viruses and Ca$^{2+}$ in the infected cell falls generally into three major categories: 1) viral proteins directly or indirectly disturb Ca$^{2+}$ homeostasis by altering membrane permeability and/or manipulating key components of the Ca$^{2+}$-signaling apparatus; 2) viral proteins directly bind to Ca$^{2+}$ for structural integrity or functionality; and 3) critical virus-host interactions depend on cellular Ca$^{2+}$-regulated proteins or pathways. EF-hands have been found abundantly in eukaryotes and bacteria. However, literature reporting EF-hand or EF-hand-like Ca$^{2+}$-binding motifs in virus proteins is scarce, possibly due to lack of accurate prediction methods and robust validating methodologies. With our developed method, 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 [254]. The Ca$^{2+}$-binding capabilities of these sequences from virus such as Sindbis Virus and Vaccinia virus (poxvirus family) have been experimentally verified in this chapter.

Sindbis Virus (SINV) is a member of the Togaviridae family, in the alphavirus subfamily. The virus causing fever in humans is transmitted by mosquitoes. The symptoms include arthralgia, rash and malaise. The virus was first isolated in 1952 in Cairo, Egypt and SINV has been linked to Pogosta disease in Finland recently [255]. SINV induces apoptosis in many vertebrate cells and apoptosis of neuroblastoma cells induced by SINV is dependent on viral replication but not dependent on a rise in intracellular Ca$^{2+}$ [256]. However, apoptosis occurred more rapidly in the absence of extracellular Ca$^{2+}$ and the mechanism remains unknown because the Ca$^{2+}$ binding domain in SINV remains to be indentified.

Vaccinia virus (VACV or VV) belongs to the poxvirus family, a large complex of double-strand DNA viruses that replicate in the cytoplasm of the infected cell in a highly regulated manner [257]. Western
Reserve A38L protein from the vaccinia virus represents the first example of a virus protein which directly or indirectly promotes the influx of extracellular Ca^{2+} [258].

Nonstructural protein (nsP) encoded by a viral genome that are produced in the organisms they infect, but not packaged into the virus particles. Some of these proteins may play roles within the infected cell during virus replication or act in regulation of virus replication or virus assembly. We grafted two predicted 29-residue EF-hand motifs, from the nonstructural protein 1 (nsP1) of SINV (CD2.Sin.EF) and poxvirus (CD2.Pox.EF), into CD2.D1 to examine their Ca^{2+} binding capability by using aromatic residue-sensitized Tb^{3+} luminescence resonance energy transfer (Tb^{3+}-LRET).

5.2. Results

5.2.1. Probing metal binding of grafted EF-hand by FRET pair of Trp and Tb^{3+}.

**CD2.Pox.EF.**

Engineered proteins with grafted EF-hand motifs were generated as described in Chapter 2.1 and proteins were expressed and purified following the method in Chapter 2.3. Before studying the metal binding properties of these engineered proteins, the native structure of domain 1 of CD2 had to be confirmed. As shown in **Fig. 5.3A**, the structural integrity of the engineered protein with grated EF-hand from poxvirus is confirmed by the fluorescence emission peaks at 315 nm and 330 nm. The Trp emission peak at 330 nm (W32) of CD2.Pox.EF exhibits a blue shift from 350 nm which is the same as CD2 without insertion. The Trp emission peak at 315 nm is from the inserted E-helix of calcium binding loop. It indicated the hydrophobic environment of Trp residues and the correct folding of the engineered protein. Circular Dichroism (CD) spectroscopy also was employed as a way to confirm the structural integrity of the engineered protein by monitoring the secondary structure. As seen in **Fig. 5.3B**, the major secondary structure contents of CD2.Pox.EF and CD2.WT are β sheet which has a
negative peak at 218 nm. The negative shoulder peaks of CD2.Pox.EF at 224 nm and 202 nm indicate
the helical and random coil contribution from the inserted loop. The average molar ellipticity for both
the wild type and engineered protein are almost the same (Fig. 5.3B). The noise indicating in the CD
spectrum below 200 nm is mainly from Cl− in the buffer (100 mM KCl, 10 mM Tris-HCl, pH 7.2).
After confirmation of native structure of the engineered protein, FRET based Tb3+ titration of the
engineered protein from the nsP1 of poxvirus (CD2.Pox.EF) has been tested. As shown in Fig. 5.4A,
2 to 5 µM CD2.Pox.EF was prepared in 10 mM PIPES, 100 mM KCl at pH 6.8. Intrinsic Trp residues
close to grated EF-hand loop was excited at 280 nm. If Tb3+ binds to the predicted loop from pox
virus, the fluorescence signal from Trp by FRET at 545 nm will increase until it reaches saturation.
The inset is the fluorescence signal from free Tb3+ ion which also can be excited at 280 nm. The
contribution of fluorescence signal from Tb3+ ion was subtracted from Tb3+ signal in protein. As
shown in Fig. 5.4B, the fluorescence signal decreased by adding calcium into 2 µM CD2.Pox.EF pre-
incubated with 30 µM TbCl3 because Tb3+ was competed out from the binding pocket. The
dissociation constant of calcium (258 ± 12 µM) was calculated by Equation 2.7 with apparent
dissociation constant fitted in the inset.

CD2.Cal1.EF and CD2.Cal2.EF.

3 µM CD2.Cal1.EF or CD2.Cal2.EF was prepared in 10 mM PIPES, 100 mM KCl at pH 6.8. The
inserted loops in CD2.D1 of these two proteins are from calcium binding EF hand motif 1 and 2 of the
calcium buffer protein Calbinidin D9k. CD2.Cal1.EF contains the pseudo EF-hand with 14 residues in
the binding loop while CD2.Cal2.EF has the classical EF-hand with 12 amino acids in the binding
loop. Except the intensity differences, the excitation spectrum for both protein are almost same when
fix emission wavelength at 340 nm for CD2.Cal1.EF and at 332 nm for CD2.Cal2.EF (Fig. 5.5A). As
shown in Fig. 5.5B, CD2.Cal2.EF folds better than CD2.Cal1.EF because the peak with the maximum intensity for CD2.Cal2.EF in the emission spectrum shifts further to the lower wavelength than CD2.Cal1.EF and fluorescence intensity higher when two protein concentrations in the measurement are the same (3 µM). As shown in Fig. 5.6AB, when the Tb3+ signal reached saturation after titration, the final fluorescence intensity are similar for both grafted EF hand motifs. However, the total Tb3+ required to reach saturation for CD2.Cal1.EF is more than 14 folds higher than CD2.Cal2.EF. The dissociation constants of Tb3+ for CD2.Cal1.EF (174 µM) and CD2.Cal2.EF (1.2 µM) was obtained by fitting F/F0 of Tb3+ signal with 1 to 1 binding equation (Equation 2.6).

CD2.Shr.EF and CD2.Sin.EF.

Two predicted 29-residue EF-hand motifs, one from the Shr of S. pyrogenes (CD2.Shr.EF) and the viral nsP1 of Sindbis virus (CD2.Sin.EF), were grafted into CD2.D1 to examine their Ca2+ binding capability by using aromatic residue-sensitized Tb3+ luminescence resonance energy transfer (Tb3+-LRET) (Fig. 5.7). As indicated in Fig. 5.7A, both engineered proteins fold well by monitoring intrinsic Trp fluorescence. Circular dichroism studies of both engineered proteins showed a notable trough at ~216 nm which is characteristic of β-sheet structure. More negative signals were observed below 240 nm due to the contribution from the insertion of the helix-loop-helix sequences (Fig. 5.7B). Both proteins were able to bind the Ca2+ analog, Tb3+, with affinities of 25.1 µM for CD2.Shr.EF and 16.4 µM for CD2.Sin.EF (Fig. 5.7CD). The biological relevance of these EF-hand Ca2+-binding motifs will be further investigated in the virus transfected system.

5.2.2. Revealing Ca2+ dependent conformation change by CD spectroscopy.

Far UV CD spectrometry of CD2.Cal1.EF in aqueous solution yields a spectrum with minima at 202 nm and a maximum at 194 nm in both Ca2+ free and Ca2+ loaded condition (Fig. 5.8A). As shown in
**Fig. 5.8B**, far UV CD spectrometry of CD2.Cal2.EF in apo form yields a spectrum with minima at 209 nm and a maximum at 196 nm. After adding Ca\(^{2+}\), far UV CD spectrometry of CD2.Cal2.EF in holo form yields a spectrum with minima at 219 nm and a maximum at 195 nm, characteristic of a protein with substantial \(\beta\)--sheet secondary structure. **Fig. 5.8C** summarizes the secondary structure content for CD2 WT and CD2 engineered proteins predicted by DichroWeb based on experimental far UV CD spectrum. The predicted major secondary structure for domain 1 of CD2 (CD2.WT) is \(\beta\)--sheet (40.3% in Table 5.1), which is in good agreement with the secondary structure contents determined by X-ray crystallography [259]. The engineered proteins with grafted EF hand motifs from calbindin D9k increase their \(\beta\)--sheet content and decrease \(\alpha\)-helical content after addition of 2 mM Ca\(^{2+}\). For CD2.Pox.EF and CD.Chr.EF, the secondary structure remains almost the same between the EGTA form and Ca\(^{2+}\) loaded form (Fig. 5.9). CD2.Cal2.EF has the largest average molar ellipticity and reveals dramatic Ca\(^{2+}\) depend secondary structure conformation change comparing to other engineered proteins.

### 5.2.3. **Observing oligomeric stage of CD2 with grafted EF-hand by PFG-NMR measurement and fluorescence anisotropy.**

Next, the oligomeric state of the grafted EF-Hand motifs from calbindin D9k was examined by two different techniques: pulsed-filed gradient NMR (PFG-NMR) and fluorescence anisotropy. PFG NMR has been widely used to study the molecular motion, effective dimensions and oligomeric states of proteins in solution [111]. The size of proteins can be estimated by measuring the diffusion constants and the hydrodynamic radius is calculated with Equation 2.2 to 2.4 with reported lysozyme size as a reference.

The diffusion constants of CD2 grated protein CD2.Cal1.EF and CD2.Cal2.EF were measured under
Ca\(^{2+}\)-saturated and Ca\(^{2+}\)-depleted conditions to determine where the isolated EF-hand motifs from calbindin D\(_9k\) undergo dimerization on metal binding. As shown in Fig. 5.10, NMR signal decay when the field strength was increased from 0.2 to 31 G \(\cdot\) cm\(^{-1}\). The calculated hydrodynamic radius of the CD2 monomer (19.4 Å) was close to the one (19.6 Å) reported previously [112]. Table 4.2 summarizes all the calculated hydrodynamic radii value. The calculated hydrodynamic radii of CD2.Cal1.EF are 21.1 ± 0.2 Å for apo form and 20.4 ± 0.3 Å for Ca\(^{2+}\)-loaded form (Fig. 5.10A); the calculated hydrodynamic radii of CD2.Cal2.EF are 27.4 ± 0.3 Å for Ca\(^{2+}\)-loaded form (Fig. 5.10B). According to calculations using the spherical shape of macromolecules, the hydrodynamic radius of the protein will increase by 27% on formation of the dimer [260]. The increase in size for CD2.Cal2.EF in 10 mM CaCl\(_2\) is 41% to CD2 monomer and 34% to CD2.Cal1.EF monomer, indicating that it exists as a dimer in solution in the presence of Ca\(^{2+}\).

Fluorescence anisotropy was also used to determine the oligomeric status of the engineered protein in the absence of in the presence of Ca\(^{2+}\). As shown in Fig. 5.11, the fluorescence anisotropy of CD2.Cal1.EF remains at the same value (~0.08) when protein concentration varies from 2 to 10 µM and binding with metal also does not alter the value. On the contrary, the fluorescence anisotropy of 5 µM CD2.Cal2.EF increased from 0.076 (0.079 in 10 µM) in 300 µM EDTA to 0.084 (0.086 in 10 µM) in 5 mM CaCl\(_2\) (Fig. 5.12). However, there is no obvious change for 2 µM CD2.Cal2.EF. The anisotropy value is related to the molecule tumbling rate in the aqueous solution. The bigger fluorescence anisotropy value, the slower the molecule tumbling rate and the bigger the protein size. The data suggests that CD2.Cal2.EF forms a dimer in a Ca\(^{2+}\)-dependent and concentration dependent way while CD2.Cal1.EF does not.
5.3. Discussion and summary.

Table 5.3 summarizes the Tb$^{3+}$ dissociation constants for all the engineered proteins studied in this chapter. The negative charge of amino acids in the EF-hand binding loops plays a key role in the metal binding affinities. CD2.Cal2.EF, CD2.Shr.EF and CD2.Sin.EF contain 4 negative charged residues in the loop (Fig. 5.2). They also have higher Tb$^{3+}$ binding affinities than CD2.Pox.EF which has only 3 negative charged amino acids at the conserved positions. The correct folding is also important to study the side-specific metal binding properties by grafted approach. From both far-UV spectrum and Trp fluorescence, CD2.Cal1.EF does not fold very well, thus forming the improper arranged binding pockets with relative low binding affinities. When the final Tb$^{3+}$ concentration in the fluorescence titration cuvette is getting very high, the unspecific binding to CD2 host protein will occur which also affects the accurate estimation of metal binding affinities in the grafted proteins. The biological relevance of these EF-hand Ca$^{2+}$-binding motifs are necessary to understand the role of calcium in both bacterial and viral systems.

Shaw et al. [261] first reported that an isolated EF-hand III from skeletal troponin C dimerizes in the presence of Ca$^{2+}$. EF hands from parvalbumin and calbindin D$_{9K}$ have also been shown to exhibit Ca$^{2+}$-dependent dimerization [262-264]. Wojcik et al. [239] have shown that the isolated 12-residue peptide from calmodulin (CaM) EF-hand motif III does not dimerize in the presence of Ca$^{2+}$, but dimerizes to form a native-like structure in the presence of Ln$^{3+}$, which has a similar ionic radius and coordination properties to Ca$^{2+}$. They concluded that local interactions between the EF-hand Ca$^{2+}$-binding loops alone could be responsible for the observed cooperativity of Ca$^{2+}$ binding to EF-hand protein domains. To examine the key determinants for Ca binding and Ca-induced conformational change, peptides or fragments encompassing the helix--loop--helix motif produced by either
synthesis or cleavage are common methods. With our lab developed grafting approach, the isolated calcium binding loop 2 from calbindin D_9k has been determined to form dimer in a Ca^{2+} and concentration dependent way (Table 5.2). The Ca^{2+} binding to CD2.Cal2.EF could induce large conformational change in the EF-hand helix - loop - helix motif. Upon binding to Ca^{2+}, CD2.Cal2.EF behaves more like CD2.WT in the secondary structure maybe because the glycine linker is getting close induced by metal binding in the EF-hand motif and disruption of host protein is getting less.

Overall, based on sequence homology, a straightforward and fast method has been developed to detect linear Ca^{2+}-binding motifs from genomic information. Experimentally, we have also developed a robust and reliable grafting approach to study Ca^{2+}-binding properties of continuous Ca^{2+} binding sites. This novel approach has been successfully used to dissect site-specific Ca^{2+} binding affinity and cooperativity among the four canonical EF-hands in the prototypical Ca^{2+}-binding protein, calmodulin.

The combination of these two approaches is expected to enable us to explore more Ca^{2+} binding sites that are underrepresented due to the limitation of available methodology. In addition, PFG-NMR and fluorescence anisotropy provide both accurate and fast measurement for the native protein size determination in solution.
Figure 5. 1. Calcium is involved in multiple steps of virus life cycle.

The virus life cycle can be divided into several steps, attachment, penetration, un-coating, replication, assembly, release and budding. Each step is controlled tightly by calcium through calcium binding proteins.
CaM.EF.III  EIREAFRVF DKDGNGYISAAE LRHVMTNL
Bacteria  CD2.Shr.EF:  EKVLVKLGK DLDGDGKLSKTE LEQIRGEL
Virus     CD2.Sin.EF:  SKWAKERKD DLDNEKMLGTE RKLTYGCL
Virus     CD2.Pox.EF:  SKWAKERKD DYDGNGTETRGE RKLTYGCL
Mammalian CD2.Cal2.EF: TLDELFEEL DKNGDGEVSFEE FQVLVKKI
Mammalian CD2.Cal1.EF: EELKGIEFK YAAKEGDPNQLSKE ELKLLLQT

**Figure 5.2.** The Predicted EF-hand loops in various biological systems.

The entire EF-hands (E-helix, loop and F-helix) from predicted proteins were grated into CD2 domain 1 (CD2.D1) as host protein. The structure of CD2.D1 (blue) was draw by PyMOL with PDB 1hng. LRET or FRET occurs between aromatic Trp residues (orange) in CD2.D1 and Tb$^{3+}$ (green) ion in grated EF-hand loops (Magenta). All the loops in canonical EF-hands family can be aligned well with EF-hand loop III from calcium modulated protein, calmodulin. The binding ligands at positions 1, 3, 5, 7, 9, 12 for canonical EF-hands and at positions 1, 3, 5, 7, 11, 14 for pseudo EF-hands are highlighted in red color.
Figure 5. Characterization of CD2 grated protein (CD2.Pox.EF) by intrinsic Trp fluorescence emission spectrum (A) and circular dichroism spectroscopy (B). A, the structural integrity of the engineered protein with grated EF-hand from poxvirus is confirmed by the fluorescence emission peaks at 315 nm and 330 nm. The solid line with solid square stands for grated protein CD2.Pox.EF; the dashed line with hollow square stands for the domain 1 of CD2 (CD2.WT). The blue shift of the emission peak from 350 nm indicated the hydrophobic environment of Trp residues and the correct folding of the engineered protein. B, the major secondary structure contents of CD2.Pox.EF (●) and CD2.WT (○) are β sheet which has a negative peak at 218 nm.
Figure 5.4. FRET based Tb\textsuperscript{3+} titration of the engineered protein from the nsP1 of poxvirus (CD2.Pox.EF).

A, 2 to 5 µM CD2.Pox.EF was prepared in 10 mM PIPES, 100 mM KCl at pH 6.8. Intrinsic Trp residues close to grater EF-hand loop was excited at 280 nm. If Tb\textsuperscript{3+} binds to the predicted loop from pox virus, the fluorescence signal from Trp by FRET at 545 nm will increase until it reaches saturation. The inset is the fluorescence signal from free Tb\textsuperscript{3+} ion which also can be excited at 280 nm. The contribution of fluorescence signal from Tb\textsuperscript{3+} ion was subtracted from Tb\textsuperscript{3+} signal in protein. B, the binding constant of Tb\textsuperscript{3+} (86 ± 2 µM) was obtained by fitting F/F\textsubscript{0} of Tb\textsuperscript{3+} signal with 1 to 1 binding equation (Equation 2.6). The fluorescence signal decreased by adding calcium into 2 µM CD2.Pox.EF with pre added 30 µM TbCl\textsubscript{3} because Tb\textsuperscript{3+} was competet out from the binding pocket. The dissociation constant of calcium (258 ± 12 µM) was calculated by Equation 2.7 with apparent dissociation constant fitted in the inset.
Figure 5. Characterization of CD2 grated protein (CD2.Cal1.EF and CD2.Cal2.EF) by intrinsic Trp fluorescence excitation spectrum (A) and emission spectrum (B).

A, the fluorescence excitation spectrum of the engineered protein with grated EF-hand motifs (CD2.Cal1.EF, red line with solid dots and CD2.Cal2.EF, black line with solid dots) from Calbindin D9k. B, the structural integrity of the engineered proteins are confirmed by the fluorescence emission peaks at 332 nm (CD2.Cal1.EF, red line with hollow dots) and 340 nm (CD2.Cal2.EF, black line with solid dots). The blue shift of the emission peak from 350 nm indicated the hydrophobic environment of Trp residues and the correct folding of the engineered protein. All the fluorescence measurements were taken in 10 mM PIPES, 100 mM KCl at pH 6.8 with 3 µM proteins.
2 to 5 µM CD2.Cal1.EF (A) or CD2.Cal2.EF (B) was prepared in 10 mM PIPES, 100 mM KCl at pH 6.8. Intrinsic Trp residues close to grafted EF-hand loop was excited at 282 nm. If Tb$^{3+}$ binds to the grafted loops, the fluorescence signal from Trp by FRET at 545 nm will increase until it reaches saturation. The contribution of fluorescence signal from Tb$^{3+}$ ion was subtracted from Tb$^{3+}$ signal in protein. The binding constants of Tb$^{3+}$ for CD2.Cal1.EF (174 µM) and CD2.Cal2.EF (1.2 ± 0.1 µM) was obtained by fitting F/F₀ of Tb$^{3+}$ signal with 1 to 1 binding equation (Equation 2. 6).
Figure 5.7. Probing metal binding properties of predicted EF-hand Ca^{2+} binding motifs from Shr of *Streptococcus pyogenes* (CD2.Shr.EF) and from the nsP1 of Sindbis virus (CD2.Sin.EF).

A, the structural integrity of the engineered proteins are confirmed by the fluorescence emission peaks at 315 nm and 330 nm. B, far ultra-violet circular dichroism spectra of CD2 wild type and the engineered proteins. C and D, Tb^{3+}-binding curves of the engineered proteins CD2.Shr.EF and CD2.Sin.EF. The titration curve is fitted for a 1:1 binding stoichiometry (Equation 2.6). All the fluorescence measurements were taken in 10 mM PIPES, 100 mM KCl at pH 6.8 with 2 to 5 µM proteins. All the CD measurements were taken in 10 mM Tris-HCl, 100 mM KCl at pH 7.2 with 8 to 10 µM proteins.
Figure 5. Monitoring Ca^{2+} dependent secondary structure changes of CD2.Cal1.EF and CD2.Cal2.EF by far ultra-violet circular dichroism spectra.

A, the far-UV CD spectrum of CD2.Cal1.EF with (●) and without calcium (○). B, the far-UV CD spectrum of CD2.Cal2.EF with (●) and without calcium (○). C, the secondary structure prediction (DichroWeb) of CD2.WT and engineered proteins from calbindin D_{9k} with and without calcium. The dash bar stands for alpha helical content and the grey bar stands for the Beta sheet content. D, the calcium binding EF-hand loop 2 (brown) of calbindin D_{9k} was highlighted in its NMR solved high resolution structure (PDB code: 2BCA). All the CD measurements were taken in 10 mM Tris-HCl, 100 mM KCl at pH 7.2 with 8 to 10 µM proteins. 2 mM EGTA or Ca^{2+} was added in the buffer to obtain apo or holo form of engineered proteins.
Figure 5. 9. Monitoring Ca$^{2+}$ dependent secondary structure changes of CD2.Pox.EF and CD2.Shr.EF by far ultra-violet circular dichroism spectra.

A, the far-UV CD spectrum of CD2.Pox.EF with 2 mM (◆) or 4 mM (■) calcium and without calcium (□). B, the far-UV CD spectrum of CD2.Shr.EF with 2 mM (◆) or 4 mM (■) calcium and without calcium (□). All the CD measurements were taken in 10 mM Tris-HCl, 100 mM KCl at pH 7.2 with 8 to 10 µM proteins.
Figure 5. Determination of the oligomeric state of CD2.Cal1.EF and CD2.Cal2.EF by PFG-NMR.

A, the NMR signal decay of apo-CD2.Cal1.EF (○), Ca²⁺-CD2.Cal1.EF (●) as a function of field strength. B, the NMR signal decay of Ca²⁺-CD2.WT (●), Ca²⁺-CD2.Cal2.EF (■) as a function of field strength. The calculated hydrodynamic radii of the CaM and complexes are indicated on the top.
Figure 5. Determination of the oligomeric state of CD2.Cal1.EF by fluorescence anisotropy.

The fluorescence signal in different excitation and emission orientation was measured in 10 mM KCl, 50 mM Tris-HCl at pH 7.5 with 300 µM EDTA or 5 mM Ca\(^{2+}\) for 2 µM (A), 5 µM (B) and 10 µM (C) CD2.Cal1.EF. The fluorescence anisotropy was calculated by Equation 2.8. The fluorescence maximum intensity was measured at \(\lambda_{ex} = 282\) nm and \(\lambda_{em} = 340\) nm with an integration time of 15s.
Figure 5. Determination of the oligomeric state of CD2.Cal2.EF by fluorescence anisotropy.

The fluorescence signal in different excitation and emission orientation was measured in 10 mM KCl, 50 mM Tris-HCl at pH 7.5 with 300 µM EDTA or 5 mM Ca²⁺ for 2 µM (A), 5 µM (B) and 10 µM (C) CD2.Cal2.EF. The fluorescence anisotropy was calculated by Equation 2.8. The fluorescence maximum intensity was measured at λex = 282 nm and λem = 332 nm with an integration time of 15s.
Table 5.1 The secondary structure prediction by DichroWeb.

<table>
<thead>
<tr>
<th>Grafted proteins</th>
<th>CD2.WT</th>
<th>CD2.Cal1.EF</th>
<th>CD2.Cal2.EF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Apo</td>
<td>Holo</td>
<td>Apo</td>
</tr>
<tr>
<td>α Helix (%)</td>
<td>7.6</td>
<td>10.7</td>
<td>8.6</td>
</tr>
<tr>
<td>β Sheet (%)</td>
<td>40.3</td>
<td>33.2</td>
<td>37.8</td>
</tr>
</tbody>
</table>
Table 5.2 Hydrodynamic radii ($R_{\text{hydro}}$) and fluorescence anisotropy ($r$) of calbindin D$_{9k}$ grafted proteins.

<table>
<thead>
<tr>
<th>Grafted proteins</th>
<th>CD2.WT</th>
<th>CD2.Cal1.EF</th>
<th>CD2.Cal2.EF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Apo</td>
<td>Holo</td>
<td>Apo</td>
</tr>
<tr>
<td>R$_{\text{hydro}}$ (Å)</td>
<td>19.4 ± 0.4</td>
<td>21.1 ± 0.2</td>
<td>20.4 ± 0.3</td>
</tr>
<tr>
<td>Apo</td>
<td>0.077 ± 0.003</td>
<td>0.076 ± 0.003</td>
<td></td>
</tr>
<tr>
<td>Dimer</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 5.3 The summary of Tb$^{3+}$ dissociation constants ($K_d$) in grafted proteins.

<table>
<thead>
<tr>
<th>Grafted proteins</th>
<th>$K_d$ (µM)</th>
<th>Folding</th>
<th>Negative charges in the loop</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD2.Cal1.EF</td>
<td>174</td>
<td>Partial</td>
<td>3</td>
</tr>
<tr>
<td>CD2.Cal2.EF</td>
<td>1.2 ± 0.1</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>CD2.Shr.EF</td>
<td>25.1 ± 0.6</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>CD2.Sin.EF</td>
<td>16.4 ± 0.1</td>
<td>Yes</td>
<td>4</td>
</tr>
<tr>
<td>CD2.Pox.EF</td>
<td>86.0 ± 2.0</td>
<td>Yes</td>
<td>3</td>
</tr>
</tbody>
</table>
6. MASS SPECTROMETRY OF PROTEIN STUDY

6.1. Introduction

Mass spectrometry has come a long way from the discovery of isotopes for many of the chemical elements to the application to large, highly polar molecules such as proteins. Electrospray ionization mass spectrometry (ESI-MS) and Matrix-assisted laser desorption and ionization mass spectrometry (MALDI-MS) are two major techniques for protein-protein interaction study based on the mass of molecules. In 1985, Franz Hillenkamp, Michael Karas and their colleagues found that the alanine could be ionized more easily by a pulsed 266 nm laser if it is mixed with tryptophan with aromatic group [265]. They were the first to define the term Matrix-assisted laser desorption and ionization (MALDI). Later, Koichi Tanaka from Shimadzu Corp. applied MALDI mass spectrometry to analyze biological macromolecules in 1988 and received the 2002 Nobel Prize in Chemistry because of this high-impact discovery [266]. MALDI-MS is usually used to investigate protein-protein interaction due to its soft ionization technique without the fragmentation. MALDI-MS is frequently used to identify proteins (protein ID) by peptide mass fingerprinting after enzyme digestion and to search protein post translational modification sites such as methylation and phosphorylation because of its fast process and high-throughput screening. MALDI-MS is a power technique for protein study with less sample consumption, tolerance to high salt concentration and high detection limit for large protein complex. This chapter is focused on protein study using MALDI-MS.

As shown in Fig. 6.1, MALDI-MS instrument includes five major parts: sample plate, laser generator, flight tube, reflector and detector. In brief, samples with either a positive charge (Positive mode) or a negative charge (Negative mode) for analysis are excited by laser and are ejected into the flight tube through an electronic force. The speed of the samples depends on their mass over charge ratio. Detector records the time of flight (TOF) of the samples which can be used to calculate their molecular weight after standard calibration. For the smaller molecules, a reflector is used to increase the flight length of the samples in order to improve the accuracy in the electron mode. For the larger molecule, linear mode can
be used to improve the sensitivity. Reflectron mode and linear mode use different detectors, reflector
detector and linear detector, respectively. Reflectron mode is good for peptide (<5000 kDa) analysis with
internal calibration strand, especially for protein identification with the peptide mass fingerprinting. Liner
mode is suitable for identification of protein complex and protein oligomer.

The MALDI-MS samples have to be mixed with different matrixes based on the properties of
samples. The matrix refers to the crystallized molecule with UV abortion. Table 6.1 summarizes all of the
commonly used matrixes for protein study in MALDI-MS and words written in red indicate their major
applications in protein study. The MALDI-MS metal plate is pre-coated with isopropanol which forms an
organic thin layer on the surface. The sample is co-crystalized with matrix on the metal plate (Fig. 6.2).
As shown in Fig. 6.2, the matrix absorbs the UV laser light triggered by a laser beam during the
desorption process. The matrix not only passes the proton to samples, but also transfers the energy to
samples from laser during the ionization process. Charged samples are extracted into flight tube.

The protein sample preparation method in MALDI-MS can be the key to obtaining the spectrum
successfully. Proper sample preparation procedure is vital for improving the efficiency of MALDI-MS.
The formation of homogeneous, small and evenly distributed matrix/analyte co-crystal reduces the matrix
signal and yields a reproducible signal because the overall ionization process is facilitated.
Inhomogeneous crystallization could cause the “sweet spot” on the crystal surface which affects the spot--
to-spot and the sample to sample reproducibility. Most of the common matrixes are not soluble in water
and should be prepared freshly in organic/aqueous solution mixture. Although MALDI-MS allows higher
salt concentration to exist in the sample than ESI-MS, a low quality spectrum is often produced when the
sample contains a low abundance of the analyte and a high concentration of the contaminants. It is
recommended to purify the sample before performing the MALDI-MS analysis. Once the samples are
spotted on the plate with high levels of contaminants, cold water should be used to carefully wash away
the contaminants such as salt on the crystal surface. The water soluble contaminant can be reduced by this
simple method because matrix/analyte co-crystal is only soluble in organic solvent. In order to obtain
homogeneous co-crystallization for the low abundance analyte, matrix dissolved in highly volatile organic solvent such as acetone can be sprayed on the plate before spotting the sample. Under microscope, it is clear to see that matrix forms homogeneous and evenly distributed crystal surface in such a fast crystallization process.

MALDI-MS provides a fast and accurate method for identification of protein modification sites. Functional groups with certain properties are often introduced to protein for protein-protein interaction and metal binding study through protein modification. The intrinsic fluorescence of CaM comes from tyrosine and phenylalanine. The fluorescence from tyrosine and phenylalanine in CaM is weak and often overlaps with the CaM binding peptide. Dansyl chloride was first introduced by C.Gros and B. Labouesse in 1969 to study amino acids, peptides and proteins [267]. Later on, CaM was labeled with dansyl chloride to study the binding of inhibitory ligands [151]. Dansyl groups majorly interact with free amine group at N-termini of the protein at the physiological pH 7. However, more than one dansylation site have been found in CaM by MALDI-MS study and the location of dansylation sites has been suggested in this chapter. Therefore, IAEDANS dye for the specific labeling with the CaM mutant with additional cysteine added into the end of C-termini of CaM, T34C and T110C was introduced to study the domain dependent interaction between CaM and gap junction peptides. The condition of the reaction between IAEDANS and cysteine in CaM was also improved based on the reaction efficiency obtained from MALDI-MS spectrum.

6.2. Dansylation of CaM and its variants

As shown in Fig. 6.3A, CaM has 8 potential reaction sites (7 lysines and N-termini highlighted in grey color) with a free amine group in the structure. At pH 7.0, all the amine groups in the lysine are protonated and no reaction occurs between these lysines and dansyl chloride. However, up to 4 dansylation sites were observed in MALDI-MS spectrum when 1 mM Ca\textsuperscript{2+}-CaM was incubated with 5 molar excess of dansyl chloride at pH 7.0 at 4°C for overnight (Fig. 6.3B). This result indicates that at least three lysines in CaM can be dansylated at pH 7.0 and major modified CaM are single site and double
sites labeling. We kept all the condition the same but varied the initial protein concentration for dansylation. Three sites of CaM were dansylated when protein concentrations were 0.5 mM (Fig. 6.3C) and 0.25 mM (Fig. 6.3D). The peaks of CaM with three dansylation sites in these two reactions were much smaller than the one in 1 mM high protein concentration. We further decreased the initial CaM concentration to 0.125 mM (Fig. 6.3E), and the efficiency of the labeling was affected. The major protein peak became the free CaM without label. The efficiency of the dansylation reaction depends on the initial protein concentration and specific single label on N-termini is not available for CaM dansylation reaction. The best initial CaM concentration we found is 0.25 mM with single label as major peak and less double and triple labeling.

In order to locate the possible dansylation sites and to study the domain specific interaction between CaM and CaM binding domain peptide, half CaM for both N-lobe (N-CaM, 1-75 aa) and C-lobe (C-CaM, 75-148 aa) were engineered and produced. As shown in Fig. 6.4, both pH effect on the labeling efficiency and half CaM labeling also have been tested at 6.5 and 7.0. The peak intensity of free CaM at pH 6.5 (Fig. 6.4DE) is much higher than pH 7.0 (Fig. 6.4GH) for both CaM and N-CaM. Therefore, the labeling efficiency at pH 7.0 is much higher than pH 6.5. As shown in Fig. 6.4FI, C-CaM is easier to degrade than N-CaM because the major peak is the fragment of C-CaM which is the degradation product, missing either FVQMMTAK (MW 955.20 Da) from C-termini or MKDTDSEE (953.97) from N-termini. The labeling efficiency for C-CaM is almost 100% since the free CaM peak before reaction in Fig. 6.4F almost disappeared in Fig. 6.4I after the reaction. In contrast, for N-CaM, the major product after dansylation is still free N-CaM. The dansylation labeling efficiency of C-CaM is much higher than N-CaM. According to the labeling efficiency by comparing the protein and reaction product intensity for both entire CaM and two half CaM, it is highly possible that the first dansylation site in CaM is the same as the one in C-CaM (> 90%) and second dansylation sites in CaM is the same as the one in N-CaM (< 50%). According to the results of MALDI-MS, the optimal initial CaM concentration in the reaction of
dansylation on CaM is 0.25 mM and reaction at pH 7.0 has higher labeling efficiency. Even at pH 6.5, dansyl labeling on lysine group still cannot be eliminated.

6.3. IAEDANS dye labeling of CaM cysteine variants

CaM does not have a cysteine in its sequence and cysteine mutants were introduced into CaM at the C-terminal end (CaM-Cys, Fig. 6.5A) or to replace threonine at 34 (T34C, Fig. 6.5B) and 110 (T110C, Fig. 6.5B). The domain specific interaction between CaM binding peptide and these CaM mutants can be studied by fluorescence after specific labeling with IAEDANS on the cysteine residues. Similar to Dansyl groups, IAEDANS can be excited at 335 nm and have an emission wavelength with maxima intensity at ~ 490 nm, an emission that is distinguishable from the intrinsic protein aromatic group. In order to get maximum reaction efficiency, the CaM cysteine mutants were pre-incubated with 10 mM DTT, a reduced environment. The 10 mM DTT was removed by an Amicon concentrator with Nitrogen gas pressure. The reaction took place in 10 mM Tris, 100 mM KCl, pH 7.4 (degassed to remove all the oxygen) with 6 M urea for overnight incubation at 4 °C. The reaction bottle was filled with nitrogen and kept dark during labeling. As shown in Fig. 6.5 DEGH, both CaM-Cys and T34C can be labeled with IAEDANS with almost 100% efficiency. However, T110C could be labeled possibly due to its local environment (Fig. 6.5 FI). Since CaM is very stable and can be refolded back easily from a denatured, linear condition in 6 M urea to facilitate the reaction rate and efficiency (Fig. 6.6).

The semi-quantitative analysis for dansylation and IAEDANS labeling efficiency is not very accurate because both dansyl and IAEDANS groups have the same properties as the MALDI matrix. The aromatic group in these two dyes can absorb the laser excitation energy at ~335 nm in the MALDI instrument and facilitate the flight ability of the modified protein molecule. By the matrix effect of these labeling dyes, the population of modified protein molecule could be overestimated because it is more easily projected than the protein molecule without the label. Therefore, the dansyl label can be used to raise the peptide detection limit with a low abundance in the sample by the improvement of its ionization efficiency [268, 269].
6.4. **Protein-protein interaction study of CaM with gap junction peptide.**

MALDI-MS has been considered a “softly-soft” ionization mass spectrometry technique in the protein-protein interaction study. The average dissociation energy of intermolecular non-covalent bond is around 20 kJ/mole, which is almost 20 folds less than the energy required to cleave a covalent bond. The properties of the soft ionization process make MALDI-MS an advanced technique to explore the protein complex without breaking its weak non-covalent bond. MALDI-MS is more tolerant of buffers, denaturants, detergents, salts and many other contaminates than the ESI-MS. MALDI-MS does not require specific solvent combination to boost the signal in ESI-MS. MALDI-MS has become an attractive approach to study protein-protein interaction because of its simplicity, low detection limits, low sample consumption and low detection limits.

In Chapter 3 and 4, we have reported the interaction between α-gap junction Cx43/Cx50 peptides with CaM by MALDI-MS. We also performed a quick test for the possibility of interaction between CaM and γ-gap junction Cx45 peptide (Cx45p) from the second half of its intracellular loop. As shown in Fig. 6.7, both apo-CaM in 1 mM EGTA and holo-CaM in 1 mM CaCl₂ form a complex with Cx45 peptide. But the intensity of the complex in 1 mM Ca²⁺ (Fig. 6.7A) is more than two folds higher than the one in 1 mM EGTA (Fig. 6.7B) if we set the intensity of free CaM in each spectrum as a reference. Moreover, although we only added peptide to CaM as 2 to 1 ratio, holo-CaM could form the complex with four peptides but apo-CaM can only form with three peptides with lower intensity. As shown in Fig. 6.8, the interaction between Cx45p and CaM was also confirmed by NMR HSQC spectrum in the presence of calcium. The dramatic chemical shift changes for most of the amino acids of CaM were observed in the overlay NMR HSQC spectrum. The data also suggests strong binding because of the slow exchange in the spectrum. However, the binding ratio between Cx45p and CaM is not consistent with MALDI-MS because the spectrum of the CaM stops changing after adding Cx45p at 1 to 1 ratio (Fig. 6.8). As indicated in Fig. 6.9, the Cx45p binds to both lobes of CaM because the average chemical shift perturbations were observed in the whole CaM. The conclusion of the interaction between CaM and
Cx45p in the absence of calcium is difficult to make because of the insolubility of the Cx45p in 5 mM EGTA. An NMR of apo-CaM binding domain study without addition of EGTA has to be explored because most of the peptides do not dissolve well in the high concentration of EGTA. The interaction between Cx45p γ-gap junction and Cx45 peptide in the intracellular region has been identified by both MALDI-MS and NMR spectrum.

6.5. Dimerization study of engineered proteins with grafted EF-hand motif.

It has been reported that 35% or more proteins in the cell are in oligomeric state [270]. And average population of the oligomeric proteins is tetramer according to the calculation [271]. However, the number of the protein in oligomeric state is quite low because of the difficulties to get crystal for the oligomeric proteins [272]. In addition, the size of oligomeric protein is often too big to get high definition solution structure information by NMR technique. MALDI-MS is a fast approach to study the oligomeric protein. MALDI-MS also could provide the structure information in atomic level using hydrogen-deuterium exchange mass spectrometry.

EF-hand motif has been reported for involving in dimerization of many calcium binding proteins [263, 273-275]. We have reported calcium independent oligomerization of stromal interaction molecule 1 (STIM1) by the grafting approach. STIM1 is responsible for activating the Ca\(^{2+}\) release-activated Ca\(^{2+}\) (CRAC) channel [276]. The isolated EF-hand form STIM1 was inserted into the host protein, domain 1 of CD2 (CD2.D1, read Chapter 2 and 5 for more information about the grafting approach). As shown in Fig. 6.10 and Fig. 6.11, in the presence of Calcium or EGTA, peaks of STIM1 in both dimer and trimer molecule weight range were observed in MALDI-MS. The intensity of dimer peak is several folds stronger than the trimer. However, in the control experiment, the oligomeric state of the host protein CD2.D1 also was observed in the MALDI-MS spectrum but with less abundance (Fig. 6.12&13). It is very difficult to get the trimer peak of CD2.D1 by scanning many times for both calcium and EGTA condition. Therefore, we conclude that inserted EF-Hand facilitates the dimer formation of CD2.D1 but
calcium does not play a role for the oligomerization (Fig. 6. 14A). The data obtained by cross-linking SDS-PAGE also confirmed the oligomeric stage of the grafted protein CD.STIM1.EF (Fig. 6. 14B).

6.6. **Determination of chemical PEGylation sites on protein based MRI contrast agent.**

For the therapeutic protein based drug, polyethylene glycol polymer (PEG) chains are often attached to lysine group of the protein drug through a covalent bond to reduce the immunogenicity from the host immune system. This process is called PEGylation. In our protein based MRI contrast agent, PEGylation also increases the protein solubility and the relativity. The determination of the PEGylation sites in the MRI contrast agent could provide the structural information for improvement of the drug design. HPLC-ESI-Nano spray was used in this study to locate the PEGylation sites of the contrast agent protein (7E15) with much less sample consumption and better isolation.

As shown in Fig. 6. 15, 7E15 were PEGylated by different PEG with various chain length from 0.6 kDa to 5 kDa. Like other protein modification regents working on the lysine residues, more than one PEGylation sites were observed for all PEG reactions (Fig. 6. 15) because most of the lysine residues in the target protein are exposed to the solvent. 7E15 PEGylation product was digested by Trypsin which cleaves the bond after lysine and arginine. The digestion in 0.05% SDS was taken at 37 °C for overnight. The digestion solution was injected into Nano spray ESI-MS after 0.22 µm filtration. The peptide fragments were separated by HPLC in a very low flow rate following ESI-MS detection. Later, the peptide precursors were selected to perform tandem MS in order to get the peptide sequence information. As shown in Fig. 6.16, only the peptide fragment sequence with underline can be detected in ESI-MS. Peptide fragments of 7E15 containing K66 and K91 (Marine blue in Cartoon, PDB code: 1hng) cannot be observed because 1) the fragment with PEG is difficult to be ionized in current condition; 2) the size of the fragment with PEG is over the detection limit; 3) the PEGylation sites in 7E15 blocks the possible Trypsin cleavage sites and reduces the enzyme activity. Therefore, K66 and K91 are two highly possible PEGylation sites in the MRI contrast agent protein 7E15. In the future plan, Lysines in 7E15 will be replaced by arginine to specify the PEGylation locations.
6.7. Summary

MALDI-MS is a powerful technique to study protein-protein interaction and modifications. Dansylation and IAEDANS labeling on CaM have been studied by MALDI-MS. In the physiological pH, not only N-termini of CaM can be dansylated, but also the lysine residues in CaM. Although dansylation requires deprotonated lysine side chains and lysine has high pka (~10.5), we cannot get CaM dansylation product from a single labeling at even lower pH (6.5) possibly due to the special local environment of certain lysine residues. The interaction of CaM and Cx45 gap junction peptide was also confirmed by both MALDI-MS and NMR spectrum. The oligomeric stage of engineered protein with grated EF-hand form STIM1 has also been demonstrated by MALDI-MS. The possible PEGylaiton sites on the protein based contrast agent 7E15 has also been suggested by ESI-MS.
Figure 6.1 Voyager-DE™ PRO workstation.

Figure is from Applied Biosystems.
1. Sample (A) is mixed with excess matrix (M) and dried on a MALDI plate.

2. Laser flash ionizes matrix molecules.

3. Sample molecules are ionized by proton transfer from matrix:

\[ \text{MH}^+ + A \rightarrow \text{M} + \text{AH}^+. \]

Figure 6.2 Desorption and Ionization processed in MALDI-MS.
Figure 6.3 MALDI-MS spectrum of CaM dansylation in various concentrations.

A, Cartoon of CaM structure with potential dansylation sites highlighted in light grey sphere. B, the CaM dansylation product with 1 mM initial protein concentration. C, the CaM dansylation product with 0.5 mM initial protein concentration. D, the CaM dansylation product with 0.25 mM initial protein concentration. E, the CaM dansylation product with 0.125 mM initial protein concentration.
Figure 6.4 MALDI-MS spectrum of CaM and variants dansylation at various pH.

Cartoon picture of CaM (A, 1-148 aa), N-CaM (B, 1-75 aa) and C-CaM (C, 76-148 aa) with potential dansylation sites highlighted in brown sticks. The dansylation product of CaM (D) and N-CaM (E) at pH 6.5. The dansylation product of CaM (G), N-CaM (H) and C-CaM (I) at pH 7.0. F, the spectrum of C-CaM without dansylation. C-CaMF stands for the major degradation fragment of C-CaM.
Figure 6. 5 MALDI-MS spectrum of CaM variant IAEDANS labeling with cysteine mutation or insertion.

Cartoon picture of CaM-Cys (A), T34C (B) and T110C (C) with IAEDANS labeling sites highlighted in pink color. MS spectrum of CaM-cys (D), T34C (E) and T110C (F). MS spectrum of CaM variant with IAEDANS dye modification for CaM-Cys (G), T34C (H) and T110C (I).
Figure 6.6 The effect of urea on CaM modification with IAEDANS dye.

A, MALDI-MS spectrum of reaction product without 6 M urea. B, MALDI-MS spectrum of reaction product with 6 M urea. The efficiency of CaM chemical labeling in denature condition is much higher than native solution.
Figure 6.7 MALDI-MS spectrum of CaM and C45 peptide complex.

The spectrum were obtained after 100 μM CaM with 200 μM Cx45p peptide was incubated in 10 mM Tris-HCl, 100 mM KCl (pH 7.4) at room temperature with either 1 mM CaCl₂ (A) or 1 mM EGTA (B). The MALDI-MS spectrum of Cx45 peptide from the intracellular loop was indicated in the inset of A. Both apo-CaM and holo-CaM forms complex with Cx45 peptide.
Figure 6.8 Monitoring the interaction between CaM and Cx45 peptides (Cx45p) by (1H, 15N)-HSQC spectroscopy. An overlay of HSQC spectra of holo-CaM (red) with the spectrum of the holo-CaM-Cx45p (green). The NMR spectrum was obtained with 200 µM N\textsuperscript{15}CaM in 5 mM CaCl\textsubscript{2}, 10 mM Bis-Tris, 5 mM MES, 0.1% NaN\textsubscript{3} at pH 6.5 and 20 µl 3 mM Cx45p dissolved in DD H\textsubscript{2}O.
Figure 6.9 Chemical shift perturbation in CaM induced by addition of 2-fold molar excess of Cx45p
A, The chemical shift change of the assigned CaM residues in $^1$H NMR spectrum. B, The chemical shift change of the assigned CaM residues in $^{15}$N NMR spectrum. C, The weight average chemical shift change ($\Delta \delta$) was calculated using Equation 2.5.
Figure 6. 10 MALDI-MS spectrum of CD2.STIM1.EF in the presence of CaCl2.

A, the monomer of CD2.STIM1.EF. B, the dimer of CD2.STIM1.EF. C, the trimer of CD2.STIM1.EF. Sample of 50 µM CD.STIM1.EF with 500 µM CaCl2 in 10 mM Tris, pH 7.4 was incubated at room temperature for 2 hours before spotting on the MALDI plate.
Figure 6. 11 MALDI-MS spectrum of CD2.STIM1.EF in the presence of EGTA.

A, the monomer of CD2.STIM1.EF. B, the dimer of CD2.STIM1.EF. C, the trimer of CD2.STIM1.EF. Sample of 50 µM CD.STIM1.EF with 500 µM EGTA in 10 mM Tris, pH 7.4 was incubated at room temperature for 2 hours before spotting on the MALDI plate.
Figure 6. MALDI-MS spectrum of CD2.D1 in the presence of CaCl₂.

A, the monomer of CD2.D1. B, the dimer of CD2.D1. C, the trimer of CD2.D1. Sample of 50 µM CD2.D1 with 500 µM CaCl₂ in 10 mM Tris, pH 7.4 was incubated at room temperature for 2 hours before spotting on the MALDI plate.
Figure 6.13 MALDI-MS spectrum of CD2.D1 in the presence of EGTA.

A, the monomer of CD2.D1. B, the dimer of CD2.D1. C, the trimer of CD2.D1. Sample of 50 µM CD.D1 with 500 µM EGTA in 10 mM Tris, pH 7.4 was incubated at room temperature for 2 hours before spotting on the MALDI plate.
Figure 6.14 Analysis of oligomer status of CD2.STIM1.EF by both MALDI-MS and crosslinking SDS-PAGE.

A, Semi-quantitative analysis of monomer and dimer population of CD2.STIM1.EF with additional calcium and EGTA in the MALDI-MS. B, SDS-PAGE result of CD2.STIM1.EF after crosslinking with additional calcium and EGTA (data from Dr. Yun huang). 100 μg proteins in 20 mM HEPES buffer (pH 7.5) with 0.2% (w/v) glutaraldehyde reacted at 37 °C for 10 min.
Figure 6. 15 MALDI-MS spectrum of MRI contrast agent 7E15 labeling with PEG.

The exposed lysine residues of 7E15 in solution were labeled with PEG0.6 (A, 0.6 kDa), PEG2.4 (B, 2.4 kDa) and PEG5 (C, 5 kDa).
7E15-2T calculated M.W. = 11290.65 Da

**GS** RDSGTVGAL GHGI\_INIPN FQMTDDIDEV
WERGSTLVA EFKR\_KMKPFL KSGAFEIDAN
GDL\_IKNLTR DDSGTYNVT\_YSTNGTRILN
KALDLRILE

Figure 6.16 The possible PEGylation sites determined by Nano spray ESI-MS.

Only the peptide fragment with underline can be observed in the mass spectrum. Red and green color letters are the amino acids forming the calcium binding pocket in 7E15. All the lysine residues are in purple color in the sequence. K66 and K91 (Marine blue in the cartoon) are the possible PEGylation sites.
<table>
<thead>
<tr>
<th>Compound</th>
<th>Common names</th>
<th>Solvent</th>
<th>Wavelength (nm)</th>
<th>Applications</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,5-dihydroxy benzoic acid</td>
<td>DHB Gentiamic acid</td>
<td>acetonitrile</td>
<td>266</td>
<td>Peptides Glycoproteins</td>
<td><img src="image1" alt="Structure" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>water</td>
<td>337</td>
<td>Nucleotides</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>methanol</td>
<td></td>
<td>Oligonucleotide</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>acetone</td>
<td></td>
<td>Oligosaccharides</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>chloroform</td>
<td>355</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,5-dimethoxy-4-hydroxycinnamic acid</td>
<td>sinapic acid sinapinic acid SA</td>
<td>acetonitrile</td>
<td>266</td>
<td>Peptides Proteins</td>
<td><img src="image2" alt="Structure" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>water</td>
<td>337</td>
<td>Lipids</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>acetone</td>
<td>355</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-cyano-4-hydroxycinnamic acid</td>
<td>CHCA</td>
<td>Acetonitrile</td>
<td>337</td>
<td>Peptides Lipids</td>
<td><img src="image3" alt="Structure" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water</td>
<td>337</td>
<td>Nucleotides</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ethanol</td>
<td>355</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>acetone</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7. MAJOR FINDING AND THE SIGNIFICANCE OF THIS DISSERTATION

The performed studies of this dissertation are divided in three major parts: 1) investigating CaM regulation of gap junction proteins, 2) identifying EF-hand viral calcium binding proteins using grafting approaches, and 3) probing protein modification by mass spectrometry.

7.1. Investigating CaM regulation of gap junction proteins

Connexin 50 (Cx50), a member of the α family of gap junction proteins expressed in the lens of the eye, has been shown to be essential for normal lens development. Working with Dr. Yubin Zhou in Yang’s lab, we have first reported a Ca\(^{2+}\)-dependent regulation of Cx43, and Cx44, the sheep ortholog of human Cx46 via the association of Ca\(^{2+}\)-CaM at the intracellular loop domain of alpha-family gap junction proteins [103, 104]. However, several major questions remain to be answered:

1. Is lens specific Cx50 also regulated by CaM via interacting at this cystosolic region of gap junction protein?
2. If yes, does CaM bind to these alpha-family gap junction proteins with different affinities?
3. What are the CaM binding modes to the Cx?
4. Can CaM interact with Cx43 via multiple regions? If yes, what is the possible working model?

Chapters 3 and 4 of this dissertation addressed the above-unanswered questions. First, we have predicted and identified a CaM-target sequence (CaMBD, residues 141-166) at the conserved second half of the intracellular loop region of Cx50. Using a peptide model encompassing this predicted CaM region, we have determined its strong ability to bind CaM using a variety of in vitro biophysical and electrophysiological approaches. NMR and fluorescence spectroscopy further revealed that the peptide stoichiometrically binds to Ca\(^{2+}\)/CaM with an affinity of ~5 nM. The binding of the peptide expanded the Ca\(^{2+}\) sensing range of CaM by increasing the Ca\(^{2+}\) affinity of the C-lobe of CaM, while decreasing the Ca\(^{2+}\) affinity of the N-lobe of CaM. In addition, we have shown that elevations in intracellular Ca\(^{2+}\) concentration affected a 95% decline in junctional conductance (g\(_j\)) of Cx50 in N2A cells that is likely
mediated by CaM, because inclusion of the CaM inhibitor, calmidazolium, prevented this Ca\(^{2+}\)-dependent decrease in \(g_j\). The direct involvement of the Cx50 CaMBD in this Ca\(^{2+}/\)CaM-dependent regulation was further demonstrated by inclusion of a synthetic peptide encompassing the CaMBD in both whole cell patch pipettes, which effectively prevented the intracellular Ca\(^{2+}\)-dependent decline in \(g_j\). Overall, these results demonstrate that the binding of Ca\(^{2+}/\)CaM to the intracellular loop of Cx50 is critical for mediating the Ca\(^{2+}\)-dependent inhibition of Cx50 gap junctions in the lens of the eye.

Second, we have further shown that CaM exhibits the differential affinities for Cx43, 44, and Cx50 at this conserved loop regions using peptide models. The observed CaM binding affinity order is consistent with their helical propensity at the CaM binding regions of the Cxs.

Third, we have identified additional CaM binding sites on the first half of the intracellular loop and the C-terminal domain of Cx43 by biophysical studies. Our results indicate that in the presence of Ca\(^{2+}\), synthesized Cx peptide fragments, Cx43-2p and Cx43-4p, encompassing predicted CaM binding regions are able to bind with high affinity to CaM using NMR spectroscopy. Unlike Cx43-3p, the suggested stoichiometry of these newly discovered peptides which bind with CaM is greater than one. The NMR RDC result suggests that the CaM binding mode of Cx43-2p to CaM is similar as CaMBD in SK channel in an extended mode.

Fourth, in addition, the conserved CaM binding region in the second half of the intracellular loop has also been identified as a gamma-class of gap junction Cx45.

Based on our studies on CaM, it directly interacts with Cx43 at three predicted regions including two cytosolic loop regions (Cx43-2p and Cx43-2p) and the C-terminal (Cx43-4p). Based upon dye transfer studies via w.t Cx43 and mutations at the predicted CaM binding region, and the electrophysiological studies using patch clamp and peptide competition, we propose the working model for CaM regulation of alpha-gap junction proteins shown in Fig. 4.11. The binding of CaM to the cytosolic region Cx43-3p is the major regulation site. Upon response to the cytosolic calcium increase, CaM binds to the cytosolic
loop to result in the conformational change of gap junction and close the channel. It is possible CaM using Cx43-2p region as an anchor to locate close to the regulation site Cx43-3p to allow fast response. Our observation of the direct interaction of Cx43-4p at the C-terminal domain using peptide model by NMR is also interesting given the knowledge that this region is involved in trafficking, assembly and regulations such as physorylation.

The significance and high impact of this research is multifaceted. First, our results elucidate the molecular level of regulation of Cx43 by multiple CaM targeting regions and provide a proposed model about how CaM regulate gap junction assemble, trafficking and gating. Second, a large number of residues in the Cxs mutated in human diseases reside at the highly conserved CaM binding site in our prediction. Our studies provide insights that define the critical cellular changes and molecular mechanisms contributing to human disease pathogenesis as part of an integrated molecular model for the calcium regulation of GJs. Third, our proposed1 model for Ca$^{2+}$-dependent regulation of the alpha-class of gap junction proteins could provide new therapies to manage and/or prevent human maturity onset cataract (Cx46 and Cx50) and heart diseases (Cx43) by introducing CaM mutant with altered Ca$^{2+}$ binding affinities. The model of CaM interacts with multiple CaM binding sites in Cx43 will facilitate understanding of mechanism underlying gap junction regulation, and thus provide a molecular basis of potential therapeutic intervention of heart diseases related to malfunction of gap junction.

7.2. Identification of continuous calcium binding sites by grafting approach.

We applied a powerful Calcium Pattern Search (CaPS) method to predict and identify continuous Ca$^{2+}$ binding sites, EF-hand motifs, from linear protein sequences in bacteria and viruses. We have applied a grafting approach to predict and identify EF-hand motifs from various biological systems. The calcium binding sites identified in nonstructural protein 1 (nsP1) of Sindbis virus and Poxvirus, streptococcal hemoprotein receptor (Shr) of *Streptococcus pyogenes* provides not only novel insight into the roles of Ca$^{2+}$ and Ca$^{2+}$-binding proteins in pathogens including bacteria and virus, but also a method to explore
and define the role of Ca$^{2+}$ in other biological systems. We have also shown that the metal binding properties and dimerization study of two EF-hand motifs isolated from Calbindin D9k implicates the key role of Ca$^{2+}$ in protein oligomerization which is important for biological function of protein complex. The developed algorithm and grafting approach provide a new way to understand the role of calcium in channels and other biological systems.

7.3. **The study of protein-protein interactions by mass spectrometry.**

MALDI-MS has been considered a “softly-soft” ionization mass spectrometry technique in the protein-protein interaction study. The properties of the soft ionization process make MALDI-MS an advanced technique to explore the protein complex without breaking its weak non-covalent bond. The interaction of CaM and its targeting peptides fragments from gap junction has been proved by MALDI-MS in a fast, efficient and costless manner. MALDI-MS also suggests a de novo application to study the oligomerization of calcium binding proteins, and a fast and accurate method for identification of protein modification sites. Functional groups with certain properties are often introduced to protein for protein-protein interaction and metal binding study through protein modification. In this study, MALDI-MS was used to improve the protein modification with dye labeling for the fluorescence study or with PEG for the protein based contrast agent.
PUBLICATIONS AND MANUSCRIPTS


MANUSCRIPT IN REVISION OR IN PREPARATION

Chen, Y., Zhuo, Y., Zou, J., Yang, J. J., "Explore multiple calmodulin binding sites in connexin43”.

Chen, Y., Zou, J., Yang, J. J., "Elucidation of a nova extracullular calcium binding site on gap junction protein connexin26.”

Zou, J., Chen, Y., Yang, J. J., "Identification of Calmodulin binding domain in gamma family gap junction protein connexin45”

Yang, W., Chen, Y., Zhou, Y., Zou, J., Huang, Y., Lurtz, M., Louis, C., Yang, J. J., "A diffusion NMR study on the CaM-Cx43 peptide complex indicates the direct binding of CaM at the gap junction entrance.”

REFERENCES


96. !!! INVALID CITATION !!!


269. Chiappetta, G., et al., *Dansyl-peptides matrix-assisted laser desorption/ionization mass spectrometric (MALDI-MS) and tandem mass spectrometric (MS/MS) features improve the liquid


