CALPAIN INHIBITOR DESIGN INSPIRED BY THE NATURAL INHIBITOR CALPASTATIN

by

Kristin Elizabeth Low

A thesis submitted to the graduate program in Biochemistry in the Department of Biomedical and Molecular Sciences in conformity with the requirements for the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
(November, 2014)

Copyright © Kristin Elizabeth Low, 2014
Abstract

Calpains are cysteine proteases that depend on $\text{Ca}^{2+}$ for activation. They have many roles in the cell; hence modifications of calpains or changes in their levels of activity can lead to altered physiological states. Thus, it is of great interest and medical relevance to design inhibitors that can specifically target calpain without interfering with the other numerous cellular cysteine proteases. The cytosolic and ubiquitous calpains-1 and -2 have a naturally occurring and specific inhibitor, calpastatin, which binds with high affinity to the active site but in such a way that the calpastatin itself is not cleaved. Crystal structures of the full-length calpain-2 in complex with calpastatin have aided the design of novel calpain-specific inhibitors. I have helped design cyclic peptides to mimic important conserved structural elements of calpastatin and have investigated their inhibition potency against calpain. Cyclic peptides and derivative peptidomimetics were able to inhibit cysteine proteases including calpain. Macrocycle and peptidomimetic products showed competitive, non-competitive, and mixed inhibition of calpain-2 with $K_i$ values in the micromolar range while conferring some specificity. In a second strategy inhibitors mimicking the calpastatin subdomain B $\alpha$-helical region were designed to be calpain-specific by binding outside of the common cysteine protease active site cleft but extend into it. Indeed, stabilized $\alpha$-helices synthesized from conserved calpastatin residues were found to inhibit calpain specifically and potently, with potential use as fluorescent activity-based probes. Previously reported allosteric inhibitors of calpain were investigated, and their ability to inhibit the calpain enzyme through binding to the PEF domain questioned. The compound PD150606 was able to inhibit the protease core of calpain, which lacks the PEF domain of the full enzyme, and the penta-peptide LSEAL was found to have no effect at all on calpain \textit{in vitro}. Thus it is now clear that binding to
hydrophobic grooves in the calpain PEF domains does not confer inhibition and the mechanism of non-competitive calpain inhibition is still unknown.
Co-Authorship

A description of the contributions of each author to the experimental work and manuscript preparation for each chapter is presented here.

Chapter 3: Rational design of calpain inhibitors based on calpastatin peptidomimetics.

This chapter was written as a research paper for subsequent submission once intellectual property has been protected. The authors are as follows: Kristin E. Low (1), Kevin Chen (1), Robert L. Campbell (1), Peter L. Davies (1), Spencer Ler (2), Andrei K. Yudin (2), Jennifer L. Hickey (3), Joanne Tan (2), Conor C. G. Scully (2), and Serge Zaretsky (2).

(1) Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON, K7L 3N6, Canada
(2) Department of Chemistry, University of Toronto, Toronto, ON, M5S 3H6, Canada
(3) Encycle Therapeutics Inc., 101 College Street, Suite 332, Toronto, Ontario M5G 1L7, Canada

Kristin E. Low assisted in the synthesis of 3rd generation compounds and performed modeling, docking simulation, and molecular dynamics simulations studies, enzyme activity assays, and kinetic analyses. Kevin Chen performed some of the inhibition screening and kinetic analyses, with supervision by Kristin E. Low. Serge Zaretsky planned and carried out or supervised synthesis of all compounds, planned library and SAR screening, and performed structural analysis of peptides. Spencer Ler
carried out synthesis of the 1st generation peptide library, with supervision by Serge Zaretsky. Joanne Tan performed re-analysis by NMR to determine compound class.

Jennifer L. Hickey assisted in the NMR re-analysis and assignment of peptide structures. Conor C. G. Scully guided initial docking attempts and the generation of computational structures of compounds. The manuscript was written by Kristin E. Low and Serge Zaretsky, with input and overall guidance from Andrei K. Yudin, Robert L. Campbell, and Peter L. Davies.

Chapter 4: Development of α–helical calpain probes by mimicking a natural protein-protein interaction.

This chapter was published in the Journal of the American Chemical Society.


Kristin E. Low modeled the stabilized helices based on the calpain-calpastatin structure in order to provide a foundation for inhibitor development with the retention of key residues and structural features. Kristin E. Low also designed the epoxide-linked peptide inhibitors to provide ideal linker lengths for efficient covalent attachment to the active site cysteine. Finally, Kristin E Low provided guidance on the protease activity assays needed to accurately assess inhibition. Hyunil Jo performed the cross-linker screen and “selection of fitness” screen. Nataline Meinhardt performed the crosslinking
of unpurified peptide, preparation of cross-linked peptides, CD spectroscopy, protease activity assays, kinetic analysis, determination of IC$_{50}$, and activity-based probe experiments. Yibing Wu performed the NMR spectroscopy. The manuscript was written by Nataline Meinhardt with input from me and other authors.

Chapter 5: Allosteric inhibitors of calpains: reevaluating inhibition by PD150606 and LSEAL


Kristin E. Low designed the experiments and performed all of the experimentation with the exception of the thermal shift assay, performed by Sarathy Karunan Partha. The manuscript was written by Kristin E. Low with input from Peter L. Davies and Robert L. Campbell.
Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Peter Davies. With his mentorship, he has pushed me to be a better researcher as well as a better person. Peter has always promoted the feeling of a scientific community as well as that of incredible support. He has helped to provide an atmosphere of hard work, critical thinking, and pursuit of scientific knowledge, while still encouraging us to be healthy and wear our bike helmets. It has been a fantastic learning experience to be a part of his research team.

I would also like to extend many thanks to Dr. Robert Campbell. From crystallography, to simulations, to drug design, to enzyme kinetics, to countless other computational problems, Rob has been an incredible resource and fantastic help even when we’re both learning the intricacies of the issues and programs involved. I could not have done much of this work without his expertise and enthusiasm. I will miss being able to pop in and ask him if he saw the latest news on Slashdot or having him beat me to it.

Thanks to all the members of the Davies Lab, past and present. Sherry Gauthier has provided tremendous help with technical support through the years, and has also become an important friend for emotional support. Dr. Rachel Hanna showed me the ropes with calpain from day one and I will never forget her help. The 6th floor of Botterell Hall needs to be thanked as well, including past students; we’re much like a big, happy family and I can’t imagine graduate school without it.
I would like to thank my committee members, Dr. John Allingham and Dr. Steven Smith, for their advice and guidance. Extra thanks go out to Dr. John Allingham for his help and support with in-house X-ray beam time and helpful discussions, and to Dr. Bruce Hill for helpful fluorescence and enzyme discussions. I would also like to thank Dr. Zongchao Jia, Dr. Laura van Staalduinen, Dr. Qilu Ye, and Dr. Frédérick Faucher for extensive help with X-ray beam time and complex crystallography problems. Diane Sommerfeld has also been a great help in ensuring graduate studies run smoothly for us students.

My parents, Edith and Richard Low, should be given due credit as well. Without their foundations of science, computers, and engineering, I could never be the researcher I am today. Whether it was instilled genetically or from their encouragement and enthusiasm, my interests lie where they do because of them. Thanks especially to Dad for the love and support through all these years. I hope I would have made Mom proud.

Finally, I would like to thank the rest of my family for encouragement and support, and all my friendships forged from Ottawa, Montréal, and Kingston.
Table of Contents

Abstract.............................................................................................................................................. ii
Co Authorship ...................................................................................................................................... iv
Acknowledgements .............................................................................................................................. vii
List of Figures...................................................................................................................................... xiv
List of Tables...................................................................................................................................... xvii
List of Abbreviations............................................................................................................................ xviii
Chapter 1 General Introduction ........................................................................................................... 1
Chapter 2 Literature Review .................................................................................................................. 2
  2.1 Calcium signaling .......................................................................................................................... 2
    2.1.1 Role for calcium signaling in the cell ....................................................................................... 3
    2.1.2 Regulation of calcium levels .................................................................................................... 4
    2.1.3 Calcium-binding proteins ....................................................................................................... 7
  2.2 Cysteine proteases ....................................................................................................................... 11
    2.2.1 Reaction mechanism .............................................................................................................. 12
    2.2.2 Predominant cysteine proteases ............................................................................................ 14
    2.2.3 Nomenclature of protease specificity ..................................................................................... 16
  2.3 Calpain family of proteases ......................................................................................................... 18
    2.3.1 Mammalian and non-mammalian calpain isoforms ............................................................... 19
    2.3.2 Physiological roles and pathological conditions ................................................................. 22
    2.3.3 Calpain substrates and cleavage preferences ....................................................................... 26
  2.4 Calpain structures ....................................................................................................................... 27
    2.4.1 Activation mechanism (Ca^{2+}-free vs. Ca^{2+}-bound structures) ....................................... 30
    2.4.2 Calcium requirement for activation \textit{in vitro} and \textit{in vivo} ....................................................... 37
    2.4.3 Autoproteolysis ....................................................................................................................... 38
    2.4.4 Heterodimerization and subunit dissociation ......................................................................... 39
    2.4.5 Using the protease core domains as a “mini-calpain” .......................................................... 39
  2.5 Calpastatin, the endogenous calpain inhibitor ............................................................................. 41
    2.5.1 Regulation of calpain activity .................................................................................................. 42
    2.5.2 Structure and calpain inhibition ............................................................................................. 43
    2.5.3 Stabilization of calpain .......................................................................................................... 44
2.6 Exogenous calpain protease inhibitors ................................................................. 47
  2.6.1 Active site-directed calpain inhibition ............................................................. 47
  2.6.2 Peptidomimetics or peptide-derived inhibitors ............................................... 52
  2.6.3 Non-peptidic small-molecule inhibitors .......................................................... 57
  2.6.4 Selectivity of inhibitors for calpain over other cysteine proteases ................. 58
  2.6.5 Theoretical and existing therapeutic uses of calpain inhibitors .................... 61
2.7 Research objectives ................................................................................................. 62

Chapter 3  Rational design of calpain inhibitors based on calpastatin peptidomimetics. 64
  3.1 Abstract .................................................................................................................. 64
  3.2 Introduction ............................................................................................................. 65
  3.3 Experimental Procedures ...................................................................................... 73
    3.3.1 Materials ........................................................................................................... 73
    3.3.2 Cyclic peptide and peptidomimetic synthesis. .................................................... 74
    3.3.3 RP-HPLC/MS .................................................................................................. 75
    3.3.4 Peptide NMR .................................................................................................... 75
    3.3.5 Modelling, docking simulation, and molecular dynamics simulation studies... 76
    3.3.6 Peptide library screening with calpain and inhibition kinetics in a fluorescence-
          based hydrolysis assay ......................................................................................... 77
    3.3.7 Inhibition of papain and cathepsins L in a fluorescence-based hydrolysis assay
          ............................................................................................................................ 78
    3.3.8 Analysis of enzyme kinetics with lead compounds ......................................... 78
  3.4 Results and Discussion .......................................................................................... 79
    3.4.1 Modelling potential inhibitors on the β-turn structure of calpastatin in the
          calpain active site cleft. ........................................................................................ 79
    3.4.2 Macrocyclic and amidine-bearing peptidomimetic library derived from
          macrocyclization of peptides .............................................................................. 81
    3.4.3 Screening of peptide library using fluorescence-base protease assays ........ 82
    3.4.4 Enzyme kinetics analyses and inhibition mechanisms ................................... 86
    3.4.5 Molecular docking of cPGALK NMR structures to calpain-2 .................... 91
    3.4.6 Improvement of potency and selectivity of the peptidomimetic amPGLdO .... 94
  3.5 Conclusion .............................................................................................................. 98
Chapter 4  Development of α-helical calpain probes by mimicking a natural protein-protein interaction

4.1 Abstract .................................................................................................................. 100
4.2 Introduction .......................................................................................................... 101
4.3 Results and Discussion ....................................................................................... 104
  4.3.1 Design of template-constrained cyclic peptides stabilizing an α-helix
       conformation .......................................................................................................... 104
  4.3.2 Application of i, i+4 m-xylene crosslinker-based stabilization for calpain
       inhibitor design ..................................................................................................... 116
4.4 Conclusions ......................................................................................................... 129
4.5 Materials and Methods ...................................................................................... 130
  4.5.1 Crosslinker screen .......................................................................................... 130
  4.5.2 “Selection of the fitness” Screen ..................................................................... 131
  4.5.3 Crosslinking with the unpurified peptide ....................................................... 131
  4.5.4 Preparation of crosslinked peptides 3c from model peptide 3c by solid-phase
       peptide crosslinking ............................................................................................. 131
  4.5.5 CD spectroscopy .............................................................................................. 132
  4.5.6 NMR spectroscopy .......................................................................................... 133
  4.5.7 Protease activity assays .................................................................................... 133
  4.5.8 Kinetic analysis of calpain-1 by 3c ................................................................. 134
  4.5.9 Determination of IC\textsubscript{50} against enzymes ........................................ 135
  4.5.10 Activity based probe linker experiments ...................................................... 136
  4.5.11 ABP labeling experiments ............................................................................ 136

Chapter 5  Allosteric inhibitors of calpains: reevaluating inhibition by PD150606 and
LSEAL .......................................................................................................................... 138
5.1 Abstract ................................................................................................................ 138
5.2 Introduction .......................................................................................................... 139
5.3 Materials and Methods ...................................................................................... 142
  5.3.1 Materials ......................................................................................................... 142
  5.3.2 (EDANS)-EPLFAERK-(DABCYL) hydrolysis assay ..................................... 142
  5.3.3 Calpain autolysis/proteolysis assays .............................................................. 143
  5.3.4 Casein hydrolysis assay .................................................................................. 144
5.3.5 Thermal shift in a differential light scattering assay .......................................................... 144
5.4 Results ........................................................................................................................................ 145
  5.4.1 PD150606 inhibits the protease core of calpain-1 .............................................................. 145
  5.4.2 LSEAL inhibits neither the protease core of calpain-1 nor full-length calpain-2. ...... 147
  5.4.3 PD150606 inhibits calpain autoproteolysis by a direct effect on the protease core. .......... 147
  5.4.4 PD150606 is a relatively weak calpain inhibitor as shown by the fixed-point casein hydrolysis assay ........................................................................................................ 149
  5.4.5 PD150606 binding stabilizes full-length calpain-2 but not the calpain-1 protease core .................................................................................................................................. 151
  5.4.6 PD150606 does not follow traditional Michaelis-Menten kinetics with full-length calpain-2 .......................................................................................................................... 153
  5.4.7 PD150606 appears to inhibit the protease core of calpain-1 in a noncompetitive manner ......................................................................................................................... 157
5.5 Discussion .................................................................................................................................... 157
  5.5.1 Calpastatin binds to PEF domains only to anchor its inhibitory B peptide .... 157
  5.5.2 PD150606 does not induce changes in structure or dynamics of calpain .... 158
  5.5.3 Thermal shift assays are neutral with the protease core ................................................. 159
  5.5.4 Mode of inhibition by PD150606 is not clear ................................................................. 160
  5.5.5 LSEAL does not inhibit calpain and shows no affinity for PEF domains ... ... 161
5.6 Conclusion ................................................................................................................................. 161
5.7 Acknowledgements .................................................................................................................. 162
Chapter 6 General Discussion ........................................................................................................ 164
  6.1 Monitoring and assessing calpain activity using enzyme kinetics .................................... 166
  6.2 Calpastatin-based peptidomimetic calpain inhibitors ...................................................... 170
    6.2.1 Exploiting the conserved α-helix of calpastatin subdomain B .................................... 170
    6.2.2 Cyclic peptide mimics of calpastatin secondary structure - complications and binding site determination .............................................................. 172
  6.3 The possibility and significance of calpain allosteric inhibition ...................................... 177
  6.4 Potential “activators” of calpain and inhibition of calpain autoproteolysis ..................... 179
  6.5 Future development of calpain inhibitors ............................................................................ 180
References.................................................................................................................................................. 183

Appendix A Supplementary Information for Chapter 4: Development of α-helical calpain probes by mimicking a natural protein-protein interaction.......................................................... 209

Appendix B Kinetic analyses of the calpain-specific substrate (EDANS)-EPLFAERK-(DABCYL) ........................................................................................................................................... 250

Appendix C The potential for “activation” of calpain via inhibition of autoproteolysis... 256

Appendix D Crystallography of whole calpain bound to inhibitors ...................................................... 263

Appendix E Determination of the inhibitor binding site via photoreactive labeling ....... 268
List of Figures

Figure 2.1: EF-hand motif from calmodulin and coordination of Ca\(^{2+}\). ........................................... 9
Figure 2.2: Catalytic mechanism of cysteine proteases ................................................................. 13
Figure 2.3: Standard protease substrate and subsite nomenclature ............................................... 17
Figure 2.4: Members of the human calpain family ........................................................................... 20
Figure 2.5: Crystal structure of the apo-form of rat calpain-2 ......................................................... 31
Figure 2.6: The protease cores of apo calpain-2 and holo calpain-1 ............................................... 33
Figure 2.7: Crystal structure of Ca\(^{2+}\)-bound rat calpain-2 ............................................................ 35
Figure 2.8: Calpain inhibitor structures in the active site cleft ........................................................ 40
Figure 2.9: Structure of calpastatin-bound Ca\(^{2+}\)-activated rat calpain-2 ........................................ 44
Figure 2.10: Common reversible and irreversible warheads for covalent reaction with calpain .......................................................... 49
Figure 2.11: The warhead and address region of peptidomimetic inhibitors .................................... 51
Figure 2.12: Peptidomimetic calpain inhibitors ............................................................................. 53
Figure 2.13: Comparison of active site clefts between calpain, papain, and cathepsin .......... 59
Figure 3.1: Mimicking the active site of calpastatin-bound calpain-2 with cyclic peptides ......... 67
Figure 3.2: Macrocyclization of peptides to mimic natural calpastatin \(\beta\)-turn motifs .................. 71
Figure 3.3: Inhibition screening of peptide compounds ............................................................... 83
Figure 3.4: Comparison of the active site region of cathepsin-L and calpain-2 ......................... 85
Figure 3.5: Michaelis-Menten curves for the inhibition of calpain-2 by amPLKG, cPGALK, amPGLdO, and cPGSGO .......................................................... 88
Figure 3.6: Representative lowest binding energy docking pose of cPGALK docked to calpain-2 .......................................................... 93
Figure 3.7: Inhibition of cysteine proteases by variants of amPGLdO ........................................... 95
Figure 4.1: X-ray crystal structure of the calpain 2-calpastain complex ....................................... 103
Figure 4.2: Conformational restriction via crosslinking ............................................................... 106
Figure 4.3: Helix stabilization via screening of 24 crosslinkers ..................................................... 109
Figure 4.4: CD spectra of the model peptide and the crosslinked peptides in phosphate buffer [50 mM, pH = 7.0, 25 °C] ....................................................................... 112
Figure 4.5: NMR of m-xylyl c15-constrained cyclic peptide. ...................................................... 115
Figure 4.6: Sequence of double cysteine mutants (3a, 3b, and 3c) and their crosslinked counterparts (3a, 3b, and 3c). ..................................................................................................................... 117
Figure 4.7: CD spectra of uncrosslinked peptides 3a-c and crosslinked peptides 3a-c. [-125 µM peptide, 50 mM Tris (pH 7.5), 40% TFE]. ...................................................................................... 118
Figure 4.8: Lineweaver-Burke analysis shows that calpain inhibitor 3c to be a competitive inhibitor. ................................................................................................................................. 121
Figure 4.9: Design of a calpain-specific ABP ........................................................................................................... 125
Figure 4.10: FITC-NM-02 as a calpain specific ABP ................................................................................................. 128
Figure 5.1: Inhibition of calpain-2 and calpain-1 protease core. ............................................................... 146
Figure 5.2: The effect of PD150606 inhibition on calpain-2 autoproteolysis and calpain-1 protease core cleavage of inactive C105S calpain-2. ................................................................. 148
Figure 5.3: The inhibition of calpain-2 and calpain-1 protease core in a fixed-point casein hydrolysis assay..................................................................................................................... 150
Figure 5.4: Differential binding of PD150606 to apo- and Ca^{2+}-bound calpain-2. ............... 154
Figure 5.5: Enzyme kinetics curves for calpain-2 and calpain-1 protease core inhibited by PD150606. ................................................................................................................................................ 155
Figure A.1: MALDI spectra of low concentration linker with the model peptide. ........... 219
Figure A.2: Crude HPLC profile (Low Concentration Reaction with the model peptide). .................................................................................................................................................. 217
Figure A.3: Crude HPLC profile – “selection of the fittest” reaction................................. 223
Figure A.4: Characterization of helix formation in the peptide by NMR spectroscopy... 225
Figure A.5: CD spectrum of uncrosslinked peptide 3a-c in Tris buffer (50 mM, pH = 7.5). .................................................................................................................................................. 232
Figure A.6: CD spectrum of crosslinked peptide 3a-c in Tris buffer (50 mM, pH = 7.5). .................................................................................................................................................. 233
Figure A.7: Representative calpain activity assay progress curve................................. 236
Figure A.8: Michaelis-Menten curve for determining K_m of calpain substrate. ......... 238
Figure A.9: IC_{50} Curves for enzyme assays ......................................................................................... 243
Figure A.10: Michaelis-Menten plot of initial velocities at different substrate and inhibitor concentrations ....................................................................................................................... 242
Figure A.11: IC_{50} Curves for enzyme assays using inhibitor 3c ........................................... 245
Figure B.1: Fluorescence of increasing concentrations of free EDANS fluorophore in solution................................................................. 252
Figure C.1: Activity of calpain-2 in the presence of cPGLGF and B27 peptides........ 258
Figure E.1: Potential binding sites for amPGLdO on calpain-2............................................. 270
List of Tables

Table 2.1: Residue preference of the calpain-1 protease core for each position in a peptide substrate ............................................................................................................. 28
Table 3.1: Summary of inhibition kinetics ......................................................................................................................................................................................... 89
Table 4.1: $K_i$ against calpain-1 ........................................................................................................................................................................................................... 120
Table 4.2: The $K_i$ of crosslinked inhibitor 3c against other papain family proteases. .... 123
Table 5.1: The effect of PD150606 and LSEAL on calpain-2 and calpain-1 protease core aggregation temperatures .......................................................................................................................................................................................... 152
Table A.1: Calculated and observed model peptide masses. ........................................... 212
Table A.2: Alanine Scanning Mutagenesis. .................................................................................................................................................................................................................................................. 228
Table A.3: Uncrosslinked and crosslinked calpastatin fragment peptides. ................. 230
Table A.4: $V_{\text{max}}^{\text{app}}$ and $K_m^{\text{app}}$ values obtained from the above Michaelis-Menten plot... 243
Table A.5: Calculated and Observed Masses for Calpain Activity Based Probes. ......... 247
Table B.1: Inner filter effect corrections for (EDANS)-EPLFAERK-(DABCYL) calpain substrate. ................................................................................................................................. 253
Table C.1: Rate constants comparing product formation and autoproteolysis during calpain “activation” .......................................................................................................................... 261
Table D.1: Attempted crystallization screens of calpain and calpain-inhibitor complexes. ........................................................................................................................................ 266

xvii
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABP</td>
<td>activity-based probe</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>ATA</td>
<td>aurintricarboxylic acid</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>ATPase</td>
<td>adenosine triphosphatase</td>
</tr>
<tr>
<td>C2L</td>
<td>C2-like</td>
</tr>
<tr>
<td>CAPNS1-/-</td>
<td>calpain-1 small subunit knockout</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>CFP</td>
<td>cyan fluorescent protein</td>
</tr>
<tr>
<td>CPP</td>
<td>cell penetrating peptide</td>
</tr>
<tr>
<td>DABCYL</td>
<td>4-((4-(dimethylamino)phenyl)azo)benzoic acid</td>
</tr>
<tr>
<td>DCM</td>
<td>dichloromethane</td>
</tr>
<tr>
<td>DIPEA</td>
<td>diisopropylethylamine</td>
</tr>
<tr>
<td>DMF</td>
<td>α,α’-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DSF</td>
<td>differential scanning fluorimetry</td>
</tr>
<tr>
<td>DSLS</td>
<td>differential static light scattering</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E-64</td>
<td>trans-epoxysuccinyl-L-leucylamido (4-guanidino)butane</td>
</tr>
<tr>
<td>EDANS</td>
<td>5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid</td>
</tr>
<tr>
<td>EDT</td>
<td>1,2-ethanedithiol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>EGF</td>
<td>epidermal growth factor</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethyleneglycoltetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>ERK</td>
<td>extracellular signal-regulated kinases</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>FRET</td>
<td>fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>GR</td>
<td>glycine-rich</td>
</tr>
<tr>
<td>HBTU</td>
<td>O-(benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate</td>
</tr>
<tr>
<td>HCL</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>HCTU</td>
<td>O-({1H-6-chlorobenzotriazol-1-yl})-N,N,N',N'-tetramethyluronium hexafluorophosphate</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>HPLC</td>
<td>high-performance liquid chromatography</td>
</tr>
<tr>
<td>HRP</td>
<td>horseradish peroxidase</td>
</tr>
<tr>
<td>IC₅₀</td>
<td>concentration at which 50% of the enzyme is inhibited</td>
</tr>
<tr>
<td>IFE</td>
<td>inner filter effect</td>
</tr>
<tr>
<td>INFIT</td>
<td>inverse Fourier transformation of in-phase multiplets</td>
</tr>
<tr>
<td>IP₃R</td>
<td>inositol triphosphate receptor</td>
</tr>
<tr>
<td>ITD</td>
<td>isothermal denaturation</td>
</tr>
<tr>
<td>Kᵢ</td>
<td>inhibition constant</td>
</tr>
<tr>
<td>Kₘ</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>matrix-assisted laser desorption/ionization-time of flight</td>
</tr>
<tr>
<td>MAPK</td>
<td>mitogen-activated protein kinase</td>
</tr>
<tr>
<td>mEF</td>
<td>mouse embryo fibroblast</td>
</tr>
<tr>
<td>Mmt</td>
<td>methoxytrityl</td>
</tr>
<tr>
<td>MPT</td>
<td>mitochondrial permeability transition</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometer</td>
</tr>
<tr>
<td>N</td>
<td>Ala(2-naphtyl)</td>
</tr>
<tr>
<td>NMDA</td>
<td>N-methyl-D-aspartic acid</td>
</tr>
<tr>
<td>NMP</td>
<td>N-methylpyrrolidone</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>PC1/2</td>
<td>protease core 1/2</td>
</tr>
<tr>
<td>PDB</td>
<td>protein database</td>
</tr>
<tr>
<td>PEF</td>
<td>penta EF-hand</td>
</tr>
<tr>
<td>pfF</td>
<td>pentfluorophenylalanine</td>
</tr>
<tr>
<td>PMCA</td>
<td>plasma membrane calcium-dependent ATPase</td>
</tr>
<tr>
<td>PVDF</td>
<td>polyvinyl difluoride</td>
</tr>
<tr>
<td>RCM</td>
<td>ring closing metathesis</td>
</tr>
<tr>
<td>RFU</td>
<td>relative fluorescence units</td>
</tr>
<tr>
<td>RMSD</td>
<td>root-mean-square deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate – polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SER</td>
<td>sarcoendoplasmic reticulum</td>
</tr>
<tr>
<td>SERCA</td>
<td>sarco/endoplasmic reticulum Ca^{2+} ATPase</td>
</tr>
<tr>
<td>SLY-MCA</td>
<td>succinyl-Leu-Tyr-7-amido-4-methylcoumarin</td>
</tr>
<tr>
<td>SPR</td>
<td>surface plasmon resonance</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SR</td>
<td>sarcoplasmic reticulum</td>
</tr>
<tr>
<td>STD-NMR</td>
<td>saturation-transfer difference NMR</td>
</tr>
<tr>
<td>Succ-LLVY-AMC</td>
<td>N-succinyl-Leu-Leu-Val-Tyr-7-amido-4-methylcoumarin</td>
</tr>
<tr>
<td>TBI</td>
<td>traumatic brain injury</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>TFE</td>
<td>trifluoroethanol</td>
</tr>
<tr>
<td>TIPS</td>
<td>triisopropylsilane</td>
</tr>
<tr>
<td>TRIS</td>
<td>tris(hydroxymethyl) aminomethane</td>
</tr>
<tr>
<td>Trt</td>
<td>trityl</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
<tr>
<td>YFP</td>
<td>yellow fluorescent protein</td>
</tr>
<tr>
<td>Z-FR-AMC</td>
<td>benzyloxycarbonyl-Phe-Arg-7-amino-4-methylcoumarin</td>
</tr>
</tbody>
</table>
Chapter 1

General Introduction

In order to understand calpain activation and the issues involved in calpain activity assays, one needs to understand calcium signaling in the cell and how the various cellular calcium-binding proteins interplay to regulate calcium levels. This thesis begins by summarizing calcium signaling in the cell, regulation of calcium levels, and the basics of calcium-binding proteins. Introducing cysteine proteases next and their mechanism of action provides the background for discussing calpain, its mechanism of activation in relation to other similar cysteine proteases, and the importance of calpain in mammalian cell physiological processes and associated pathological conditions. Calpain structures have been intrinsic in determining the activation mechanism of calpains, as well as informing about calcium requirement, autoproteolysis, heterodimerization, and the functioning of the protease core domains. The endogenous calpain-specific inhibitor calpastatin is then introduced in the context of calpain structures solved so far to further lead into exogenous calpain protease inhibitors and their mechanisms of action, while the importance of calpain inhibitors for in vitro tools and in vivo therapeutics is elaborated upon. Finally, the main research objectives are summarized in order to relate exogenous inhibitor design back to the calpastatin mode of calpain inhibition, and provide an introduction for the subsequent research chapters.
Chapter 2

Literature Review

2.1 Calcium signaling

The use for intra- and inter-cellular signaling arose as multi-cellular organisms evolved to be increasingly complex, involving many interactions between cells and proteins, as well as compartmentalization of cell contents. Many different molecules are used for such signaling, including proteins for complex signal transduction pathways, and charged atoms and small effector molecules for the more basic and intrinsic pathways. One of the most common mechanisms has been shown to involve calcium and phosphate ions as the major players, allowing the cells to adapt to their environment, and importantly enabling the modification of the response based on concentration of the signaling molecule[1].

Due to the charged nature and chemical properties of calcium, it is highly beneficial to use as a signaling molecule in the cell. The cellular calcium concentrations between internal and external locales range over four orders of magnitude from ~100 nM
to mM, respectively. The positive charge prevents Ca$^{2+}$ from moving across the plasma membrane, while it binds water much less readily than other more ubiquitous cations such as Mg$^{2+}$. Most importantly, calcium has a larger atomic radius than magnesium, allowing for more flexibility in coordination by protein residues. Thus, proteins can bind calcium preferentially over magnesium with a wide range of binding constants. In addition, contrary to more complex signaling molecules, calcium cannot be chemically altered. The formation of insoluble salts with such ions as phosphate, an important molecule for major cellular processes and signaling, is another hugely problematic property of calcium. For these reasons, calcium must be sequestered from the cytosol into such organelles as the sarcoendoplasmic reticulum (SER) and mitochondria, and its cellular concentrations tightly regulated. Specific control over localized calcium allows the cell to remove a potentially dangerous ion from the cytoplasm and use it as an important signaling molecule.

2.1.1 Role for calcium signaling in the cell

Through activating downstream proteins via calcium binding, this simple cation can function as a chemical relay from external sources or internal repositories to cytosolic components, eliciting an appropriate cellular response. Calcium is used in many key intracellular pathways including muscle contraction[2], cell migration[3], neurotransmitter release[4], and cell cycle regulation[5], among others. As a secondary
messenger, calcium can interact with proteins triggered by first messengers for three main functions: storing calcium, transporting calcium across membranes, and sensing and modulating calcium signals.

2.1.2 Regulation of calcium levels

Calcium repositories in the body consist mostly of bone, where 99.5-99.9% of the entire body's calcium content is found[6]. The remaining 0.1-0.5%, equating approximately 100 mmol, is distributed evenly between intracellular and extracellular compartments. On a cellular level, calcium differs greatly between the cytosol and the extracellular space with the cytosol concentrations being 3-4 magnitudes less than the low mM concentrations found outside the cell. This large difference results in an easy and spontaneous influx of Ca\(^{2+}\) into the cytosol across the plasma membrane and from internal sources including sarco/endoplasmic reticulum and mitochondrial stores. Calcium is quickly taken up into these organelle stores to remove the ion from the cytoplasm, but any movement of calcium is highly regulated and tightly controlled by varying ion transport channels to allow for an appropriate and localized response to stimuli.

Sodium/calcium exchangers are the main transport proteins responsible for efflux of calcium from the cytoplasm in order to restore low resting calcium concentrations.
They are, however, reversible and are thus involved in allowing controlled import of calcium[7]. \(\text{Na}^+\text{Ca}^{2+}\) exchangers function across plasma membranes or intracellular compartments, whereby one \(\text{Ca}^{2+}\) is exported for the import of three \(\text{Na}^+\) ions. Binding of calcium ions to the cytosolic portion of the protein activates transport while contrarily the binding of sodium will inhibit ion transport. Two distinct calcium-binding domains, each containing a pair of EF-hand loops, of the \(\text{Na}^+\text{Ca}^{2+}\) exchangers have different binding affinities for calcium, allowing for fine-tuning of calcium sensing, responding to a wide range of calcium concentrations[8].

Another transport protein responsible for movement of calcium across the plasma membrane, which is a major player in calcium export from the cytosol, is the P-class plasma membrane calcium-dependent ATPase (PMCA). To account for the various mammalian cellular responses and needs, there are four distinct PMCA genes while PMCA-like genes have been found in other eukaryotic cell systems. One ATP is hydrolyzed to export one molecule of \(\text{Ca}^{2+}\). PMCA pumps have low affinities for calcium with \(K_d\) values in the range of 10 \(\mu\)M, and are thus inactive at physiological cytosolic calcium concentrations. Many mechanisms are responsible for modifying the PMCA pumps in order to increase affinity to calcium to \(K_d\) values of around 100-200 nM including binding and activation by calmodulin, interaction with the surrounding membrane phospholipids, phosphorylation by protein kinase A, and cleavage by proteases such as calpain[9]. The \(\text{Na}^+\text{Ca}^{2+}\) exchangers function at low affinity but high capacity while the PMCAs function at high affinity but low capacity. This allows for the
main two plasma membrane calcium pumps to complement each other; Na\(^+\)/Ca\(^{2+}\) exchangers are mostly responsible for rapid adjustments needed while PMCA pumps function to maintain constant cytosolic calcium concentrations. Many other calcium channels that span the plasma membrane help to adjust calcium levels in the cytoplasm including voltage-gated Ca\(^{2+}\)-selective channels and transient receptor potential ion channels.

Sarco/endoplasmic reticulum Ca\(^{2+}\) ATPase (SERCA) pumps are responsible for transport of calcium to the sarcoplasmic reticulum. Functionally and structurally, SERCA pumps are quite similar to PMCA pumps. There are five main isoforms of SERCA pumps, each being tissue specific and with distinct functions. They are the main components responsible for calcium uptake to the ER in skeletal muscle[10]. In order for muscle relaxation following muscle contraction, calcium must be pumped back into the SR for storage and to terminate the calcium signal. SERCA pumps are efficient at maintaining this concentration difference with two Ca\(^{2+}\) molecules being exported via high affinity calcium-binding sites for every molecule of ATP hydrolyzed[11]. In addition to restoring resting calcium levels after muscle contraction, there is a need to pump back the calcium that naturally leaks out of the SR into the cytoplasm. If SERCA pumps are blocked and/or PMCAs continually export calcium from the cell faster than the SERCA pumps can restore calcium to the SR, store-operated Ca\(^{2+}\) entry occurs mediated by IP\(_3\) receptors (IP\(_3\)R) and Ca\(^{2+}\) enters back into the SR[1]. IP\(_3\)R-mediated Ca\(^{2+}\) flux has been
shown to have a much more prominent role in non-muscle cells for calcium flux into the ER[12].

Mitochondria are important in coupling cellular respiration and oxidative phosphorylation with calcium signaling. Although calcium transport across mitochondrial membranes is poorly understood, it is known that calcium uniporters are responsible for the uptake of Ca\(^{2+}\) into the mitochondria, while antiporters are responsible for the efflux[13]. K\(^+\) membrane potentials drive the uniporter uptake of calcium into the mitochondria, and other associated proteins have been shown to modulate this uptake. Antiporters transport calcium back out of the mitochondria, and this has been shown to be dependent on Na\(^+\) and H\(^+\) membrane potentials. A large pore in the inner mitochondrial membrane, called the mitochondrial permeability transition (MPT), also allows for calcium flow the permeability of which can be increased by fatty acid association.

### 2.1.3 Calcium-binding proteins

After deposition in the cytoplasm or organelles from the various transporters mentioned previously, calcium can enact downstream cellular responses through interaction with calcium-binding proteins. Many of the main calcium-binding players are localized to membranes close to calcium stores in order to be involved in signal...
transduction without overloading calcium across the entire cytosol. This allows for localized increases in calcium to the concentrations needed for signal propagation without detriment to cellular function.

Ca\(^{2+}\)-buffering proteins are required to maintain the necessary physiological levels of calcium in the cytosol and lumen of intracellular storage compartments. SR and ER have been shown to be responsible for calcium stores in skeletal muscle and non-muscle cells, respectively. The muscle-cell specific protein calsequestrin is responsible for binding and sequestering calcium in the SR lumen, while the ubiquitously expressed protein calreticulin sequesters calcium in the lumen of both the ER and SR\(^\text{[14,15]}\). Both proteins have high binding capacities for Ca\(^{2+}\); calsequestrin can bind 40-50 Ca\(^{2+}\) with low affinity\(^\text{[16,17]}\), while calreticulin binds 20-30 molecules of Ca\(^{2+}\) with high affinity\(^\text{[18]}\). The sequestration of calcium inside storage organelles helps to decrease soluble (free) calcium levels, which in turn helps the pumps not to have to transfer Ca\(^{2+}\) against a high gradient.

Intracellular transduction of calcium signals is achieved most commonly via EF-hand motifs, which are helix-loop-helix structural motifs that preferentially bind calcium (6-8 coordination residues) over other more abundant divalent cations like magnesium (6 coordination residues) (Figure 2.1). Through interaction with side chain carboxylates and
Figure 2.1: EF-hand motif from calmodulin and coordination of Ca$^{2+}$. Crystal structures of the N-terminal EF-hand motifs of calmodulin in the (A) apo (green) (PDB: 1CFD) and (B) calcium-bound (cyan) forms (PDB: 1CLL) are shown in cartoon representation. Calcium ions are shown as grey spheres. (C) Calcium coordination in the calmodulin EF-hand shown in (B). Residues responsible for coordination are labeled, shown in stick representation, and coloured by atom-type (red – oxygen, blue – nitrogen, cyan – carbon). Coordinating water is also shown as a red sphere. Interactions between Ca$^{2+}$ and each of the seven coordinating groups are shown as dashed lines.
protein backbone carbonyls in a loop straddled by two alpha-helices, calcium is coordinated usually by six residues[19,20]. Upon ligand binding, EF-hand motifs can confer a conformational change on a host protein like calmodulin, thereby inducing a signal as a structural response to the Ca\(^{2+}\) effector.

Arguably the most important calcium-binding proteins are the ones that propagate or elicit the response from calcium signaling: the messengers. Of the many calcium-dependent proteins in the cell, calmodulin is a prevalent messenger requiring calcium to bind to its many interactors and is often considered the prototypical calcium sensor[21]. It is also the major calcium-binding protein in non-muscle cells[22]. Calmodulin contains four EF-hand motifs, a pair in each of two globular domains, which induce large structural rearrangements upon calcium binding exposing the linker between the domains[23]. This area between N- and C-terminal domains is responsible for interaction with over 100 target proteins of immensely varied shapes and functions including kinases, phosphatases, nitric oxide synthases, and phosphodiesterases[24].

The cysteine protease calpain is another important Ca\(^{2+}\) sensor and is involved in the cleavage of downstream targets based on local high concentrations of calcium. This messenger contains 10 Ca\(^{2+}\) binding sites, 8 of which reside in EF-hand motifs of varying fidelity, and this calcium activation is unique among intracellular proteases. Active calpain is responsible for the controlled, limited proteolysis of many targets for a variety
of cellular functions: reorganization of cytoskeletal proteins in both cell differentiation and cell motility; modifications of proteins in signal transduction, necrosis, and apoptosis pathways; regulation of gene expression. Calpain is required for normal cellular functions. However, during unregulated calpain activity that follows a loss of calcium homeostasis many pathological conditions occur. Calpain and its inhibition, by both endogenous and exogenous inhibitors, is the main focus of my thesis.

2.2 Cysteine proteases

Proteases are enzymes that catalyze the hydrolysis of peptide bonds resulting in the irreversible modification of target proteins. They can either target the terminal ends of a polypeptide or cleave the protein at internal amide bonds. They can be either highly specific and cleave substrates at specific extended recognition sequences or more promiscuous, recognizing one or more amino acid types anywhere in the protein sequence. This allows for proteases to have a wide variety of roles in the cell, involving turnover of proteins and peptides or functional activation of a substrate in a signal transduction pathway.

There are eight groups of proteases, each named for the residue involved in the catalytic mechanism and classified by the MEROPS protease database: cysteine, serine, threonine, aspartic acid, glutamic acid, mixed, unknown, and metalloproteases[25].
Metalloproteases are the exception to the naming convention, as they use metal ions for activation of the water molecule instead of an amino acid. From these groups, proteases are further classified into superfamilies, or clans, based on structure and catalytic mechanism. As a result of convergent evolution, the protease superfamilies can have very different protein folds and yet use the same catalytic mechanism. Cysteine proteases consist of 14 different protease superfamilies, with one group of unclassified families. These clans are further divided into families denoted with a letter, identifying the catalytic nucleophile, and number. For example, calpains belong to clan CA and peptidase family C2, while bearing similarity to the papains in peptidase class C1.

2.2.1 Reaction mechanism

Cysteine proteases typically use the cysteine-histidine-asparagine catalytic triad to catalyze the cleavage of the peptide bond via nucleophilic attack at the targeted carbonyl moiety (Figure 2.2). Histidine deprotonates the active site cysteine to result in a strongly nucleophilic thiolate ion. The thiolate attacks the carbonyl carbon of the scissile bond, forming an acyl-enzyme intermediate, while the thiol-imidazolium ion pair is stabilized by the nearby neutral asparagine residue. A tetrahedral intermediate forms when the electron pair on the carbonyl oxygen moves from the double bond to the oxygen alone, and the transition intermediate is stabilized by the oxyanion hole, a pocket common among cysteine proteases. Collapse of the transition intermediate occurs when
Figure 2.2: Catalytic mechanism of cysteine proteases. The residue numbering shown are based on those from calpain-1, with curved lines representing the rest of the protein. The enzyme is coloured black, the peptide substrate is coloured blue, and the water molecule used for hydrolysis is coloured red. Hydrogen bonds are represented by dashed lines.
the oxygen electron pair reforms a double bond with the carbon and the catalytic histidine protonates the nitrogen, creating a leaving group with the C-terminal end of the scissile bond. The N-terminal end is released from the enzyme when water acts as a nucleophile to attack the carbonyl carbon, once again forming a tetrahedral intermediate that undergoes a similar collapse[26].

2.2.2 Predominant cysteine proteases

Papain-like cysteine proteases are the largest clan of cysteine proteases (clan CA), expressed widely across the animal and plant kingdoms while also found in some viruses and bacteria. They are generally between 20-35 kDa in size and the catalytic portion resembles the enzyme papain consisting of two domains, left (L-) and right (R-) [27,28]. The active sites of the enzymes are formed by the cleft between the two domains with the catalytic residues coming from both sides. The catalytic cysteine is found on a highly conserved α-helix on the L-domain and the histidine and asparagine residues reside in close proximity on the R-domain.

Cathepsins are a family of papain-like lysosomal proteases, abundantly expressed across a wide variety of cell types. There are approximately 12 cathepsins, differentiated by structure, substrates, and catalytic mechanism[27]. Although the majority of cathepsins are cysteine proteases, a few are serine or aspartyl proteases. Cathepsins, along with most proteases in the CA clan, are proenzymes in that the N-
terminal propeptide occupying the active site must be proteolytically removed for enzyme activation[29]. Most cathepsins function as endopeptidases however cathepsin B has been found to function as a dipeptidyl carboxypeptidase[30]. Cathepsin B is the most widely expressed cysteine cathepsin and typically functions as a housekeeping protein within the lysosome[31]. Cathepsins K and L are also cysteine-dependent, widely expressed cathepsins with papain-like catalytic domains.

Caspases are aspartate-specific cysteine proteases belonging to the CD protease clan of cysteine proteases, and are essential for apoptotic cell death. Recently, caspases have also been implicated in cell survival, proliferation, and differentiation, and inflammation response pathways[32]. Some caspases are proenzymes and require cleavage of the prodomain by other caspases or proteases. This removes the propeptide from the active site cleft and allows substrates to bind. While the structure of the active caspases seem to be common and tetrameric in nature, the structure of the prodomains can range from simple peptides to large, structured domains[33]. Although caspases do not belong to the papain-like cysteine protease family links between caspases and calpains have been seen in apoptosis and necrosis, including both direct and indirect cross-talk between cascading pathways[34].
2.2.3 Nomenclature of protease specificity

In 1967, Schechter and Berger devised a nomenclature system to describe the papain protease groove-like active site and the ability to accommodate different side-chains on either side of the cleaved peptide bond in specific enzyme subsites[35] (Figure 2.3). The system was primarily created to allow for discussion about the preference for specific residues at each subsite and the minimum length of substrate required for papain activity. This system was then ported over to other proteases with similar active sites such as cathepsins and calpains. The substrate PX (or unprimed) and PX’ (or primed) positions are found on the N-terminal and C-terminal sides of the peptide bond respectively, where X designates the sequential number of residues as distance from the scissile bond. Similarly, SX and SX’ are used to describe the respective protease subsites to accommodate PX and PX’.

This nomenclature system is generally used to describe substrate and inhibitor binding to a protease active site cleft, however it is somewhat limited in that the active site and enzyme-ligand interactions are over-simplified. Many inhibitors do not contain peptidic residues and are not as likely to fit into the designated positioning system. Despite this limitation, the system remains popular and is widely used for discussing the active site of proteases.
Figure 2.3: Standard protease substrate and subsite nomenclature. The protease substrate and subsites are labeled as per the system devised by Schetcher and Berger [35], with the N and C termini of the substrate displayed and the location of the scissile bond labeled appropriately. Substrate positions and subsites on the unprimed side or N-terminal side of the scissile bond are labeled in blue, while those on the primed side or C-terminal side are labeled in red.
2.3 Calpain family of proteases

The calpain enzyme, originally known as calcium-activated neutral protease, is named for the requirement of calcium for activation (cal-), as used in calmodulin, and the papain-like cysteine protease catalytic domains (-pain) [36]. Since its original discovery in 1964 by Guroff[37], calpain has been extensively studied and consists of a large family of proteases across animals, plants, fungi, yeast, and bacteria[38,39]. In humans alone, 15 calpain and calpain-associated genes have been identified, one of which (CAPN4) consists of a catalytically inactive subunit common to some active calpain enzymes[40].

Calpains are papain-like proteases and as such belong to the CA clan of cysteine proteases. They differ somewhat from standard papain-like enzymes. Calpains share the two proteolytic core domains common to other CA clan enzymes, however calpain is not a proenzyme; even though it may go through limited autoproteolysis, it does not contain a propeptide that needs to be cleaved for substrate to bind. Instead, calcium is required for activation and to align the catalytic triad residues and form the active site cleft. Calpains are defined by the presence of protease core domains and activation by calcium, however they may differ highly in their peripheral domains. This allows for a variety of physiological roles for calpains.
2.3.1 Mammalian and non-mammalian calpain isoforms

Members of the calpain family of proteases have been found across many mammalian and non-mammalian species[38,39,41]. Mammalian forms of calpain can be classified as both classical and non-classical calpain proteases (Figure 2.4). The classical or “typical” calpains are those that share domain structure with the 80 kDa large subunit of the ubiquitous calpain-1 and calpain-2 proteins, including the presence of the calcium-binding PEF-domain. They consist of the protease core domains PC1 and PC2, followed by the C2-like (C2L) and the penta EF-hand (PEF) domains. The PEF domain is responsible for dimerization with the calpain small subunit 1, containing a homologous PEF domain preceded by an N-terminal glycine-rich (GR) domain. However, some calpains, like calpain-3, homodimerize through their PEF domain as described below[42].

Several tissue-specific classical calpains contain a domain structure similar to calpains-1 and -2. These other classical calpains include: skeletal muscle calpain-3[43], gastric calpains-8 and -9, testicular calpain-11[44], hair follicle calpain-12[45], lung and testicular calpain-13[46], and the poorly-studied calpain-14[46]. Despite the presence of a PEF-domain, classical calpains are not guaranteed to heterodimerize with the small subunit. Calpain-3 has been shown to form a homodimer via the PEF-domain[42,47],...
**Figure 2.4: Members of the human calpain family.** The domain structures of members of the human calpain family are shown in linear bar representation. Calpain isoforms are listed next to the domain structures with the first being the classical calpains followed by non-classical calpains. The catalytic triad (cysteine, histidine, asparagine) is labeled for each calpain in the protease core domains with the exception of calpain 6, which lacks the cysteine and histidine residues labeled in red. Abbreviations used: PEF, penta-EF hand; GR, glycine-rich; NS, N-terminal insertion sequence; IS1 and IS2, insertion sequence 1 and 2; T, tra-3; PBH, palB homologous domain; SOL, SOL homologous domain; Zn, zinc-finger motif. Adapted from Ph.D. theses of T. Moldoveanu and R. Hanna, with updated domain nomenclature.[40].
while an active gastric calpain complex is formed from the heterodimerization of calpains-8 and -9[48]. Although the structure of this complex is not known, it is quite likely that the two constituent calpains associate through their C-terminal PEF domains. Human calpains arise from a total of 15 genes, and many have homologues in other mammals.

The non-classical or “atypical” calpains lack PEF-domains, and are not known to dimerize. Calpains-5 and -6 contain a C-terminal T domain, named for its similarity to the Caenorhabditis elegans protein tra-3[49]. Calpain-6 is unusual as it lacks the catalytic residues necessary for proteolytic activity. Catalytic activity is presumably not necessary for its function in stabilizing microtubules[50]. Calpain-7, a homolog to the fungal calpain PalB, varies from the “typical” calpains by containing a tandem repeat of microtubule-interacting and transport domains at the N terminus[51]. Calpain-10 is involved in type 2 diabetes mellitus and contains a second C2L domain in place of the PEF domain[52]. Calpain-15, also known as SOLH, is homologous to the SOL gene from Drosophila melanogaster and contains 5 consecutive zinc-finger domains[53].

Recent forays into genome sequencing have allowed for the identification of many calpain-like proteins in non-mammalian organisms. As previously described, organisms such as Drosophila melanogaster and Caenorhabditis elegans possess genes for calpain-like proteins. Among the most well-studied calpain-like proteins outside
of mammals, 12 have been found in *C. elegans*, one of which is the aforementioned tra-3 involved in sex determination[54,55]. Of the four genes found in the *D. melanogaster* genome, two of their gene products were found to be calcium-dependent and proteolytically active[56]. In addition to those in nematodes and insects, calpain-like genes have been found in kinetoplastids[57], plants[58], and fungi[59].

2.3.2 Physiological roles and pathological conditions

Calpain is involved in a large number of roles in the cell. It has been hard to pinpoint and study the primary roles of calpains because they have been many linked to so many different cellular processes. Studies have shown calpain to be involved in a number of physiological events[39]: progression through the cell cycle[5], cell migration[60], regulation of transcription factors and gene expression[61], cell necrosis and apoptosis[62], signal transduction pathways[63], long-term potentiation of neurons[64], and repair of wounded plasma membrane[65].

Calpains generally make specific cuts in their substrates by cleaving at a limited number of exposed sites. Thus calpains are responsible for remodeling or regulation of proteins in signal transduction rather than full digestion of substrates as occurs with components of the proteasome or lysosomal proteases. There are many signal cascades involving the limited cleavage of targets by calpains. For example, apoptotic
cell death can be triggered by the release of cytochrome c from mitochondria, initiating a signaling cascade that results in the activation of caspases. The Bcl-2 family of proteins regulates this process for either pro-survival or pro-apoptosis; the anti-apoptotic Bcl-2 protein prevents the release of cytochrome c, while the pro-apoptotic Bax protein promotes release of cytochrome c[66]. Increases of intracellular $\text{Ca}^{2+}$ levels have been shown to induce apoptosis in a variety of cells and are an example of cross-talk between caspase and calpain pathways[67]. Specifically, N-terminal cleavage of Bax by calpain releases a pro-apoptotic protein fragment that induces cytochrome c release from mitochondria, independent of Bcl-2, followed by cell death[68]. Calpain also cleaves the tumour suppressor protein p53 at a single N-terminal site, thereby destabilizing the p53 protein that would otherwise promote cell survival. p53 mutants resistant to calpain cleavage are seen to be more stable than wild-type protein, avoiding apoptosis[69]. Calpain involvement in cell migration has also been extensively studied showing calpain-mediated limited cleavage of proteins at focal adhesions include paxillin and talin[70] and dissociation of the integrin-cytoskeletal bonds[71]. In addition, the cleavage of signal transduction proteins in the ERK/MAPK pathways by calpains is involved in EGF-mediated migration[72].

As one can imagine from the sheer number of different physiological processes with which calpain is involved, it is at least partly responsible for a number of pathological disorders, and as such is of interest to the medical field. Some disease states are caused by mutations in the sequence of the calpain protein or changes in
gene regulation. Mutations in the gene encoding calpain-3 can result in limb-girdle muscular dystrophy type 2A[73], through direct inactivation of calpain-3 or loss of function caused by accelerated autoproteolysis[74]. In some people, mutations in the gene for calpain-10 contribute to type 2 diabetes[75]. Additionally, down-regulation of the calpain-9 gene can cause gastric cancers[76].

Another problematic situation is any loss of calcium homeostasis resulting in an influx of Ca$^{2+}$ into the cells. This can overactivate calpain and cause a variety of issues. One of the most devastating losses of calcium homeostasis occurs following ischemia and the subsequent reperfusion in heart attack and stroke. After a loss of oxygen due to reduced blood flow, the cellular processes responsible for producing ATP fail, causing depolarization of the cell due to diminished ATP-dependent ion pump activity. Due to this depolarization, voltage-gated ion channels are activated, causing an influx of Ca$^{2+}$. While other problems occur in the ischemic cascade from the depolarization, the calcium release triggers the release of glutamate, which, in turn, promotes the release of more calcium via Ca$^{2+}$-permeable NMDA receptors[77]. Excess calcium activates the many calcium-dependent proteins, including phospholipases and proteases, and results in intertwined signaling pathways for the degradation of membranes and proteins, and the initiation of apoptosis and/or necrosis. Restoring blood flow to these tissues causes many similar problems: a large increase in intracellular Ca$^{2+}$, resulting from reactivation of the Ca$^{2+}$-ATPase in the SR, is seen immediately with myocardial reperfusion and leads to many of the same cascades as with ischemia[78]. From the increased levels of
calcium due to ischemia and reperfusion, calpains are among some of the proteases activated and their overactivation can lead to a number of pathological conditions.

Recently, much work has linked overactivation of calpains with many brain-related pathological conditions including neurodegeneration, such as Alzheimer’s, Parkinson’s, and Huntington’s diseases, and traumatic brain injury. Neurodegeneration can occur post-ischemia from such events as stroke. Calcium influx in neurons is mostly generated during reperfusion, at concentrations high enough for calpain activation, and calpain cleavage of glutamate receptors is one of the ways calpain can cause neurodegeneration[79]. Alzheimer’s disease (AD) is caused by neuronal cell death in the hippocampus and the cortex, and is characterized by the formation of extracellular plaques consisting of amyloid-β peptide aggregates and tangled networks of the tau protein[62]. Calpain itself has been found to be involved in the processes that control amyloid precursor protein[80] and tau protein[81]. In traumatic brain injury (TBI), calpain is activated shortly after the injury, initiating several previously-discussed signaling cascades, and treatment with calpain inhibitors has been shown to prevent TBI-associated neurodegeneration in mice[82]. With pathological conditions such as AD and TBI, activation of calpain by Ca\(^{2+}\) influx can lead to both necrotic and apoptotic cell death through a variety of mechanisms[83].
2.3.3 Calpain substrates and cleavage preferences

Calpain cleaves a wide variety of substrates. *In vitro* studies have shown over 100 different proteins are cleaved by calpains, although not all of these are physiological substrates *in vivo*. Most calpain substrates identified *in vitro* can be classified into four main categories: i) cytoskeletal proteins, ii) kinases and phosphatases, iii) membrane-associated proteins, and iv) transcription factors. Substrates such as spectrin and casein have been among the most widely used for monitoring calpain activity both *in vitro* and *in vivo* yet do not share a consensus sequence. Work was done to compile cleavage sequences from the then 106 known calpain substrates, and provided a sequential preference of calpain cleavage sites from $P_4$ to $P_7^\prime$ of TPLKSPPPSR[84]. This was based on protein substrates however, of which local secondary structure could influence calpain cleavage thus the abundance of Pro residues in this sequence could break any local secondary structure, enhancing cleavage by calpain. In fact, other groups have shown that unstructured backbone regions or specific tertiary structures were more important for cleavage by proteases[85,86].

An active gene construct of the calpain-1 protease core has been a great tool for studying the enzymatic mechanism of calpain because it is a less complicated system than using full-length enzyme. Inhibitor-bound calpain structures have shown the importance of the S1 and S2 subsites for substrate and inhibitor recognition. However,
these have provided limited knowledge about the ideal substrate cleavage sequence and if the other positions along the active site cleft confer any preference. Using the calpain-1 protease core, the ideal calpain substrate sequence was deduced independent of structural factors[87]. Based on previous inhibitor-bound structures, it was seen that calpain had a preference for leucine at the substrate P2 position and this was confirmed in the studies by Cuerrier et al. Calpain selectivity was determined for substrate positions across the active site ranging from P5 to P3', which proved to be somewhat consistent with the previously described studies on a compilation of known calpain cleavage sites by Tompa et al. [84](Table 2.1). These studies allowed for an ideal cleavage sequence to be determined followed by the design of an ideal substrate that, when positioned between a fluor and a quencher, provides a tool for monitoring cleavage by calpain and demonstrates a higher turnover rate with respect to calpain activity than other fluorescent substrates. The peptide (EDANS)-EPLFAERK-(DABCYL) has aided in observing in vitro calpain activity and to further quantify inhibition via fluorimetric assays.

2.4 Calpain structures

The classical and ubiquitous calpains-1 and -2 are the most studied members of the calpain family. The first calpains to be discovered, these proteases are abundant in mammalian tissues. They are similar in structure and sequence, sharing 55-65%
Table 2.1: Residue preference of the calpain-1 protease core for each position in a peptide substrate[87]

<table>
<thead>
<tr>
<th>Position</th>
<th>P5</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1’</th>
<th>P2’</th>
<th>P3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference</td>
<td>P</td>
<td>F</td>
<td>F&gt;L&gt;P</td>
<td>L&gt;V</td>
<td>L=F</td>
<td>M&gt;A&gt;R</td>
<td>E</td>
<td>R&gt;K</td>
</tr>
</tbody>
</table>
identity, yet differ in the required concentration of calcium for activation,[39]; calpain-1, previously known as μ-calpain, requires multi-micromolar calcium concentrations while calpain-2, or m-calpain, requires sub-millimolar levels of calcium. These two calpains are the subject of the majority of studies on calpain.

Over the years, the ability to produce recombinantly expressed calpains has allowed researchers to study the overall structure, activation, enzymatic mechanism, and the relationships between enzyme and ligand, whether inhibitor or substrate. Importantly, recombinant expression of rat calpain-2 and the protease core domains of calpain-1 in *E. coli* allowed for the first crystal structures of calpain enzymes.[88,89] No structures have been solved for native calpains isolated from natural sources.

Structural studies into the calpain family of enzymes began with the crystal structures of the calcium-free and calcium-bound homodimers of the small subunit PEF-domains[90,91]. The structures of the PEF (S) domains showed the tight interaction of two monomers with a conformational change induced upon Ca\(^{2+}\) binding to four of the five EF-hand motifs. The 5th EF-hand motif was found to interact with that of the other monomer, without binding Ca\(^{2+}\). These structures were important to present a model for the heterodimerization of calpain large and small subunits, due to high sequence similarity between the two PEF-domains (L) and (S). In addition, this now suggested the binding of eight Ca\(^{2+}\) molecules to calpain, four in each of the PEF-domains, and began
to elucidate the mechanism of conformational change induced by the calcium binding necessary for calpain activation.

2.4.1 Activation mechanism (Ca\(^{2+}\)-free vs. Ca\(^{2+}\)-bound structures)

Prior to the first heterodimer calpain structures, the calpain domains were defined by predictions based on amino acid sequence and shared homology with papain-like protease and calmodulin Ca\(^{2+}\)-binding domains[92]. The crystal structures of inactive full-length calpain-2 [93,94] supported the tight interaction demonstrated in the PEF-domain homodimer structure and allowed for domain boundaries to be defined based on structural features (Figure 2.5).

The N-terminal domain forms an α-helix that interacts with the small subunit and is described as an anchor helix. The anchor helix does not block the active site, meaning that calpains are not pro-enzymes. This finding was unique and contrary to other cysteine proteases. In addition, the protease core domains PC1 and PC2 of calpain consist of more residues resulting in a more extensive structure relative to papain than previously thought. The catalytic residues are positioned in an active site cleft consistent with other papain-like proteins. Important to note is that in the Ca\(^{2+}\)-free enzyme the catalytic triad is not formed, unlike other papain-like cysteine proteases, with the inactive
Figure 2.5: Crystal structure of the apo-form of rat calpain-2. The crystal structure of inactive rat calpain-2 (PDB: 1DF0) is shown in cartoon representation and coloured by domain as per the domain structure representation below. In order from N to C termini: anchor helix (red); PC1, protease core domain 1 (orange); PC2, protease core domain 2 (yellow); C2L, C2-like domain (green); transducer arm (grey); PEF(L), penta-EF hand domain on the large subunit; PEF(S), penta-EF hand domain on the small subunit. Not shown is the GR (glycine-rich) domain of calpain, missing from the construct used for crystallography.
catalytic cysteine protonated and >8.5 Å away from the histidine (Figure 2.6B). This is far from the ~3.5 Å distance between the deprotonated cysteine and histidine in other cysteine protease structures. In addition, assembly of the catalytic triad is hindered by the presence of Trp288 (calpain-2 numbering) projecting into the active site cleft.

The third major domain, the C2L domain, in the calpain structure consists of an anti-parallel beta-sandwich formed by 8 sheets and bears some resemblance to a C2 domain. Since C2 domains typically contain Ca$^{2+}$ and phospholipid-binding sites, it was hypothesized that the C2L domain of calpain contains these sites with the potential to bind a total of 3 Ca$^{2+}$ ions. The C2L domain makes contacts with the PEF domain from the large subunit and is connected to PEF (L) by a long, flexible ‘transducer region’. The PEF domain on the large subunit interacts with the PEF domain on the small subunit through the previously discussed pairing of the 5th EF-hand motif. The glycine-rich (GR) domain was omitted from the construct used by Hosfield et al. and is not seen in the structure by Strobl et al., despite having been included in their construct. It is assumed that the GR domain is too disordered to be seen in the crystal structure and its function remains a mystery.

Exogenous proteolysis assays were performed to probe the structure of calpain. Using the inactive mutant C105S to avoid the complications of autoproteolysis, calpain was slowly digested in the absence of calcium ions. In the presence of these ions the
Figure 2.6: The protease cores of apo calpain-2 and holo calpain-1. (A) Shown on the left in cartoon representation is the crystal structure of the protease core domains of Ca$^{2+}$-free rat calpain-2 with PC1 coloured orange and PC2s coloured yellow (PDB: 1DF0). Shown on the right are the protease core domains of Ca$^{2+}$-bound rat calpain-1 (PDB: 1KXR) with the same colouring. Calcium ions are represented as grey spheres. (B) Similar to (A), the protease core domains of Ca$^{2+}$-free rat calpain-2 and Ca$^{2+}$-bound rat calpain-1 are shown on the left and right, respectively. Catalytic residues and Trp298 (calpain-1 numbering) are shown as sticks, while the distance between the catalytic cysteine (serine mutation in this structure) and histidine is noted with a dotted line.
enzyme was more rapidly cleaved but some domains were stabilized by calcium activation against the external proteolysis, implying major conformational changes caused by calcium-binding[95].

The structure of the Ca$^{2+}$-bound protease core of calpain-1 was a crucial step in studying the properties of activated calpains[89]. Calcium bindings sites are present at two non-EF hand sites, one on either side of the active site cleft. Upon binding calcium, the PC1 and PC2 domains rotate relative to one another by approximately 25°[96] (Figure 2.6A). This rotation aligns the catalytic triad and moves the active site Cys and His closer from $>8.5$ Å to the $\sim3.5$ Å required for activity (Figure 2.6B). In addition, the obstructing Trp298 (calpain-1 numbering) residue moves out of the active site to occupy a hydrophobic patch that is exposed upon rearrangement of the calcium-binding loops. The protease core of calpain-1 resembles papain-like cysteine proteases in its active form. Since the protease core of calpain can be active without the PEF domains, albeit at a lower activity than the whole enzyme, it can be said that full-length calpain activity depends more on the binding of Ca$^{2+}$ to the protease core non-EF hand sites to create an active site and less on the binding to the PEF (L) and (S) domains.

Many pieces of the puzzle were put together when the active calpastatin-bound calpain-2 structure was solved by X-ray crystallography[97,98] (Figure 2.7). Active
Figure 2.7: Crystal structure of Ca\textsuperscript{2+}-bound rat calpain-2. The crystal structure of Ca\textsuperscript{2+}-bound rat calpain-2 (PDB: 3BOW) is shown in cartoon representation and coloured by domain as per the domain structure representation below. In order from N to C termini: anchor helix (red); PC1, protease core domain 1 (orange); PC2, protease core domain 2 (yellow); C2L, C2-like domain (green); transducer arm (grey); PEF(L), penta-EF hand domain on the large subunit; PEF(S), penta-EF hand domain on the small subunit. Calcium ions are represented as grey spheres. Note that the anchor helix present in the construct is missing from the solved structure.
calpain-2 binds to 10 Ca\textsuperscript{2+} ions, one in each of PC1 and PC2 domains and four in each of the PEF (L) and PEF (S) domains. Surprisingly, no calcium ions were bound to the C2L domain, and the topology of the β-sheets in the C2L domain differs from conventional C2 domains. The active conformation of full-length enzyme is compacted when compared to the inactive conformation, and the rigid body rotation of the protease core domains is consistent with that of the protease core alone. The active site cleft is formed with catalytic residues in position for cleavage of substrate. Comparing Ca\textsuperscript{2+}-free and Ca\textsuperscript{2+}-bound structures, the PEF (L) and PEF (S) domains undergo a small conformational change, quite unlike the related protein calmodulin but consistent with the earlier small subunit homodimer structures. With the structure of the whole enzyme in place, it appears that the calcium-binding sites on the protease core domains are for catalytic activation and enzyme activity, whereas calcium-binding sites on the PEF domains are for stabilizing the protease core and allowing for increased protease activity only under high enough concentration of Ca\textsuperscript{2+}. That is the PEF domains provide additional safeguards against casual stimulation of calpain and ensure that activation is cooperative and concerted event that requires the binding of numerous calcium ions at multiple sites over the enzyme.
2.4.2 Calcium requirement for activation *in vitro* and *in vivo*

The concentration of calcium required for activation differs for each calpain. For example, the Ca$^{2+}$ required for half-maximal *in vitro* activation of calpain-1 is 30-50 μM but for calpain-2 is 400-800 μM[39,99]. The mechanism for activation of calpains by calcium has been solved as one can now study structures of both inactive and active forms of the enzyme. However, the basis for the difference in calcium concentrations remains unknown. The apo structure of a chimeric calpain containing mostly calpain-1 but with N- and C-terminal regions borrowed from calpain-2 was solved by Pal *et al.* who suggested that the inactive calpain-1 whole enzyme may have its active site closer to fully formed than calpain-2, resulting in enhanced protease activity and reduced Ca$^{2+}$ requirement[100].

Even if the slight structural differences between calpains-1 and -2 can account for the different calcium requirements between the calpains, the overall intracellular calcium levels *in vivo* are much lower than that required for activation of either calpain (nM vs μM and up). There have been several mechanisms suggested for reducing the calcium requirement of calpains to *in vivo* levels including: activator proteins[101,102], covalent modifications including phosphorylation[103], autoproteolysis[99,104-107]; and subunit dissociation[108-111]. One of the better proposals may be the localization of calpain to membranes[112-114] or similar structures for benefitting from the very high localized...
calcium signal produced near the calcium channels and pumps. Once this local high calcium concentration diffuses, calpain will become inactive, demonstrating a safety mechanism whereby calpain is prevented from being overactive for too long a period of time.

2.4.3 Autoproteolysis

Calpain is known to undergo significant intermolecular autoproteolysis upon calcium-activation beginning with the rapid removal of the N-terminal helix and loss of the GR-domain in vivo[106,115]. At the same time, cuts appear in the C2L domain and accompany the loss of enzyme activity[116]. There are reports in the literature that limited autoproteolysis can cause activation of the enzyme, and cleavage of the N-terminal region may reduce the requirement for calcium[99]. This remains not entirely clear, but with recent structures could be explained as a structural constraint release. Without the N-terminal helix, the PC1 and PC2 domains may be freer to rotate into the active conformation, and the energy required for the conformational change upon calcium binding may be lower.
2.4.4 Heterodimerization and subunit dissociation

Subunit dissociation is another mechanism put forward by some researchers to account for calpain activation at the low calcium levels found in vivo [108-111]. The current active calpain structures reveal tight binding (extensive, hydrophobic) between large and small subunits, which is difficult to reconcile with the idea that the large subunit can release the small subunit [97,98]. Furthermore, recent studies have shown that heterodimerization is required for activity and a 1:1 stoichiometry of subunits can be seen in SDS-PAGE results, suggesting that earlier reports of subunit dissociation may be due to misidentification of autoproteolytic products [116] (Rachel A. Hanna, et. al, unpublished).

2.4.5 Using the protease core domains as a “mini-calpain”

Since full-length calpain can undergo autoproteolysis and aggregation upon activation by calcium, the protease core of calpain-1 has been a widely used tool in studies to avoid these complications, and analyze the catalytic triad and activation mechanism. The protease core domains are crystallizable and the structures of numerous calpain core-inhibitor complexes have been solved [117-121] (Figure 2.8). Structures of the protease core have revealed differences in the primed-side residues in the active site from other cysteine proteases, which some inhibitors have taken
Figure 2.8: Calpain inhibitor structures in the active site cleft. The protease core of calpain-1 is displayed as a surface representation while inhibitors are displayed as sticks. The enzyme carbons are coloured white, with charged and uncharged atoms as vibrant or pale colours respectively, or other colours for inhibitors. Inhibitors shown are E64 (yellow), leupeptin (magenta), ZLAK-3001 (cyan), WR-18(S,S) (salmon), and ZLLYCH2F (pale blue) (PDB: 1TLO, 1TL9, 2R9C, 2NGQ, and 2G8J, respectively). Atoms are coloured by type (red – oxygen, blue – nitrogen, yellow – sulphur).
advantage of. Examples include the ZLAK compounds\cite{117}, where an adenine ring on the inhibitor is involved in stacking interactions with a tryptophan residue on calpain, and SNJ-1715\cite{118}, with a molecule of buffer occupying the primed side. The use of the calpain protease core for calpain-inhibitor complex structures has limitations because interactions with distal domains are missing. Unfortunately, these are interactions that could make inhibitors more calpain-specific because these distal domains are peculiar to calpain and are not present in other cysteine proteases. For example, potential interactions with the C2L domain of the whole enzyme may influence the specificity of the inhibitor SNJ-1945 for calpain whole enzyme over other cysteine proteases with similarity to only the core domains\cite{118,122}. Future progress on structure-based calpain-specific inhibitors will likely require crystallography with the whole enzyme.

2.5 Calpastatin, the endogenous calpain inhibitor

Calpastatin is the endogenous polypeptide inhibitor of calpains-1 and -2, and possibly other calpain family members\cite{123,124}. Calpastatin does not inhibit other papain-like cysteine proteases, unlike other less-specific proteinaceous inhibitors such as cystatins. There is only one gene coding for the calpain-specific inhibitor. However, different calpastatin isoforms can be produced through alternative splicing and the use of four different promoters\cite{125-127}.
2.5.1 Regulation of calpain activity

The classical calpastatin isoform is approximately 75-80 kDa and consists of an L domain followed by four homologous inhibitory domains 1-4. Each domain 1 through 4 contains subdomains A, B, and C. While subdomain B is required for inhibition of calpain, subdomains A and C are responsible for increased affinity of calpastatin for calpain[128]. Calpastatin is responsible for the inhibition of calpain in the cell and can bind to enzyme only in conditions where calpain is in the active calcium-bound conformation.

Initially, several studies suggested calpastatin might allosterically inhibit calpain. First, a region corresponding to the residues key for inhibition in subdomain B were found to have no effect on the protease core "mini-calpain" [89]. Second, calpastatin was cross-linked to calpain and mapped to a region in the C2L domain[129]. Third, the cysteine protease inhibitor E64 was able to bind and inhibit calpain even in the presence of calpastatin indicating the active site cysteine was freely available[130]. All this changed when the crystal structures of calpastatin-bound calpain were solved and showed calpastatin subdomain B occupying the active site cleft[97,98].
2.5.2 Structure and calpain inhibition

Calpastatin is an intrinsically disordered protein when not bound to calpain but some areas do have a predisposition to form secondary structure\[131\]. Subdomains A and C have a tendency to form α-helices\[132\] and residues in the middle of subdomain B form a β-turn\[132,133\].

The first structures of calpastatin subdomains bound to calpain included the structure of a 19-residue fragment of subdomain C bound to the PEF (S) domain homodimer. It was bound in a hydrophobic pocket that expands upon Ca\(^{2+}\)-binding\[90,91\]. Based on homology, an equivalent fragment of subdomain A was modeled to bind to the similar hydrophobic groove on the PEF (L) domain. On their own, subdomains A and C do not inhibit calpain and inhibition of the enzyme requires subdomain B\[134\]. The sequence TIPPEY within subdomain B is also necessary for inhibition, but not sufficient. In fact, a 27-residue section of subdomain B is responsible for calpain inhibition and has been used as an inhibitor on its own with a lower affinity than a fully intact calpastatin domain\[128,135\]. Two residues key for two β-turns in this 27-residue section were found to be intrinsic to the inhibition of calpain\[136\], and modeling and cyclization studies indicated that these β-turns induced a loop-like conformation of the peptide\[137,138\].
The crystal structures of calcium-bound calpain inhibited by calpastatin have revealed the mechanism of binding and mode of inhibition of calpain[97,98] (Figure 2.9).

Supporting earlier studies, it was found that subdomains A and C act as anchors forming amphipathic α-helices on binding to hydrophobic clefts in the two PEF domains as previously modeled. Subdomain B is responsible for inhibition of calpain by binding and occluding both sides of the calpain active site cleft formed by the protease core domains while binding extensively across the C2L domain. Subdomain B avoids cleavage by looping up and away from the active site cysteine with a β-turn at the conserved Gly(613). The loop up and away from the active site cysteine, at the β-turn positions identified previously, allows for calpastatin to block substrate binding to the enzyme while avoiding cleavage itself. Subdomain B also forms a two-turn amphipathic α-helix involving the conserved TIPPEY residues on the primed side of the cleft, increasing potency and aiding in specificity of the inhibitor for calpain. The looping away of calpastatin from the active site Cys explains how this active site residue is available to react with E64 even when calpain is bound by its protein inhibitor.

2.5.3 Stabilization of calpain

Calpain has a tendency to aggregate when activated by Ca$^{2+}$ in the presence of divalent cations[111]. The aggregation of the large subunit is said to occur due to
Figure 2.9: Structure of calpastatin-bound Ca\(^{2+}\)-activated rat calpain-2. (A) The crystal structure of calpastatin (purple cartoon) bound to activated rat calpain-2 (domain coloured pale surface) (PDB: 3BOW). (B) Subdomain B of calpastatin bound to the active site cleft. Residues that form the loop away from the enzyme are coloured yellow, while residues forming the C-terminal helix are green and residues N-terminal of the loop responsible for interacting with the C2L domain are in blue. (C) and (D) The helices from calpastatin subdomains C and A, respectively, bound to hydrophobic clefts on calpain PEF domains. Breaks in the calpastatin structure are areas that lacked density in the crystal structure and are disordered.
autoproteolysis and subunit dissociation[110], however aggregation still happens with the use of an inactivating cysteine to serine calpain mutation[111]. Aggregation can be prevented with the use of monovalent ions but this limits crystallization and experimental conditions. When the crystallization of the calpain-calpastatin structure was initially pursued, co-purification of the calpain-calpastatin complex was necessary in order to prevent aggregation of calpain with the addition of calcium[97]. It was thought that the binding of calpastatin helped to stabilize the calpain enzyme, avoiding aggregation. Use of peptides corresponding to the A and C subdomains of calpastatin were found to prevent aggregation of calpain supporting the hypothesis that aggregation of calpain is caused by the exposure of hydrophobic clefts upon calcium-activation which can be negated by interaction with calpastatin (Rachel A. Hanna, et al., unpublished).

Previous work has suggested that hydrophobic interactions are responsible for calpain sub-cellular localization[108,111,139,140], or interaction with ligands and substrates[141,142]. Other work also suggested that small peptides with homology to subdomains A and C of calpastatin prevent cleavage of large substrates such as tau protein[123,143,144]. Thus, it is conceivable that these hydrophobic clefts open upon calpain activation and are responsible for binding calpain substrates. Calpastatin would then be responsible for inhibiting cleavage of substrates via both the occlusion of the active site cleft and the anchoring pockets.

46
2.6 Exogenous calpain protease inhibitors

The MEROPS database contains not only a categorized list of peptidases but also a large repository of proteinaceous and small-molecule inhibitors[25]. Calpastatin is a unique peptide inhibitor in that it is specific for the heterodimeric PEF calpains like calpain-1 and -2. For this reason, calpastatin and its isoforms are classified under the inhibitor family I27 (calpastatin family). Other proteins such as those belonging to the cystatin superfamily are more promiscuous, inhibiting many members of the papain-like clan of cysteine proteases by blocking the active site cleft[145]. Cystatins inhibit, as does calpastatin, by blocking the active site but with the P1 residue pointing away from the active site cysteine to prevent nucleophilic attack and cleavage[146]. Exogenous calpain inhibitors are generally directed to the active site of the enzyme, interacting with the catalytic cysteine in order to achieve high potency with small molecules. There exists a wide variety of exogenous calpain inhibitors constantly expanded by the development of new, increasingly potent and enzyme-specific inhibitors.

2.6.1 Active site-directed calpain inhibition

Enzyme inhibitors can irreversibly inhibit their target through covalent modification of the enzyme, or reversibly inhibit by competitive, non-competitive, uncompetitive, or mixed means. Allosteric inhibition of enzymes can occur, but to this
date the little evidence for this in the calpain family members is controversial and will be
addressed later in this thesis. Thus the focus of calpain inhibition has been on active
site-directed competitive inhibition.

Active site-directed inhibition by exogenous peptidyl inhibitors uses a substrate
mimic attached to a reactive warhead for covalent modification of the active site
cysteine. Nucleophilic attack by the active site thiol occurs, replacing the scissile amide
bond of the substrate with that of the electrophilic warhead. This reaction can occur
reversibly or irreversibly, depending on the nature of the warhead group (Figure 2.10).
Many reversible electrophilic carbonyl-containing groups have been used in calpain
inhibition including aldehydes and α-keto carbonyl compounds such as α-ketoacids, α-
ketoamides, α-ketoesters, α-diketones, and α-ketophosphorous groups[147]. These
compounds act as transition intermediate analogues, forming stable covalent bonds with
the active site cysteine but the reaction is reversible. Aldehyde inhibitors are numerous
in the calpain inhibition field, however their high cross-reactivity in vivo, limited solubility,
and poor metabolic stability makes them less desired as clinical treatments. Compounds
like α-ketoamides, in addition to a higher metabolic stability, allow for interaction with
both sides of the active site, occupying both primed and unprimed sites and allowing for
improved design towards specificity and potency[148].
Figure 2.10: Common reversible (left) and irreversible (right) warheads for covalent reaction with calpain.
In addition to a warhead, peptidyl inhibitors contain an address region required for enzyme recognition (Figure 2.11). The address region varies between proteases, reflected in the specificity for particular substrates, and spans part of the active site cleft. In calpain inhibitors many have a leucine residue at the P2 position and this preference is supported by calpain substrate studies[84,87]. Attempts have been made to modify the P2 position in order to increase potency and specificity over other papain-like cysteine proteases such as cathepsins[149].

Irreversible inhibitors of calpain activity are also plentiful. Irreversible warheads for calpain include acyloxymethyl ketones, epoxysuccinates, and aziridines [147,150]. While irreversible inactivation of calpains may not be desired in a clinical setting, it remains a key tool for studying the effects of calpain and downstream pathways.

Warheads of protease inhibitors can be specific for one class of protease over others despite similarities in the catalytic mechanism involved. For example, acyloxymethyl ketones and diazomethyl ketones have been shown to be selective for cysteine proteases (including calpain) over serine proteases while chloromethyl ketones affect both cysteine and serine proteases[147,151,152]. Modification to the histidine residue in the catalytic triad is possible and has been done especially for serine proteases.
Figure 2.11: The warhead and address region of peptidomimetic inhibitors. Shown is leupeptin with the reactive aldehyde warhead and peptidic address region labeled. The positions P1-P3 are also labeled.
2.6.2 Peptidomimetics or peptide-derived inhibitors

Peptidomimetics or peptide-derived inhibitors comprise the majority of calpain and cysteine protease inhibitors (Figure 2.12). Peptidyl compounds are good at mimicking the tight interactions between enzyme and substrate, and allow for chemical modifications or additions to increase specificity, potency, and pharmacokinetic properties. Small molecule compounds are used to inhibit protein-protein interactions[153,154] however peptides, whether based on natural structured peptides[155] or designed for disruption of interactions[156], are increasingly being used against proteases as well for their increased target selectivity[157].

Following the identification of the 27-residue fragment of calpastatin subdomain B (B27) responsible for inhibition, attempts were made to modify this polypeptide and develop it into a cell-permeable potent inhibitor with therapeutic potential. Conjugating B27 via a disulfide linker to the cell penetrating 16-residue peptide penetratin increased cell permeability without decreasing inhibitory potency[158]. Reducing the length of penetratin to just 7 residues still allowed for cell permeability[159]. The B27 peptide could be taken up into neurons following the addition of a poly-arginine peptide[160]. While B27 does not lose potency or specificity for calpain, it remains a tool for studying calpain activity; as a peptide, B27 would be highly susceptible to endopeptidases present in cells and would be much less successful in a clinical setting. The penta-
Figure 2.12: Peptidomimetic calpain inhibitors. Example calpain inhibitors are shown and include leupeptin, E-64, SNJ-1945, AK-295, and MDL 28170, aligned by their warhead groups.
peptide LSEAL is another calpastatin-based peptide sequence that has been reported to have some calpain inhibition[143,161], however the mechanism of this inhibition remains unclear.

Many naturally-occurring calpain inhibitors exist but tend to not be specific, inhibiting other cysteine and serine proteases such as cathepsins and trypsin[147]. Leupeptin is a peptidyl aldehyde inhibitor originally discovered in Streptomyces species and is a potent reversible inhibitor of calpain[162]. It is also found to inhibit cathepsins, papain, trypin, and plasmin. E64, a peptidyl epoxysuccinate, is an irreversible inhibitor of calpains from Aspergillus japonicus but again lacks specificity for calpain over other proteases. Crystal structures of both leupeptin and E64 bound to the calpain-1 protease core have allowed for the increased understanding of calpain selectivity[120]. Calpain and other papain-like proteases can accommodate the inhibitors equally well in their wide active site clefts, while interactions with mobile loops and calpain-specific sequences could potentially confer specificity.

More recent forays into developing non-natural peptidomimetic inhibitors has allowed for some increase in specificity and pharmacokinetic properties over natural peptidyl inhibitors. The development of calpeptin was a successful use of a norleucinal modification of leupeptin to increase cell permeability[163]. Further removal from non-
natural peptidomimetic inhibitors involves the use of warheads and chemical synthesis to produce novel calpain inhibitors.

Peptidyl α-keto carbonyl compounds, especially α-ketoacid compounds, were originally used as serine protease inhibitors but Li et al. began to develop them as inhibitors for calpain and other cysteine proteases[148]. α-ketoamide derivatives seemed to retain potency while increasing solubility. Many families of α-ketoamide peptidomimetics were developed. The potent compound AK-275 was developed followed by the more soluble AK-295[148,164]. A series of compounds was designed to investigate modifications at the P1’ position for increased potency and specificity[165]. SJA and SNJ series compounds contain P3 amphiphiles with aims to increase oral availability while retaining potency[122]. ZLAK series compounds extend well into the primed side of the active site cleft for increased specificity [117]. Di-peptide α-ketoamide compounds were derived from the peptidic aldehyde MDL28170 for malaria treatments[166]. Potent benzoylalanine-derived α-ketoamide compounds were also developed as potent, water-soluble, orally bioavailable alternatives[167].

Following in the footsteps of E64 research, epoxysuccinate-based inhibitors are another lead in calpain inhibitor development for irreversible modification despite this being a less desirable strategy for drug development. Cuерrier et al. created epoxide-based peptidomimetic libraries probing P4, P3, and P2 positions of the basic
inhibitor[119]. Testing the inhibition by WR compounds against calpain and cathepsins demonstrated specificity of calpain over other cysteine proteases while reaching nM potency (in terms of IC$_{50}$) for calpains-1 and -2, and elucidated additional preferences in terms of interaction with calpain in the active site. The use of epoxide-incorporating calpain inhibitors remains popular, as exemplified by the recent research by Schiefer et al. Using E64 as a lead, epoxide-containing compounds were developed with aims to maintain E64 potency but increasing calpain selectivity and drug-like characteristics[168].

Peptide-based macrocyclic inhibitors have also been in development in recent years for calpain. In order to present peptides with a favoured β-strand formation and greater resistance to peptidase degradation, Abell et al. have made a series of aldehyde-containing macrocyclic calpain inhibitors with the preferred leucine in the P2 position[169].

The ability for peptidomimetic or peptidyl-based inhibitors of calpain to resemble calpain-specific substrates increases the chance for calpain specificity. For example, Bátholy et al. have recently investigated the use of azapeptide inhibitors using the calpain-substrate consensus sequence previously identified by Tompa et al.[84,170]. The combined use of calpain substrates and calpain-specific structural characteristics can aid in the development of calpain-specific inhibitors.
2.6.3 Non-peptidic small-molecule inhibitors

Non-peptidic small-molecule inhibitors of calpains are feasible. Small-molecules could potentially bind to calpain-specific domains and pockets outside of the active site cleft, like the protease core calcium-binding sites, resulting in inactivation of calpains with no effect on other papain-like proteases. There has been limited success in this approach to calpain inhibitor development. Mercaptoacrylate compounds have been presented as potential allosteric calpain inhibitors[171,172]. The PD series of compounds have been shown to bind to the PEF domains of calpains, notably the hydrophobic clefts normally occupied by calpastatin subdomains A and C[91,123]. Looking at the recent crystal structure of calpastatin-inhibited calpain, it remains unclear as to the mechanism of allosteric calpain inhibition.

Other non-peptidic inhibitors of calpain are being developed with many compounds still binding the active site of calpain in a competitive manner[149,152]. Aurintricarboxylic acid (ATA) is an inhibitor of calpains that also affects other enzymes such as endonucleases[173]. Quinolinecarboxamides have also been presented as calpain inhibitors with a lesser effect on cathepsins that may function in a non-competitive manner[174]. Other non-peptidyl compounds have been reported to inhibit calpain including isoquinolines[175], 6-hydroxy-3-morpholinones[176], and dihydroxychalones[177]. Many complications arise with the identification of non-peptidic
calpain inhibitors. Notably, the previously reported compounds diketopiperazines[178] and HIV protease inhibitor ritonavir[179-181] have failed to produce reliable inhibition results[182-184].

2.6.4 Selectivity of inhibitors for calpain over other cysteine proteases

Calpain inhibitor development carries with it a few intrinsic problems. Many compounds lack ideal pharmacokinetic properties including cell permeability and metabolic stability. The main issue with calpain inhibitors remains selectivity. Papain-like cysteine proteases share homologous protease core domains especially in the active site cleft of the active enzymes such as seen with calpain and cathepsins (Figure 2.13). The protease core domains are so similar that papain has been used as a model for developing inhibitors against cathepsin K[185] and cathepsin L[186].

The specificity of calpain inhibitors is driven by the preference for other residues farther away from the scissile bond. For example, leucine at the P2 position is preferred by many cysteine proteases so modifications farther out could potentially increase specificity as the active site begins to differ. Calpain-specific hydrophobic pockets near P3’ and interactions with flexible loop regions are two such areas that could be manipulated. Some modifications of existing compounds could be made to drive specificity through interaction with domains present in calpain - like the C2L domain - but
Figure 2.13: Comparison of active site clefts between calpain, papain, and cathepsin. Shown in cartoon representation are: Ca$^{2+}$-bound calpain-2 (white), cathepsin L (pink), papain (orange), and cathepsin K (green). (PDB: 3BOW, 2YJ2, 1PE6, 1ATK, respectively.) All four structures were aligned by the catalytic triad and surrounding secondary structure. Side chains for catalytic residues are shown with atoms other than carbon coloured by type (red – oxygen, blue – nitrogen, yellow – sulphur). Note the presence of the C2L domain on the unprimed side of the calpain active site cleft, but lack thereof in the other papain and papain-like enzyme structures.
absent from other cysteine proteases. The SNJ class of inhibitors is one such example with the crystal structure of SNJ-1945 showing a long ether tail stretching far enough on the unprimed side of the protease core active site to potentially interact with pockets on the C2L domain in the whole enzyme[118]. With the advent of the active calpain crystal structure, such structure-based designs to increase specificity are now possible.

Non-peptidic calpain inhibitors can also be developed to interact with a number of extraneous domains present in calpains. The exact mechanism of allosteric inhibition is hard to imagine, but such compounds might prevent the key conformational changes from occurring with calcium activation. Novel small-molecule inhibitors of calpain that satisfy both specific and pharmacokinetic requirements, bind to calpain uniquely, while lacking the chemically reactive groups responsible for cross-reaction, would be ideal.

Despite the wide variety of calpain inhibitors, their lack of selectivity for calpains over other cysteine proteases has limited their usefulness. As clinical treatments for calpain pathologies, the cross-reactivity of calpain inhibitors could be dangerous for patients. Even as tools for studying the physiological roles of calpains, any lack of specificity could cloud conclusions about the function of calpains and downstream effects of calpain inhibition. Having an inhibitor to distinguish between the two main isoforms of calpain would be highly desirable but might be practically impossible.
2.6.5 Theoretical and existing therapeutic uses of calpain inhibitors

Several calpain inhibitors have shown to be somewhat more selective for calpains over other cysteine proteases. The SNJ series of compounds have been used in studies that demonstrate good pharmacokinetic properties\[187\] in addition to neuroprotective\[188\] and cardioprotective\[189\] abilities. SNJ-1945 has been investigated as a treatment for both traumatic brain injury\[190\] and retinopathies\[191,192\]. Cataractogenesis has been a hot topic for calpain inhibitor research and other groups have developed alternative inhibitors to SNJ-1945. Modeling work by Stuart et al. has been done to use the SNJ-1945 precursor as a lead for further drug development\[193\].

Calpain inhibitors have been used to treat brain injuries, especially in the last decade or so. Traumatic brain injury research has identified calpain inhibitors like AK-295 as a therapeutic treatment to reduce neuronal pathologies\[194\] with potential future use of calpain inhibitors to prevent extensive brain damage\[195\]. Calpains have also been identified as key drug targets to treat brain diseases such as Alzheimer's\[196\] and Parkinson's\[197\] diseases. Treatments with calpain inhibitor A-705253 managed to mitigate Alzheimer's Disease-like pathology in mice\[198\], while treatment with MDL-28170 and overexpressed calpastatin attenuated the loss of neurons in the mouse model for Parkinson's Disease\[199\].
2.7 Research objectives

Currently, calpain inhibitors are much sought-after as treatments for a wide variety of calpain-associated pathologies. The lack of calpain-specific potent inhibitors in the field was the main driving force behind my research. The objective of my thesis research is to design novel calpain inhibitors that are both potent and specific. The new calpain-calpastatin structure allowed me to begin to figure out the best way for structure-based design: mimicking the tight active-site interactions and key structural features of the calpain-specific inhibitor calpastatin. Described in Chapter 3 is the development of calpain inhibitors with cyclic peptides in order to mimic β-turns key for calpastatin inhibition. Cyclic peptides made in collaboration with the Yudin lab included macrocyclic pentamers and amidine-forming peptidomimetics. Original designs were based on the sequence of calpastatin in the calpain active site and the synthesis of peptide libraries allowed me to evaluate the inhibition potential of these compounds. Cyclization of calpain peptidic inhibitors can reduce susceptibility to peptidases and increase potency through reduction in entropy of binding. This research led to the development of both competitive and non-competitive inhibitors, and could potentially provide lead compounds for future drug development. Alternative calpain-specific inhibitors were designed based on the calpastatin subdomain B helix and are described in Chapter 4. Chapter 5 comprises research assessing current models for allosteric inhibition of calpains. My research provides new avenues for calpain inhibition with aims to increase
potency and specificity, and provide a spring-board for future calpain inhibitor research and development.
Chapter 3

Rational design of calpain inhibitors based on calpastatin peptidomimetics

3.1 Abstract

Calpastatin-bound calpain structures and new avenues for the cyclization of peptides have inspired the design of cyclic peptides and peptidomimetics as calpain inhibitors, based on conserved calpastatin structural features in the active site cleft. We have tested a library of 68 peptidic compounds, out of which, four molecules (cPGALK, cPGSGO, amPLKG, and amPGLdO) have shown reproducibly high levels of inhibition against calpain-2 and cathepsin L enzymes. Activity against calpain-2 was shown to be in non-competitive, competitive, and mixed inhibition models. Systematic sequence changes and further design led to the development of amPGIdO, amPGVdO, amPGLdDab, and amPGLdQ sequences, which displayed increased potency and specificity of inhibition against calpain over cathepsin L. Calculated Ki values were in the low μM range, similar to other potent peptidomimetic protease inhibitors. Compared to other macrocyclic calpain inhibitors, these compounds displayed increased specificity for calpain over other papain-like cysteine proteases. The inhibition of calpain by the peptidic protease inhibitors presented herein provides a stepping-stone for novel
protease inhibitor design, and will allow for the development of calpain-specific inhibitors with increased potency.

3.2 Introduction

Calpains are a family of calcium-activated cytosolic multi-domain cysteine proteases that are responsible for specific, limited cleavage of a wide variety of proteolytic targets. They convert a localized Ca$^{2+}$ signal into cellular responses such as reorganization of the cytoskeleton, cell cycle regulation, and apoptosis[39,124,200]. Over-activation, either by specific mutations or upstream effects, cause numerous pathological effects in conditions such as heart attack, stroke, Alzheimer’s disease, cancer, neurodegeneration, muscular dystrophy, and traumatic brain injury[62,83,200-202].

The most extensively studied isoforms, calpains-1 and -2, are heterodimers composed of an isoform-specific 80-kDa large subunit and a common 28-kDa small subunit[39-41,201,203-206]. Calpain-1 requires multi-μM concentrations of Ca$^{2+}$ while calpain-2 requires an order of magnitude higher level in the sub-mM range. Crystal structures of activated calpain have shown the binding of up to ten Ca$^{2+}$: four to each of the two C-terminal penta-EF-hand (PEF) domains (PEF(L) and PEF(S)) and two to the protease core domains (PC1 and PC2)[89,97,98]. Activation by calcium induces
conformational changes in the enzyme, aligning the domains of the protease core for active site cleft formation and resulting in an overall more compact enzyme.

Endogenously, calpain is inhibited by a highly-specific inhibitor, calpastatin, which serves to control the proteolytic activity of calpain in the cell. Calpastatin binds the enzyme only when the latter is in the calcium-activated form and is released at low concentrations of calcium[207]. Calpastatin is an intrinsically unstructured 70-kDa protein comprised of multiple domains, four of which (CAST 1 through 4) are independently inhibitory and each contain subdomains A, B, and C[124]. Crystal structures of calcium-bound calpain inhibited by calpastatin have revealed the mode of inhibition[97,98]. Subdomains A and C form amphipathic α-helices upon binding to hydrophobic clefts in the two PEF domains and help anchor the inhibitor to the enzyme (Figure 3.1A). Subdomain B is responsible for inhibition of calpain by binding and occluding both sides of the active site cleft formed by the protease core domains. Subdomain B avoids cleavage by looping up and away from the active site cysteine with a β-turn at the conserved Gly(613). It also forms a two-turn amphipathic α-helix that binds on the primed side of the cleft.

Calpain’s vast number of roles in the cell and involvement in a variety of physiological conditions and diseases clearly shows the protease is a potential therapeutic target[201,203,204]. Calpain inhibitors display a wide range of potency but, to date, have not been specific enough to avoid inhibiting other cysteine proteases such
Figure 3.1: Mimicking the active site of calpastatin-bound calpain-2 with cyclic peptides. (Caption on following page.)
Figure 3.1: Mimicking the active site of calpastatin-bound calpain-2 with cyclic peptides. (A) The structure of Ca^{2+}-activated calpastatin-bound calpain-2 (PDB: 3BOW). Calpain-2 is shown in surface representation with calpastatin in ribbon format. Calpastatin subdomains are labelled in purple and calpain domains in black. The black dashed square outline represents the area of the active site viewed in (B) and (C). (B) Close-up view of calpastatin (purple) in the active site of calpain-2, with the active site cysteine 105 (serine mutation in this structure) displayed and indicated. In stick representation (purple) are key residues of calpastatin mimicked and overlaid by the initial cyclic peptide design cPKLG (orange) shown with polar hydrogens displayed and residues labelled. N and O atoms of both calpastatin and the cyclic peptide are coloured blue and red, respectively. (C) Close-up view of the calpastatin-bound calpain-2 active site. A representative cyclic peptide docking result from Glide is shown for cPGLGK (cyan), with polar hydrogens displayed and residues labelled. N and O atoms are coloured blue and red, respectively.
as cathepsins. While there are many types of calpain inhibitors including site-directed reversible and irreversible inhibitors with various reactive warheads, their gains in activity have not supplanted the need for specificity. The recent crystal structures of calcium-activated calpain have now allowed for the structure-based design of inhibitors along the entire active site cleft, including those regions of potential interaction with the C2L domain. This had not been possible before, as the only structures available of calpain in the active conformation were those of the uninhibited and inhibited protease core (domains PC1 and PC2)[89,117,118,120,121].

Proteases universally bind to the extended β-strand conformation of their substrates[208]. This insight provides a possible explanation for the superior inhibition achieved by many peptidomimetic molecules with conformational restrictions that favour a β-turn-like formation[169,209]. Since recognition sequences are often 4-5 residues in length, one successful approach to inhibitor synthesis is to lock the sequence into a specific conformation such as a cyclic peptide. Cyclization can result in a reduced entropic penalty for binding to a protein host, thereby making it a favourable interaction. In addition, cyclic peptides exhibit greater stability to proteolytic degradation by peptidases in the cell. Cyclic peptides have previously been used to successfully inhibit HIV protease[210], HCV NS3 protease[211], and even calpain[169].

Building on the initial work of peptide macrocyclization with aziridine aldehyde dimers[212], we recently observed a propensity for the cyclic products to form β-
turns [213]. In the cyclization process, three additional atoms are introduced into the peptide backbone. This cyclization linker is comprised of an exocyclic amide and an aziridine moiety, which can react with a nucleophile (Figure 3.2A). The presence of the carbonyl group in close proximity allows the peptide to adopt a multitude of conformations with transannular hydrogen bonds (Figure 3.2A). Peptide backbones of pentamers cyclized with the aziridine aldehyde methodology have been found to adopt β-turn structures in both solution phase (NMR) and solid phase (X-ray) [213]. In the solid phase structure of cPGLGF, a type II β-turn centered on the Gly-Leu residues was observed, while in the solution phase, the turn was indexed by one residue, such that the turn was centered on the Pro-Gly residues and a type I β-turn was formed (Figure 3.2B).

The aziridine aldehyde-mediated peptide cyclization is performed at 0.05 – 0.2 M concentration without appreciable amounts of cyclodimerization and polymerization by-products formed, which is a distinct improvement over conventional peptide cyclization novel and potent calpain inhibitors (Figure 3.2D). Here we have designed peptides based on the structure of calpastatin in the active site cleft of calpain-2 and investigated the inhibition of calpain by such compounds. We show that cyclic peptide and amidine-based cycle-tail motif peptidomimetics are able to inhibit cysteine proteases including calpain, and that both show various modes of inhibition of calpain-2 with $K_i$ values in the micromolar range and instead peptidomimetics bearing an amidine functionality have been isolated (Figure 3.1C) (Zaretsky et al., unpublished). The amidine products are a novel structural class bearing a cycle-tail motif with a piperazinone-like core and linear
Figure 3.2: Mimicking the active site of calpastatin-bound calpain-2 with cyclic peptides. (Caption on following page.)
Figure 3.2: Macrocyclization of peptides to mimic natural calpastatin β-turn motifs. (A) Aziridine aldehyde based macrocyclization provides access to β-turn motifs in peptides. (B) Cyclic peptide cPGLGF arranges into distinct β-turn when observed by X-ray in the solid state versus the solution phase structure in DMSO. (C) Nucleophilic rearrangement with some peptide sequence may prevent carboxylic acid attack and cyclization. Instead, amidine-containing peptidomimetics are produced. (D) A rational approach towards replicating the β-turn inhibitory sequence of calpastatin with β-turns formed from 18-membered rings.
peptide chain. The cyclization reaction favours cyclic peptide products when all L-amino acid pentamers are cyclized and forms the amidine products predominantly with smaller linear peptides substrates or pentamers with varied side chain stereochemistry.

With the methodology to make two types of peptidic structures, β-turns with cyclic peptides and cycle-tail motifs with the amidine peptidomimetics, in hand, we turned to the crystal structures of the Ca\(^{2+}\)-activated full-length calpain protease as a model for inhibitor development. We hypothesized that by using the calpain-specific inhibition characteristics of calpastatin we would be able to design novel and potent calpain inhibitors (Figure 3.1D). Here we have designed peptides based on the structure of calpastatin in the active site cleft of calpain-2 and investigated the inhibition of calpain by such compounds. We show that cyclic peptide and amidine-based cycle-tail motif peptidomimetics are able to inhibit cysteine proteases including calpain, and that both the macrocycle and peptidomimetic products are responsible for competitive, non-competitive, and mixed inhibition of calpain-2 with \(K_i\) values in the micromolar range.

3.3 Experimental Procedures

3.3.1 Materials

Active rat recombinant calpain-2[88], and calpain-1 protease core[89] were expressed and purified as previously described. Papain, cathepsin L, the cysteine
protease substrate Z-FR-AMC, and leupeptin were purchased from Calbiochem. The calpain substrate (EDANS)-EPLFAERK-(DABCYL) was synthesized by Biomer Technology.

3.3.2 Cyclic peptide and peptidomimetic synthesis.

Peptides were synthesized using the telescopic cyclization-aziridine ring opening synthesis[213]. Unless otherwise noted, a Pro-terminated peptide was cyclized at a 0.10 mmol scale in a 2 dram vial. To it, was added aziridine aldehyde dimer (0.10 mmol, 2.0 eq. as monomer), TFE (1.0 mL, 0.1 M), and tert-butyl isocyanide (0.2 mmol, 2.0 eq.). The reaction was left to stir at room temperature for 4 h and monitored throughout by RP-HPLC/MS. Ring-opening of the aziridine moiety was initiated by addition of thiobenzoic acid (0.4 mmol, 4.0 eq) and left to stir for 2 more hours. Afterwards, Raney®-Nickel slurry (approx. 1.5 mL) was added, reaction capped, and allowed to stir for a final 16 h. The mixture was filtered through Celite with methanol and concentrated under reduced pressure. TFA/DCM (1:1, 4 mL) was then added to the mixture to deprotect the side chains. The cleavage cocktail was evaporated with a stream of nitrogen after 2 h of stirring at room temperature. The TFA salt was then purified by semi-preparative RP-HPLC/MS. The pure fractions (by RP-HPLC/MS) were pooled together and lyophilized to afford the final products as white or off-white powders. Products were identified by NMR.
3.3.3 RP-HPLC/MS

Low resolution mass spectra (ESI) were collected on an Agilent Technologies 1200 series HPLC paired to a 6130 Mass Spectrometer. Compounds were resolved on an Agilent Poroshell 120 EC-C\textsubscript{18}, 2.7 µm, 4.6 x 50 mm\textsuperscript{2} column at room temperature with a flow of 1 mL/min. The gradient consisted of eluents A (0.1% formic acid in double distilled water) and B (0.1% formic acid in HPLC-grade acetonitrile). The gradient method started with a step of 5% of B for the first 0.99 min, followed by a linear gradient from 5% to 95% B in 8.0 min. The column was then washed with 95 % B for 1.0 min and equilibrated at 5% B for 1.5 min.

3.3.4 Peptide NMR

\textsuperscript{1}H and \textsuperscript{13}C NMR spectra were recorded on Varian Mercury 400 and Agilent 500 MHz or 600 MHz spectrometers. \textsuperscript{1}H NMR spectra were referenced to CDCl\textsubscript{3} (7.26 ppm), CD\textsubscript{3}OD (3.30 ppm), DMSO-\textit{d}_\textsubscript{6} (2.50 ppm). \textsuperscript{13}C NMR spectra were referenced to CDCl\textsubscript{3} (77.2 ppm), CD\textsubscript{3}OD (49.0 ppm), and DMSO-\textit{d}_\textsubscript{6} (39.52 ppm). NMR spectra were recorded at 25 °C unless otherwise specified. Peak multiplicities are designated by the following abbreviations: s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet; ds, doublet of singlets; dd, doublet of doublets; ddd, doublet of doublet of doublets; bt, broad triplet; td, triplet of doublets; tdd, triplet of doublets of doublets. 
3.3.5 Modelling, docking simulation, and molecular dynamics simulation studies

Initial designs for the cyclic peptides were modelled using PyMOL[360] based on the crystallographic structure of calpastatin-bound calpain-2[97] (PDB: 3BOW).

Molecular dynamics (MD) calculations were done with the GROMACS package[214] on calpain-2. The protein was solvated in a box using 30365 water molecules. The net charge of the system was neutralized with 18 sodium ions. The system was then subjected to energy minimization, position restrained molecular dynamics to settle the water molecules, and unrestrained molecular dynamics simulations. The simulations were run at a temperature of 273 K using a time step of 2 fs for a total duration of 20 ns. Long-range electrostatics were treated with the particle mesh Ewald method. The simulations were conducted under isothermal conditions using V-rescale temperature coupling. MD studies with calpain-2 used the CHARMM[215] and OPLS[216] forcefields for protein and the SPCE water model. MD calculations were also done on select cyclic peptides. Topology files and parameters for the cyclic peptides were provided by SwissParam[217]. MD studies with cyclic peptides were run in a similar method to those of calpain-2 and used the CHARMM forcefield[215]. Results were analyzed using VMD[218]. Docking simulations were done using AutoDock4 and AutoDockTools4[219], and Glide[220] in the Maestro Software Suite[221] for docking the cyclic peptides to the active site cleft of calpain-2, calpain-1 protease core, and papain. The protein receptors were used in rigid docking mode. For the cyclic ligands, residue side-chains maintained flexibility while the peptide backbone was torsionally constrained based on NMR structures. The AutoDock4 dockings were performed using the Lamarckian genetic
algorithm with parameters as follows: 100 dockings, population size of 150, random starting position and conformation, translation steps of 2.0 Å, rotation steps of 50°, mutation rate of 0.02, crossover rate of 0.8, local search frequency of 0.06, and 2.5 million energy evaluations. Final docked conformations were clustered using a root-mean-squared deviation (RMSD) of 2.0 Å. Glide dockings were performed using default settings with some exceptions: Glide-XP (extra precision) option was selected, torsional constraints were set for cycle backbone and peptide bonds starting from NMR structures, and at most 100 poses per ligand were produced. Glide docking results were analyzed by Prime MM-GB/SA[222] to determine the free energies of binding.

3.3.6 Peptide library screening with calpain and inhibition kinetics in a fluorescence-based hydrolysis assay

The rate of cleavage of the (EDANS)-EPLFAERK-(DABCYL) fluorescent substrate was monitored in a 0.5-mL cuvette using a Perkin-Elmer LS55 fluorescence spectrometer with $\lambda_{\text{ex}} = 335$ nm, $\lambda_{\text{em}} = 500$ nm and slit widths of 10 nm. Triplicate readings were obtained at 0.1 s intervals. The cuvette contained 100 nM calpain-1 protease core or 10 nM calpain-2, 5 µM substrate, and 1-100 µM peptide inhibitor (dissolved in DMSO at stock concentrations of 50 mM) in 10 mM HEPES pH 7.4 and 10 mM DTT. Control assays were performed with an equivalent amount of DMSO. The total concentration of DMSO in each reaction was 0.4%. The reaction was initiated with the addition of CaCl$_2$ to 4 mM in a final volume of 0.5 mL.
3.3.7 Inhibition of papain and cathepsins L in a fluorescence-based hydrolysis assay

For papain activity assays, the rate of cleavage of the (EDANS)-EPLFAERK-(DABCYL) fluorescent substrate was measured in a similar manner as the calpain activity assays. For cathepsin L activity assays, the rate of cleavage of the Z-FR-AMC fluorescent substrate was quantified again in a similar manner but with $\lambda_{\text{ex}} = 360$ nm, $\lambda_{\text{em}} = 460$ nm and slit widths of 8 nm. Triplicate readings were obtained at 0.1 s intervals. The cuvette contained 5 $\mu$M (EDANS)-EPLFAERK-(DABCYL) or 5 $\mu$M Z-FR-AMC, and 100 $\mu$M peptide inhibitor (dissolved in DMSO at stock concentrations of 50 mM) in 10 mM HEPES (pH 7.4) and 10 mM DTT (papain) or 0.1 M NaOAc-HCl (pH 5.5), 1 mM EDTA, and 1 mM DTT (cathepsin L). Control assays were performed with an equivalent amount of DMSO. The total concentration of DMSO in each reaction was 0.4%. The reactions were initiated by the addition of 30 nM papain or 1 nM cathepsin L, respectively. The initial rates were determined according to the procedure used for the calpain activity assay.

3.3.8 Analysis of enzyme kinetics with lead compounds

Fluorogenic hydrolysis assays were performed with concentrations of substrate ranging from 1-100 $\mu$M and inhibitor concentrations from 0-50 $\mu$M. The initial rates were determined by fitting a straight line to the linear portion of the progress curves (first 10 s after calcium addition). The resulting rate data were fitted to Michaelis-Menten curves for the visualization of enzyme inhibition. $K_i$, $K_m$, and $V_{\text{max}}$ values for the reactions were
obtained from simultaneously fitting the uninhibited and inhibited data to the full mixed enzyme inhibition equation below,

\[
v = \frac{V_{\text{max}}[S]}{1 + \frac{[I]}{K_{i1}} \left( \frac{K_{m}}{1 + \frac{[I]}{K_{i2}}} + [S] \right)}
\]

Simultaneous non-linear regression fitting was done by a similar method as Kakkar et al. [223] using the non-linear least squares fitting function of gnuplot[224]. All reactions were corrected for the inner filter effect as previously described[225] by measuring the fluorescence of increasing amounts of substrate with a fixed concentration of free EDANS fluorophore and thereby calculating a correction factor.

3.4 Results and Discussion

3.4.1 Modelling potential inhibitors on the \(\beta\)-turn structure of calpastatin in the calpain active site cleft.

The crystal structures of calpastatin-bound calpain-2 revealed the mode of inhibition to be the result of calpastatin occupying both sides of the active site cleft (Figure 3.1A). Sharp \(\beta\)-turns at calpastatin residues Gly613 and Thr618 keep calpastatin
away from the calpain active-site cysteine and avoid cleavage of the otherwise unstructured inhibitor[97,98]. Previously published methods for making cyclic peptides[212] provided an option for the design and modelling of these peptides to mimic the conformation of calpastatin in the calpain active site (Figure 3.1B).

Some initial models were analyzed using the molecular dynamics software package GROMACS[214]. These models were then used in docking simulations. MD simulations of calpain-2, taken from the calpastatin-bound calpain-2 structure, were used as representative structures with which to dock cyclic peptides using AutoDock4[219] and Glide[220]. The lowest energy conformations for the docked molecules clustered in the unprimed side of the active site cleft and involved interactions with the C2-like domain of calpain (Figure 3.1C).

The tendency for our initial cyclic peptides designs to interact with the C2L domain in docking and MD experiments suggested that the synthesized peptides could potentially be specific inhibitors of calpain, because this domain is absent from other cysteine proteases.
3.4.2 Macrocyclic and amidine-bearing peptidomimetic library derived from macrocyclization of peptides.

All peptides were made with Pro at the N terminus to ensure efficient cyclization. Based on previous reports that cycles sourced from tetramers and pentamers containing 2 d-amino acids tend to undergo amidine formation, we split the library design into two groups.

The first group, made of cyclized pentamers, was aimed at mimicking the inhibitory calpastatin β-turn sequence at KLG. At Lys, the side chain length requirement was investigated with the increasingly shorter amine-bearing side chain residues Orn, Dab, and Dap. Based on the crystal structure of cPGLGF, which positioned the β-turn around Gly-Leu, the basic residues were placed in the fifth position. Finally, as it was previously seen that the solution phase structure of cPGLGF can index the turn by one residue (Figure 3.2B), the basic residues were also placed in the third position.

Amidine-bearing peptidomimetics comprised the second group of peptides formed from the aziridine aldehyde based macrocyclization. This set of inhibitor probes was also designed to screen permutations of KLG tripeptide sequence and length of amine-bearing side chains. As the binding mode was less certain, the order of the KLG tripeptide was also perturbed in a subset of the peptidomimetic probes.
3.4.3 Screening of peptide library using fluorescence-base protease assays.

The cyclized pentamers and peptidomimetics were screened using fluorescence-based enzyme activity assays against calpain-2. Macrocyclic compounds demonstrated significant levels of inhibition of calpain-2 (Figure 3.3A). Approximately 60% of the macrocyclic compounds showed levels of inhibition of calpain-2 that were ≥ 25% at concentrations of 100 µM. Two macrocyclic peptides, cPGALK and cPGSGO, showed reproducibly high levels of inhibition of calpain-2 (> 85%) at concentrations of 100 µM.

Levels of inhibition by the peptidomimetic set of compounds were similar to those of the macrocyclic forms with approximately half of second generation cyclol compounds showing ≥ 25% inhibition at concentrations of 100 µM. One peptidomimetic, amPLKG, also showed reproducibly high levels of inhibition of calpain-2 (> 85%) at concentrations of 100 µM while another, amPGLdO, similarly showed high levels of inhibition of calpain-2 at concentrations of 100 µM, but with more variation (from 16-72%) (Figure 3.3B).

Calpains share many substrates and inhibitors with other cysteine proteases, and are structurally similar with respect to their active site. As a representative example, the active site cleft region of calpain-2 is shown overlaid on cathepsin L, a physiologically relevant papain-like enzyme (Figure 3.4). The cathepsin L protease was aligned by backbone carbon atoms to calpain-2 residues of the catalytic triad and adjoining regions (calpain-2 residues 105-114, 252-268, and 286-292). The alignment between the two
Figure 3.3: Inhibition screening of peptide compounds. (Caption on following page.)
**Figure 3.3: Inhibition screening of peptide compounds.** All compounds were screened at 100 μM. Error bars represent the standard deviation resulting from triplicate assays. Inhibition is determined as a percentage of the initial reaction velocity for the control reaction containing equivalent amounts of DMSO. A lowercase “c” designates the cyclic peptide ring closure, while a lowercase “am” designates amidine-bearing peptidomimetics. “d” designates a D-amino acid. The amino acid ornithine, Orn, is represented by “O”. (A) The hydrolysis of fluorescent substrate by calpain-2 in the presence of cyclic peptides, and (B) amidine-bearing peptidomimetics. (C) The hydrolysis of fluorescent substrate by calpain-2 (black), calpain-1 (light grey), calpain-1 protease core (diagonal lines), cathepsin L (white), and papain (dotted) in the presence of each of the four selected calpain-2 inhibiting peptides.
Figure 3.4: Comparison of the active site region of cathepsin-L and calpain-2. The structures of Ca$^{2+}$-activated rat calpain-2 (grey) (PDB: 3BOW[97]), and human cathepsin L (pink) (2YJ2[226]) are displayed in ribbon representation except for the catalytic triad residues in sticks, where N, O, and S atoms are coloured blue, red, and yellow, respectively. The general area of the active site cleft is highlighted by a dotted black box. Note the absence of an equivalent region in cathepsin-L to the calpain-2 C2L domain (labelled, green circle) on the unprimed side of the active site cleft and the loop regions from residues 160-176 on the primed side of the active site cleft (labelled, orange circle).
proteases is close in these active site regions with an RMSD of 0.415 Å. Due to the high structural similarities seen, it was necessary to test any potential inhibitors for protease selectivity. Inhibition assays using fluorogenic substrates were performed against calpain-1 protease core, papain, and cathepsin L. The four selected lead compounds, cPGALK, cPGSGO, amPLKG, and amPGLdO, showed little to no inhibition of calpain-1 protease core and papain, but high levels of inhibition (> 85%) of cathepsin L (Figure 3.3C).

In general, for both macrocycle and amidine peptidomimetics, the widely variable inhibition levels suggest a strong dependence on amino acid choice and that calpain is amenable only to certain side chains at various positions in the cyclic compounds. Tight protein-ligand interactions depend on the structure of the peptide, whether amidine or macrocycle, and the orientation in the binding site, both of which are directly affected by peptide sequence and the cyclization chemical groups.

3.4.4 Enzyme kinetics analyses and inhibition mechanisms

To determine the mode of inhibition of the lead cyclic compounds with calpain-2, extensive analyses of enzyme kinetics were performed using the fluorogenic substrate assay. Previous attempts to determine the kinetics of cleavage of the (EDANS)-EPLFAERK-(DABCYL) fluorescent substrate by calpain failed to appropriately incorporate errors in fluorescence from the inner filter effect and used only the protease
core domains of calpain-1[87]. The latter are not subject to reduced activity due to autoproteolysis. We moved forward in determining the kinetics parameters with respect to full-length calpain-2. The initial reaction rates were determined by fitting a straight line to the linear portion of the progress curves (first 10 s after calcium addition) that, with the high sensitivity of the fluorimeter, provided enough data points to record the initial rate of reaction while avoiding the complications of autoproteolysis. We used the calpain-1 protease core, thereby avoiding autoproteolysis, to cleave a set concentration of substrate to completion in order to obtain a conversion factor of 1358 RFU/μM of product formed. \( K_m \) and \( V_{\text{max}} \) values for the reactions were obtained from direct fits of the Michaelis-Menten equation to the data using the non-linear least squares fitting function of gnuplot (version 4.6)[224], and the rate of substrate turnover \( (k_{\text{cat}}) \) subsequently calculated. For the cleavage of the calpain-specific FRET substrate by calpain-2, \( K_m = 41.4 \pm 2.3 \ \mu\text{M} \) and \( k_{\text{cat}} = 2.94 \pm 0.07 \ \text{s}^{-1} \).

Using the four potent compounds, the initial rates of substrate cleavage by calpain-2 were fit to the Michaelis-Menten plot for uninhibited and inhibited reactions (Figure 3.5A-D). To accurately obtain reaction parameters, the initial rate of substrate cleavage by calpain-2 was fit to the full mixed enzyme inhibition equation for uninhibited and inhibited reactions simultaneously. This allowed for the calculation of the inhibition constant, \( K_i \), as well as \( K_m \) and \( V_{\text{max}} \) (Table 3.1). Taking into account the error of the fitted equations and parameters, it was determined that of the four lead compounds,
Figure 3.5: Michaelis-Menten curves for the inhibition of calpain-2 by amPLKG, cPGALK, amPGLdO, and cPGSGO. The initial rates of substrate cleavage by calpain-2 in the absence and presence of each of (A) amPLKG, (B) cPGALK, (C) amPGLdO, and (D) cPGSGO were obtained and fit to the Michaelis-Menten equation. The resulting curves were then plotted as shown. Data were obtained for increasing concentrations of compound, each in triplicate: 0 μM (solid line, ●), 10 μM (2-dot dashed line, ■), 20 μM (dot dashed line, ▲), 35 μM (small dotted line, △), 50 μM (dashed line, □), and 100 μM (dotted line, ○).
Table 3.1: Summary of inhibition kinetics

<table>
<thead>
<tr>
<th>Compound</th>
<th>Mode of Inhibition</th>
<th>Kinetic Constants (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B27 WT peptide</td>
<td>competitive</td>
<td>( K_{i1} = 16.8 \pm 1.7 \times 10^{-3} )</td>
</tr>
<tr>
<td>amPLKG</td>
<td>competitive</td>
<td>( K_{i1} = 12.3 \pm 0.8 )</td>
</tr>
<tr>
<td>cPGALK</td>
<td>competitive</td>
<td>( K_{i1} = 16.7 \pm 1.0 )</td>
</tr>
<tr>
<td>cPGSGO</td>
<td>mixed</td>
<td>( K_{i1} = 19.4 \pm 1.6 ) ( K_{i2} = 52.7 \pm 21.1 )</td>
</tr>
<tr>
<td>amPGLdO</td>
<td>non-competitive</td>
<td>( K_{i1} = K_{i2} = 142 \pm 14.1 )</td>
</tr>
</tbody>
</table>
amPGLdO functions as a non-competitive inhibitor, amPLKG and cPGALK show competitive inhibition, and cPGSGO fits the mixed inhibitor category when tested with calpain-2. The B27 WT peptide inhibitor based on calpastatin was used as a control to ensure accurate data analyses, and showed competitive inhibition in the range of previously reported values[136].

Despite initial efforts to design competitive macrocyclic peptide inhibitors with specificity for calpain through interactions with the C2L domain, the four most potent inhibitors showed a mix of competitive, non-competitive, and mixed mode inhibition across both macrocycle and peptidomimetic compounds. Two of the four most potent inhibitors are macrocyclic compounds. Ideally, the macrocycle conformation could help form a calpastatin-like structure, while the residues leucine and lysine, responsible for interactions between calpastatin and calpain, could be key ones for inhibition by cyclic peptides. Macrocyclic compounds without those crucial residues might then result in mixed inhibition, as is the case with cPGSGO. The amidine-containing peptidomimetics on the other hand display a novel peptide structure for calpain inhibition, yet there is seemingly little information to be gathered as to a pattern for preferred residues. It is possible that, as is the case for cPGSGO, the peptide sequence can change the mode of inhibition and thus the binding site of the peptide. Investigating these amidine-based peptidomimetics compounds further by structural studies may in the future allow for more definitive comparisons and conclusions.
One potent peptidomimetic of all analyzed study compounds was the competitively-inhibiting amPLKG with a $K_i$ of $12.3 \pm 0.8 \mu$M. $K_i$ values were calculated over IC$_{50}$ values in order to assess affinity for the enzyme as well as to determine the inhibition mechanism. The level of inhibition seen with amPLKG is approaching the range of potency found with other peptidomimetic competitive inhibitors of calpain[149,226], which have been shown to inhibit calpain with nM IC$_{50}$s. Most calpain inhibition studies report IC$_{50}$ values instead of $K_i$ values. The former are easier to compute but are less rigorous than the latter. Although it is difficult to equate one to the other, as the relationship between IC$_{50}$ and $K_i$ depends on inhibition mechanism as well as reaction conditions[227,229,230], the larger macrocyclic inhibitors of calpain reported here are within the same order of magnitude as amPLKG[169,228], and show some selectivity for calpain (Figure 3.3). Interestingly, amPLKG as the most potent compound in this series retains the conserved residues of the calpastatin $\beta$-turn structure, but it remains to be seen from an enzyme-inhibitor complex structure if this is the reason for the potency especially in an amidine form.

3.4.5 Molecular docking of cPGALK NMR structures to calpain-2

Because the macrocyclic peptide cPGALK was found to competitively inhibit calpain-2, we modelled its inhibition by docking studies. cPGALK was modelled first with Macromodel (simulated annealing and MCMM with the OPLS-2005 force field) and we considered the 30 lowest energy structures that fit the NMR data. To accurately sample
interaction of the side-chains with the protein while accounting for the hydrogen-bonding network across the cycle backbone itself, all 30 NMR structures were docked individually with torsional constraints on the backbone and peptide bonds while keeping the side chains flexible. AutoDock4 and Glide provided the same structures as the most favourable poses. A representative top pose from Glide, ranked by lowest free energy of binding, can be seen in Figure 3.6. The agreement between two different docking programs and protocols gave support to the binding of cPGALK at the unprimed side of the active site cleft, consistent with the experimentally observed competitive inhibition. The vast majority of top poses from both computational suites were slight variations of this structure, either by a slight tilt in the central macrocycle or by slight movement of the side-chains, but remained nonetheless very similar. Many key contacts were formed between cPGALK and the calpain-2 structures, including five hydrogen bonding contacts as shown (Figure 3.6).

Docking results using cPGALK conformers with NMR considerations were more reliable than those of peptides starting from unbiased conformational sampling, lending more credence to the newly docked conformation of cPGALK. Thus, the determination of NMR solution structures for more compounds may help to obtain more plausible docking results in the future. Further modifications to residues pointing toward pockets on the C2L domain may increase the number and quality of interactions, and increase inhibition potency and specificity. This would allow for further assessment of the docking procedure and validation of docking results.
Figure 3.6: Representative lowest binding energy docking pose of cPGALK 
docked to calpain-2. The top docking pose from Glide, classified from the calculated 
free energy of binding by Prime MM-GB/SA, of cPGALK (residues labelled) bound to 
Ca$^{2+}$-activated calpain-2 (originating from PDB: 3BOW) seen along the active site cleft 
from the unprimed side. Calpain-2 is shown in surface representation with residues key 
for hydrogen bonding to the inhibitor in stick format and labelled. Hydrogen bonding 
contacts are labeled with cyan dashes. N and O atoms of both calpain-2 and the 
peptidomimetic are coloured blue and red, respectively.
3.4.6 Improvement of potency and selectivity of the peptidomimetic amPGLdO

True non-competitive inhibitors of calpains have been scarce, but are highly desired due to the potential for designing increasingly calpain-specific inhibitors by accessing an allosteric site. As a second generation of peptidomimetic compounds, modifications were made to the non-competitive inhibitor amPGLdO, varying both the Leu and d-Orn residues, to investigate potential improvements in both potency and specificity for calpain-2 over other similar cysteine proteases. Through initial screening of calpain-2, calpain-1, cathepsin L, and papain, using the fluorogenic substrates for each enzyme, four compounds were found to increase both potency and specificity of inhibition of calpain-2: amPGIdO, amPGVdO, amPGLdDab, and amPGLdQ (Figure 3.7). The four compounds showed significantly higher inhibition of calpain-2 over cathepsin L. For example, amPGVdO was almost twice as inhibitory against calpain-2 as amPGLdO (79% vs 44% inhibition, respectively) while showing half the inhibition of amPGLdO against cathepsin L (17% vs 32%, respectively). Attempts were made to further improve inhibition by combining the sequences of amPGIdO and amPGVdO with amPGLdDab to form amPGVdDab and amPGIdDab. Inhibition increased for calpain-1. However, there was essentially no change with respect to calpain-2 inhibition levels. The improvements so far can likely be attributed to a preference for branched hydrophobic residues at the 3rd position and size-specific charged residues at the 4th position, although this does not explain the plateau of inhibition seen for the merged sequences. More advanced conclusions will need to be drawn in the future from the structure of an amidine peptidomimetic bound to calpain.
Figure 3.7: Inhibition of cysteine proteases by variants of amPGLdO. (Caption on following page.)
Figure 3.7: Inhibition of cysteine proteases by variants of amPGLdO. The hydrolysis of fluorescent substrate by calpain-2 (black), calpain-1 (light grey), cathepsin L (white), and papain (dark grey) in the presence of second generation calpain-2 inhibiting peptidomimetics. The original sequence amPGLdO was varied at (A) d-Orn, (B) Leu, and (C) both d-Orn and Leu. Highlighted with a grey dotted box is the original compound amPGLdO in all three graphs. Amino acids ornithine (Orn), 2,4-diaminobutyric acid, and 2,3-diaminoproprionic acid are represented as O, Dab, and Dap, respectively. All compounds were screened at 100 μM. Error bars represent the standard deviation resulting from triplicate assays. Inhibition was determined as a percentage of the initial reaction velocity for the control reaction containing equivalent amounts of DMSO.
The four improved third-generation compounds seemingly displayed, in some cases, specificity for calpain-2 over calpain-1. At 100 μM, amPGIdO inhibited calpain-2 to ~ 67% but failed to inhibit calpain-1 in a significant way (~ 14%). Assuming this compound acts noncompetitively like its parent compound amPGLdO, it is possible that branched hydrophobic residues at R3 are responsible for interaction in a small allosteric pocket against calpain-2 that may be smaller on the surface of calpain-1. This would be consistent with the ability of the small side-chain alanine variant amPGAdO to inhibit calpain-2 and calpain-1 equally well (~ 66% and ~54%, respectively), and the reduced inhibition by the bulkier variant amPGFdO (~ 39% and ~ 28%, respectively). One can imagine non-competitive and mixed inhibitors, specifically amPGLdO and cPGSGO, could bind to small patches on calpain that vary slightly between calpain-2 and calpain-1, although the mechanism behind allosteric inhibition of calpains and the possibility for truly isoform-specific inhibitors remain unknown.

Additionally, looking back to first-generation compounds, cPGALK and cPGSGO also showed some specificity for calpain-2 over calpain-1 (respectively, cPGALK: ~ 84% and ~ 45%, cPGSGO: ~ 78% and ~ 38%). It is uncertain why there could be differential inhibition between calpains with these peptidomimetics because of the high sequence and structural homology of calpains-1 and -2, especially in the active site cleft for competitive inhibitors.
3.5 Conclusion

Potent, calpain-specific inhibitors are increasingly sought as tools for studying the physiological roles and substrates of calpain, and as treatments for such conditions as traumatic brain injury, Alzheimer’s disease, and heart attack and stroke. The advent of calpastatin-bound calpain structures allowed for the synthesis of calpain-specific inhibitors based on structural features of the enzyme-inhibitor complex. Small cyclic peptides presented here demonstrated an ability to function as protease inhibitors against calpain, papain, and cathepsin L. Four main lead compounds, cPGALK, cPGSGO, amPLKG, and amPGLdO, displayed competitive, non-competitive, or mixed inhibition of calpain with $K_i$ values in the mid- to low μM range, which are comparable to those of other calpain peptidomimetic inhibitors reported in literature. Two of the compounds, the competitively inhibiting cPGALK and amPLKG, showed specificity for calpain over cathepsin L. In addition, there is promise for exploiting non-competitive inhibition of calpains to achieve high specificity. Variations of the amPGLdO sequence improved the inhibition slightly, and greatly increased specificity for calpain over cathepsin L. Some of the cyclic peptides and peptidomimetic inhibitors shown here do not seem to fully inhibit calpain as competitive mimics of the conserved calpastatin β-turn motif. Instead they act with mixed or non-competitive inhibition. Further binding mode studies as well as structural work to locate the binding site will serve as a starting point for improvement and rationalizing the structure-activity relationship. As a novel
class of calpain inhibitors, the goal is now to further improve both potency and specificity of these peptide inhibitors and to structurally determine their binding sites on calpain.
Chapter 4

Development of $\alpha$-helical calpain probes by mimicking a natural protein-protein interaction

Supplementary Information can be found in Appendix A.

4.1 Abstract

We have designed a highly specific inhibitor of calpain by mimicking a natural protein-protein interaction between calpain and its endogenous inhibitor calpastatin. To enable this goal we established a new method of stabilizing an $\alpha$-helix in a small peptide by screening twenty-four commercially available crosslinkers for successful cysteine alkylation in a model peptide sequence. The effects of crosslinking on the $\alpha$-helicity of selected peptides were examined by CD and NMR spectroscopy, and revealed structurally rigid crosslinkers to be the best at stabilizing $\alpha$-helices. We applied this strategy to the design of inhibitors of calpain that are based on calpastatin, an intrinsically unstable polypeptide that becomes structured upon binding to the enzyme. A two-turn $\alpha$-helix that binds proximal to the active site cleft was stabilized, resulting in a potent and selective inhibitor for calpain. We further expanded the utility of this inhibitor by developing irreversible calpain family activity-based probes (ABPs), which retained
the specificity of the stabilized helical inhibitor. We believe the inhibitor and ABPs will be useful for future investigation of calpains, while the crosslinking technique will enable exploration of other protein-protein interactions.

4.2 Introduction

The primary goal of this work was to design and synthesize α-helical inhibitors as well as activity-based probes of human calpain, a calcium-regulated cysteine protease involved in a myriad of normal and pathological biological processes[231-242]. Although there has been considerable interest in the design of α-helical peptides for the study of protein-protein/receptor-ligand interactions and drug design, to our knowledge, there has been no work to date investigating α-helices as protease inhibitors.

Inhibitor design for this class of enzyme has historically focused on the use of peptidomimetics that fit into the active site cleft in a substrate-like manner and utilize covalent, reversible or irreversible reactive groups to react with the active site cysteine[117,119,158,243-246]. The problems with this approach are twofold: 1) the papain superfamily has a highly conserved active site cleft, which complicates identification of peptidomimetic side chains that differentially bind to individual enzymes, and 2) small peptides do not bind well to calpains.
To overcome this problem we took inspiration from the recent co-crystal structure of calpain with its endogenous protein inhibitor, calpastatin and from calpain inhibitors containing constrained scaffolds or macrocycles[97,98,169,204,247]. Calpastatin is unstructured in solution; however, upon binding to active calpain it drapes across the entire protein and undergoes structural rearrangements to form three α-helices that contact three different domains of the enzyme. One of these α-helices binds adjacent to the prime side of the active site cleft (Figure 4.1), forming a number of energetically favorable interactions between apolar sidechains that become buried upon complex formation. We therefore hypothesized that this α-helical motif would provide increased specificity via its unique binding mode since the helix avoids the highly conserved region of the active site while still inhibiting substrate access to the active site cleft.

This two-turn α-helix represents a ten-residue peptide. Previous work indicated that small peptides were poor inhibitors of calpains[39,124]. We corroborated this idea by determining that the minimal calpastatin fragment peptide that formed the two-turn α-helix (IPPKYRELLA) did not inhibit calpain (Kᵢ >100 µM). We reasoned that the entropic cost of forming an α-helix from a random coil limited the ability of small peptides to inhibit the enzyme; thus we decided to design a stabilized version of this peptide to minimize unfavorable conformational entropy.

Several strategies have previously been developed for α-helix stabilization involving main- or side-chain modifications including: disulfide bond formation[248-250],
Figure 4.1: X-ray crystal structure of the calpain 2-calpastain complex (PDB ID: 3BOW). Key residues on the inhibitor, calpastatin, (purple) and calpain-2 (black) are labeled.
hydrogen bond surrogates[251,252], ring closing metathesis[253-256], cysteine alkylation using α-haloacetamide derivatives[257] or biaryl halides[258], lactam ring formation[259-265], hydrazone linkage[266], oxime linkage[267], metal chelation[268,269], and “click” chemistry[270,271]. Of the different methods used to stabilize these structures, the inclusion of a semi-rigid cross-linker[272-278] has been particularly successful, and is explored herein.

4.3 Results and Discussion

4.3.1 Design of template-constrained cyclic peptides stabilizing an α-helix conformation.

Peptides are intrinsically flexible chains, which rapidly interconvert among a large ensemble of conformations, including canonical secondary structures (α-helices, reversed turns, β-hairpins, etc.). Generally, only one of these conformations is required to bind a given receptor/enzyme, and very large changes in affinity (>10^4) can be realized by simply restricting the structure to a single conformational state.

We were particularly interested in conformational restriction via cysteine alkylation[279-282] for its chemical stability, selectivity, cost effectiveness, and ease of introduction via standard mutagenesis into recombinantly expressed peptides or proteins or by solid-phase peptide synthesis. Importantly, a number of structurally diverse thiol
reactive crosslinkers are also commercially available. Thus, we envisioned that the bioactive conformation of a given peptide could be stabilized by identification of the optimal cysteine crosslinker from screening a library of crosslinkers on a peptide with two cysteines anchored in appropriate positions. We refer to $\alpha$-helical peptides stabilized in this manner as template-constrained peptides.

Figure 4.2 (left) shows the fundamental concept of template-constrained cyclic peptides, in this case accomplished via sidechain-to-sidechain cyclizations. To do this, a pair of cysteine residues is installed at appropriate positions in order to stabilize a local conformation. Here, we placed the cysteine residues at $i$, $i+4$ positions, because this spacing brings two thioether residues into proximity when in the $\alpha$-helix. In a series of parallel reactions we react the peptide with an indexed array of different crosslinking agents. Bis-alkylators with sufficient reactivity to alkylate thiols will cleanly form cyclic peptides, if the macrocycle can be formed in a low-energy conformation that matches one of the low-energy conformations of the peptide. For example, a meta-xylyl group, which matches the inter-thiol distance of the cysteine sidechains when in an $\alpha$-helical conformation, should stabilize this helical structure. By contrast, the much longer distance of the 4,4'-biphenylmethyl group would not be consistent with the $\alpha$-helical conformation, and would instead favor formation of a more extended conformation. Thus, depending on the template, it should be possible to stabilize any one of a number of conformations.
Figure 4.2: Conformational restriction via crosslinking (left). Kinetic “selection of the fittest” reaction. Hypothetical rate constants are denoted by $k_1$, $k_2$, and $k_3$ (right).
We use a kinetic “selection of the fittest” method, to screen for only those linkers that help select stable, low-energy conformations over more strained conformations. The kinetic scheme for cyclization requires two steps (Figure 4.2, right): The first step involves the second-order alkylation of the dithiol-peptide, which depends on the concentration of both the alkylating agent and the peptide (rate \( 1 = k_1 \) [peptide\(_{(SH)2}\) [alkylator]]. The rate of this reaction depends on the chemical nature of the alkylator, but to the first approximation is largely independent of the peptide structure, which is largely in a random coil in the linear form. Once mono-alkylated, the second-order process of reacting with a second equivalent of the alkylating agent (rate \( 2 = k_2 \) [peptide\(_{(SH)1}\) [alkylator]]) will compete with the desired first-order cyclization process (rate \( 3 = k_3 \) [peptide\(_{(SH)1}\)]). (Solvolysis reactions of the mono-alkylated product also compete with cyclization.) The cyclization reaction depends on the ability of the peptide to reach a stable, strain-free conformation as it enters the transition state for cyclization, which we presume is geometrically similar to the product for large macrocyclic rings such as those formed here. Thus, the ratio of bis-alkylated to mono-alkylated compound provides a quantitative measure of the ease of cyclization that is dependent on the conformation of the cyclic form of the peptide. Bis-alkylation is dependent on the concentration of the peptide while cyclization is independent of this parameter, therefore it is possible to select for the most efficient crosslinkers by simply running the reaction at a fixed peptide concentration with increasing concentrations of bis-alkylators and examining the product distribution by mass spectrometry.
In summary, the current method of template-constrained thioether cyclization involves several steps: 1) Screening for cross-linking agents with appropriate reactivity and ability to form cyclic products under favorable conditions with nearly equimolar amounts of peptide and bis-alkylator. 2) Examining bis-alkylator “hits” with increased stringency, using higher molar concentrations of alkylators in large excess of the peptide. This step should provide template-constrained peptides with relatively strain-free conformations. 3) Testing the template-constrained peptides to determine which have been stabilized in the appropriate conformation. This can easily be accomplished by circular dichroism (CD) spectroscopy for an α-helix. 4) Finally, determining the impact of stabilizing the helix on the ability of the peptide to bind to a protein known to recognize the sequence in a helical conformation.

To explore template-constrained cyclization to stabilize α-helices in aqueous solution, we used the model peptide 1 (sequence: Ac-YGGEAAREACARECAARECONH₂) which was similar to the FK-4 peptide previously described (Appendix A – Supporting Information Table A.1) [283]. The model peptide exhibited a low to moderate level of helicity without any stabilization.

We screened twenty-four crosslinkers for cys-thioether macrocyclizations. The crosslinkers included alkyl bromides c1-c6, c12, and c13, alkyl iodides c7-c11, benzyl bromides c14-c20, allyl bromide c21, maleimides c22 and c23 and an electrophilic difluoro-benzene c24 (Figure 4.3). The initial screening reaction was performed in a 96-
Figure 4.3: Helix stabilization via screening of 24 crosslinkers.
well plate format to identify crosslinkers that react with cysteine thiols under mild conditions (bicarbonate buffer, pH = 7.5 to 8.0) at room temperature. The crude reaction mixture was analyzed by MALDI-TOF mass spectrometry to identify any crosslinker that was a “hit”. Additional HPLC profiling can characterize product distribution.

Product distribution was analyzed using MALDI-TOF and revealed that cysteine alkylation did not occur when simple alkyl halides c1-c12 were used; only intramolecular disulfide bond formation due to oxidation was observed to occur[284]. Even when the leaving group was changed from bromide to the more reactive iodide c7-c11 alkylation reactions failed under these aqueous conditions. The crosslinking reaction with 1,4-dibromo 2,3-butanedione c13 produced a complex mixture of products. Crosslinking reactions with the maleimide crosslinkers c22-c23 also resulted in a mixture of epimeric products that were further complicated by hydrolysis of the imide (Appendix A – Supporting Information Figure A.1). Reactions using 1,5-difluoro-2,4-dinitrobenzene c24 resulted in a similar complex mixture of products. For the biaryl derivatives c17, c18, predominantly unreacted peptide was detected (MALDI-TOF and HPLC) accompanied by traces of the desired, cyclized product (Appendix A – Supporting Information Figure A.1 and Figure A.2).

The cleanest macrocyclization resulted from the reaction[285,286] with benzylic/allylic halides c14-c16 and c19-c21, which provided the major peak of the cyclization product as seen by MALDI-TOF and HPLC trace analysis (Appendix A –
Supporting Information Figure A.1 and Figure A.2). We then tested the crosslinker “hits” c14-c16 and c19-c21 under the conditions designed to increase the rate of bis-alkylation over cyclization (by increasing the concentrations of alkylating agent and peptide in solution). HPLC analysis of the “selection of the fittest” showed that the 1,3-bis(bromomethyl) benzene (α,α’-dibromo-m-xylene) crosslinker c15 and 2,6-bis(bromomethyl)pyridine crosslinker c20 gave the cleanest formation of the desired macrocycle (Appendix A – Supporting Information Figure A.3). By contrast, crosslinking with allyl crosslinker c21 produced multiple peaks. It is interesting that the m-xylene crosslinker c15 was most successful crosslinker out of the three α,α’-dibromoxylanes c14-c16, considering that all the three alkylating agents have relatively different reactivity profiles (ortho>meta>para) [280].

We next evaluated the CD spectra of these selected template constrained cyclic peptides to determine the effect of the template on their coil-helix equilibria (Figure 4.4). The determination of secondary structure was complicated somewhat by the fact that the spectra are generally interpreted using the intensity of θ222, which requires knowledge of the concentration[287], generally by measuring the absorbance of an N-terminal Tyr residue. Some of our linkers contain aromatic groups that could absorb at 278 nm and complicate concentration determination. Therefore, we use dry weight to estimate the concentration, which results up to a 25% error in concentration determination (assessed by comparing gravimetric versus spectrophotometric determination of peptides containing Tyr chromophores and lacking other groups). Because θ222 is not accurately
Figure 4.4: CD spectra of the model peptide and the crosslinked peptides in phosphate buffer [50 mM, pH = 7.0, 25 °C].
measured, we therefore interpret the data largely based on the shape of the spectra, particularly the ratio of the peak shape and relative intensities of the two exciton-coupled π-π' bands at 190 nm and 208 nm relative to that of the n-π' band near 222 nm[288]. The three xylene-based crosslinkers c14-c16 all showed an increase of the helicity in the CD spectroscopy analysis. Notably, the m-xylene based crosslinker c15 showed the most increase in helicity followed by o-xylene c14 and finally p-xylene c16.

Interestingly, the CD spectrum of the crosslinked peptides by crosslinkers c17 and c21 showed some structural differences from those seen using the xylene crosslinkers. As expected, the 4,4'-biphenyl (c17) crosslinked peptide showed little helicity, likely due to destabilization of the α-helix and stabilization of an extended conformation of the peptide because the end-to-end length of the biphenyl template is much longer than the typical α-helix pitch. Likewise, peptide crosslinked with the butenyl derivative c21 showed a CD spectrum with a deep minimum near 200 nm, similar to that of the random coil (Figure 4.4). It would be interesting to test whether this peptide, after the reduction of the double bond, could stabilize a 3_10 helix as shown in the Grubbs’s work[255]. This crosslinker could be an alternative to ring closing metathesis (RCM) stapling and subsequent double bond reduction strategy.

Heterocyclic templates were also capable of stabilizing the α-helix. 2,3-quinoxaline c19 and 2,6-pyridine c20 crosslinked peptides showed CD spectra similar to those of the o-xylene c14 and m-xylene c15 crosslinked peptides (Figure 4.4).

113
NMR spectroscopy experiments demonstrate that the cyclic template restraint strongly stabilized the helical conformation within the macrocyclic ring, and that the helix extended towards the C-terminus of the peptide (Figure 4.5). Typical stepwise NH(i)/NH(i + 1) NOE connections were observed from the first residue to the last residue, which are indicative of a helical conformation. Closer inspection showed that the cross-peak intensity became stronger after the residue 6, suggesting that the crosslinked region in the helix was more organized than frayed region of the N-terminus, which included two glycines. Furthermore, $^3J_{\text{NH-HA}}$ coupling was evaluated by the INFIT (inverse Fourier transformation of in-phase multiplets) procedure[289]. The J coupling constant is a good indicator of secondary structure. It is generally averaged to ~7 Hz if the residue is in a random coil or in equilibrium between different structures. It is less than 6 Hz if it is in $\alpha$-helical structure and is larger than 8 Hz if the secondary structure is a $\beta$-sheet. Our J coupling constant was mostly below 6 Hz suggesting an $\alpha$-helical structure. In addition, the chemical shift index of $\alpha$-H strongly demonstrated helix formation even in the fraying N-terminus. Secondary chemical shifts which were calculated by subtracting the experimental values from the intrinsic values and clearly showed the effect of the crosslinker. The most dramatic changes were observed on Cys10, Ala11, Arg12 and Cys14, influenced in part by the anisotropy effect from the benzene ring in the crosslinker (Appendix A – Supporting Information Figure A.4).
Figure 4.5: NMR of m-xylyl c15-constrained cyclic peptide (left). NOE sequential walk of backbone amide region of NOESY (250 ms) for the peptide. The cross peaks are labeled as NH($i$)/NH($i + 1$) $^3J_{\text{NH-HA}}$ coupling as function of residue (right). The small $^3J_{\text{NH-HA}}(<6\text{Hz})$ and strong sequential NH-NH NOEs denote helix formation in the peptide.
4.3.2 Application of $i, i+4$ $m$-xylene crosslinker-based stabilization for calpain inhibitor design.

Turning back to calpain inhibitor design we chose to use the calpastatin fragment IPPKYRELLA (previously shown to be inactive against calpain) as the backbone since this sequence, in the context of full-length calpastatin, forms a two-turn helix in the primed side of the active site of calpain-1 as shown in Figure 4.1. Three different sets of double cysteine mutants, $3a-c$, along with their $m$-xylene crosslinked partners, $3a-c$, were synthesized (Figure 4.6, Appendix A – Supporting Information Table A.3). Cysteine locations were chosen by both visual inspection and virtual alanine scanning mutagenesis (Appendix A – Supporting Information Table A.2) so as not to disturb key interactions at the protein-helix interface, which includes Pro51 (inhibitor) ring stacking against Trp288 (calpain) and Tyr54 (inhibitor) H-bonding to His169 (calpain) as shown in Figure 4.1.

Next, the difference in structural changes as a result of cysteine crosslinking was examined via CD spectroscopy (Figure 4.7) [287,291]. The helical content of the uncrosslinked peptides was low in the absence of added trifluoroethanol (TFE), so the experiments were conducted in the presence of 40% TFE[292]. CD analysis revealed a clear trend whereby all unlinked peptides showed little secondary structure, while the crosslinked peptides demonstrated varying degrees of a-helicity. Peptide $3c$ showed the greatest helicity after crosslinking, followed by $3b$, while $3a$ showed negligible helicity after crosslinking. The lack of increased helicity for $3a$ may be due to the fact that it lacks
Figure 4.6: Sequence of double cysteine mutants (3a, 3b, and 3c) and their crosslinked counterparts (3a, 3b, and 3c) (left). A helical wheel representation to indicate the crosslinked regions (right)[290]. \( \overline{\text{ab}} \) denotes the m-xylyl c15 crosslinking between the cysteines.
Figure 4.7: CD spectra of uncrosslinked peptides 3a-c (top) and crosslinked peptides 3a-c (bottom), [~125 μM peptide, 50 mM Tris (pH 7.5), 40% TFE]. Crosslinked peptide 3c demonstrates the greatest helical content. (See Appendix A – Supporting Information Figure A.5 and Figure A.6 for CD analysis without 40% TFE.)
the proline that is frequently found as an helix initiator of an $\alpha$-helix[293]. A possible salt bridge between the glutamic acid and lysine may also be enhancing helical content in 3c[294-296]. Thus, we believe that the primary sequence of the peptide as well as the crosslinker can influence the final helical content of the product peptide.

The inhibitors, both crosslinked and uncrosslinked, were tested for their ability to inhibit calpain-1 (Table 4.1, Appendix A – Supporting Information Figure A.7 and Figure A.9). No appreciable inhibition ($K_i > 100 \mu M$) of calpain-1 was observed for the uncrosslinked peptides 3a-c. These results corroborate previous reports stating that the minimum length of a standard calpastatin derived peptide needed to achieve reasonable calpain inhibition is 27 amino acids long[290,297]. However, the crosslinked peptide, 3c, which is only 10 amino acids long, showed good inhibition of calpain-1 in the low micromolar range (Table 4.1, Appendix A – Supporting Information Figure A.9). Furthermore, a trend relating higher helical content (Figure 4.7) positively correlated with better inhibition of calpain-1 (Table 4.1). This trend is likely directly related to helical content stabilized by the crosslinker c15, although it is also possible that the crosslinker itself could contribute to enzyme recognition of the inhibitor.

Kinetic studies were then performed to understand the mechanism of 3c inhibition of calpain-1; standard Michaelis-Menten and Lineweaver-Burke analysis showed that 3c behaved as a competitive inhibitor (Figure 4.8, Appendix A – Supporting Information Figure A.10 and Table A.4). These results are consistent with the idea that
Table 4.1: $K_i$ against calpain-1[229]. The calpain assay was done as described in Materials and Methods.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>3a</th>
<th>3b</th>
<th>3c</th>
<th>3a</th>
<th>3b</th>
<th>3c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calpain-1 (µM)</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>95.6 ± 25.5</td>
<td>10.2 ± 2.9</td>
</tr>
</tbody>
</table>
Figure 4.8: Lineweaver-Burke analysis shows that calpain inhibitor 3c to be a competitive inhibitor. Lineweaver-Burke plot was constructed from standard Michaelis-Menten kinetics.
3c binds to the α-helix binding site in the primed side of the active site of calpain and physically blocks substrate binding, and subsequently proteolysis, as predicted from the initial co-crystal data (Figure 4.1).

There has been considerable difficulty in achieving good selectivity within the papain superfamily of enzymes as these enzymes contain highly conserved active sites[97,298]. To determine whether the helical inhibitor 3c was specific for calpain we tested it against a set of canonical papain family cysteine proteases including: papain, cathepsin B and cathepsin L (Table 4.2, Appendix A – Supporting Information Figure A.11). Significantly, no inhibition (K_i > 100 µM) was observed using the crosslinked peptide 3c against papain or cathepsin B. The inhibitor was about four fold more potent against calpain over cathepsin L (K_i = 39.9 ± 1.09 µM). These results indicate that this α-helical motif may represent a uniquely selective binding element for inhibition of calpains and further validates our structure-based approach. Furthermore, structure activity relationship studies of these helical inhibitors may result in a more potent and specific inhibitors of calpain and also shed some light on to how the calpastatin helix interacts with human calpains.

The crosslinking reaction was performed with the crosslinker c15 and the three peptides in aqueous buffer system. However, in instances where there are multiple cysteines, we believe that solid-phase cysteine crosslinking could be useful for selective crosslinking. To this end, we tested the on-resin crosslinking the peptide 3c.
Table 4.2: The $K_i$ of crosslinked inhibitor $3c$ against other papain family proteases.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Calpain-1</th>
<th>Papain</th>
<th>Cathepsin B</th>
<th>Cathepsin L</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3c$ (µM)</td>
<td>10.2 ± 2.9</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>39.2 ± 1.1</td>
</tr>
</tbody>
</table>
Fmoc-Cys(Mmt)-OH was used instead of Fmoc-Cys(Trt)-OH and selective deprotection of specific cysteine side chains was achieved by 1% TFA/DCM treatment while the peptide was still resin bound[299,300]. (See the Materials and Methods). The same kinetic results were achieved with resin cross-linked inhibitor.

Based on our initial success with a stabilized, α-helical-based inhibitor of calpain we next endeavored to develop an activity-based probe (ABP) specific for calpains. ABPs are complementary chemical tools to traditional genomic and proteomic techniques; ABPs are used for identification of enzymatic targets and to evaluate dynamics of enzyme activity regardless of levels of expression[301-306]. This is important because in many cases translation and transcription do not correlate with enzyme activity[307]; this is especially true for calpains as their proteolytic activity is finely regulated post-translationally by intracellular calcium levels. Basic ABP design includes a mechanism based inhibitor, a specificity element, and a tag (Figure 4.9, top). In this case, the crosslinked peptide 3c was used for the specificity element and the succinyl epoxide functions as the warhead group that reacts with the cysteine thiol. This warhead has been established to react in a mechanism dependent manner only with active papain family proteases[308]. Three dipeptide linkers (NM-01, 02, and 03) of different lengths and rigidities were chosen via visual inspection in PyMOL [360] based on the crystallographic structure of calpastatin-bound calpain 2 (PDB code 3BOW) [97]. Lastly, we chose to use either biotin or fluorescein isothiocyanate (FITC) as a tag.
Figure 4.9: Design of a calpain-specific ABP (top). ABPs contain a mechanism based inhibitor, specificity element, and tag. Only the chemical structures ABPs containing a biotin tag are shown here. ━ denotes the m-xylyl c15 crosslinking between the cysteines. ABP binding to calpain-1 (bottom). The linker length and rigidity between the crosslinked peptide and succinyl epoxide was evaluated via reaction with calpain-1 in vitro. A five-carbon backbone, flexible linker appears optimal. Loading control lanes beneath the panel show Western blot analysis using anti-calpain-1.
We used three different amino acid sequences as linkers: alanine-alanine, β-alanine-alanine, and alanine-β-homoproline, (NM-01, NM-02, and NM-03, respectively) (Appendix A – Supporting Information Table A.5). NM-01 is the shortest linker by one carbon but has similar flexibility as NM-02. NM-02 and NM-03 should cover a similar distance between the helix and succinyl epoxide, however the β-homoproline provides more rigidity than the β-alanine.

To evaluate the best linker, we initially tested biotinylated versions of either NM-01, -02, or -03 on purified, activated calpain-1 at two concentrations, 1 and 10 µM, and on unactivated calpain at 10 µM (Appendix A – Supporting Information Figure A.9, bottom). Each ABP was added to purified calpain (pH 7.0), followed by the addition of calcium to activate the enzyme. The probe was allowed to react for 20 min. at room temperature. No calcium addition was used as a control to demonstrate that labeling only occurred with active calpain, and DCG-04, a pan-papain family cysteine protease ABP[308], was used as a positive control as it is known to label calpains. Samples were analyzed by SDS PAGE electrophoresis; proteins were transferred to PVDF membrane and analyzed by western blot for biotin using streptavidin-HRP. Our results show that two ABPs, NM-02 and NM-03, labeled calpain in an activity dependent manner, which indicated that an extra carbon in the amino acid backbone of the linker was necessary for the epoxide to react with the active site cysteine (Figure 4.9). The intensity of the bands in the blot suggested that the use of the linker β-alanine-alanine resulted in the most potent probe (NM-02) (Figure 4.9, bottom). The ABP with the alanine-β-
homoproline linker (NM-03) also bound to calpain but the rigidity in the linker induced by the pyrrolidine ring in homoproline may have contributed to less labeling. These results further support our hypothesis that the helix is binding at the active site as measurements of the probe visualized in PyMOL [360] show that a β-alanine-alanine linker would position the epoxide at the correct distance from the active site cysteine.

The presence of the succinyl epoxide warhead could reduce the specificity of the inhibitor due to its reactivity against most papain family active site cysteines. However, based on the previous kinetic studies, we reasoned that if the crosslinked peptide bound to the enzyme followed by a covalent reaction between the warhead and the active site cysteine, the ABPs had a high probability of being specific for calpain despite the addition of this reactive warhead. To investigate the specificity of NM-02, we tested a FITC tagged NM-02 against calpain-1 and calpain-2, and a panel of papain family proteases including papain, cathepsin B, and cathepsin L (Figure 4.10). FITC-NM-02 was added in increasing concentrations to either papain, cathepsin B, or cathepsin L and allowed to react for 20 min. at room temperature. Labeled enzymes were analyzed by SDS-PAGE and were visualized using a flatbed fluorescent scanner (Typhoon). We found that even at 10 µM, NM-02 did not bind to any of the other papain family cysteine proteases, which was in good agreement with the Ki (Table 4.2) determined in the binding studies of the crosslinked peptide 3c. This further suggests that NM-02 is
Figure 4.10: FITC-NM-02 as a calpain specific ABP. We tested FITC-NM-02 (probe) in vitro against purified calpain-1, calpain-2, papain, cathepsin B, and cathepsin L. Only active calpain-1 and -2 are labeled and both are increasingly labeled with increased amounts of probe. Papain, cathepsin B, and cathepsin L are not labeled by NM-02. Loading control lanes beneath each panel show colloidal blue staining or silver staining of the respective gel.
specific for calpain at concentrations that would be appropriate for protease labeling experiments.

4.4 Conclusions

In summary, we have demonstrated a simple screening of inexpensive, commercially available crosslinkers on an $i, i+4$ double cysteine mutant peptide to identify the best crosslinker to stabilize an $\alpha$-helix. We identified five crosslinkers that increase $\alpha$-helical character. Out of these five crosslinkers, dibromo-$m$-xylene, c15, reacted in a simple, one-pot reaction, both in solution and on solid-phase, with the cysteine side chain and best increased the helicity of the peptide.

We have also applied this helix stabilization method to mimic a protein-protein interaction between a protease and its endogenous protein inhibitor to create, to our knowledge, the first active site directed, $\alpha$-helical inhibitor of a protease. Importantly, we demonstrate that this inhibitor is shows good potency and high specificity for calpains over other highly similar cysteine proteases.

Lastly, we show that we can use the $\alpha$-helical inhibitor as a scaffold to create an activity-based probe for examination of calpain activity. We determined that a $\beta$-amino acid is needed in the linker to bridge the gap between the helix and the active site
cysteine. Furthermore it appeared that the ABP, NM-02, retained specificity for calpains over closely related cathepsin proteases. Given this specificity, we hope that these inhibitors and probes will allow for future studies of calpain function in multiple biological systems. We believe that the methodology used to stabilize this $\alpha$-helical inhibitor will be another useful technique for $\alpha$-helix stabilization for use in multiple biological applications.

4.5 Materials and Methods

4.5.1 Crosslinker screen

To each well of a black round-bottomed 96-well plate (polypropylene) 90 $\mu$L of the stock solution, a peptide solution (0.114 mM) in NH$_4$HCO$_3$ buffer (12 mL, 50 mM, pH = 8.0), treated with TCEP (1 M solution in the same NH$_4$HCO$_3$ buffer, 1.1 eq.) at room temperature (rt) for 1 h was added. Then 10 $\mu$L of the freshly prepared alkylating agent solution (1.5 mM in anhydrous DMF, 1.5 eq.) was applied to the well at rt and stirred for 2 h under protection from light. MALDI spectra were taken to monitor reaction progress and more alkylating agent was added if needed. The reaction was quenched by addition of 5% HCl which resulted in acidic conditions (pH = 3-4). If necessary, 100 $\mu$L of ether was added to dissolve the excess reagent and organic byproducts into the organic layer. The ether layer could be removed by pipetting. MALDI spectra were taken from the sample in the remaining aqueous solution mixture.
4.5.2 “Selection of the fitness” Screen

Screens were performed in 1.5 mL microcentrifuge tubes. 1 mL of the stock peptide solution (1 mM) in NH$_4$HCO$_3$ buffer (50 mM, pH = 8.0) was pre-treated with TCEP as described above and incubated for 1 h. Then, 100 µL of the concentrated alkylation agent solution (250 mM or saturated solution in anhydrous DMF) was added and shaken for 2 h under protection from light. The reaction was quenched by the addition of 5% HCl which resulted in acidic conditions (pH = 3-4) and purified by Reverse Phase HPLC.

4.5.3 Crosslinking with the unpurified peptide

The lyophilized crude peptide solution (app. 3-5 mg/mL) in NH$_4$HCO$_3$ buffer (100 mM, pH = 8.0) was treated with TCEP (1.5 eq.) and stirred for 1 h. The alkylation agent in DMF (app. 3 eq) was added to the solution and shaken for the 2 h. The reaction was quenched by adjusting the pH of the mixture to slightly acidic conditions through the addition of 0.5 N HCl or TFA. The crude mixture was either purified by HPLC or lyophilized for the next step.

4.5.4 Preparation of crosslinked peptides 3c from model peptide 3c by solid-phase peptide crosslinking

The uncrosslinked peptide 3c was similarly prepared on the CLEAR™ Rink Amide MBHA resin using the standard Fmoc peptide synthesis protocol (See Appendix
A – Supporting Information). Fmoc-Cys(Mmt)-OH was used for cysteine for ease of deprotection. After the final coupling and cooling down to room temperature, the resin was washed with NMP(x3) and DMF(x3) followed by DCM(x3). The resin was then treated with 1% TFA solution in DCM for 10 min then washed with dichloromethane. This process was repeated until the solution lost its yellow color, which indicated the complete removal of Mmt protecting group. Then, the resin was washed with hexane and dried. After re-swelling in DMF, a solution of a,a'-dibromo-m-xylene (2 eq.) in DMF and DIPEA (4 eq) was added. Alternatively, the resin was re-swollen in NH₄HCO₃ buffer (pH = 8.0, 100 mM) for 1 h, a solution of a,a'-dibromo-m-xylene (5 eq.) in a minimal volume of DMF was added. The solution was stirred for 3 h at room temperature. The solvent was then removed and the resin was washed thoroughly with DMF. The Fmoc group on N-terminus was removed by treatment with 20% piperidine in DMF and acetylated by Ac₂O and DIPEA. The cleavage/deprotection was done using TFA/thioanisole/EDT/anisole (90/5/3/2). The crude mixture was purified by reverse phase HPLC.

4.5.5 CD spectroscopy

Peptide solutions were prepared at ~ 50 µM in 50 mM phosphate buffer (pH 7.0) without TFE. The molar concentration of the peptide determined was by the weight (after lyophilization of the HPLC fractions) with consideration for molecular weight increase due to the presence of TFA salt for basic residues (Lys, Arg) as well as hydration (average 10%). Concentrations of the uncrosslinked peptides were determined by absorbance of Tyr residue at 280 nm with an extinction coefficient of 1280 M⁻¹ cm⁻¹[309].
Circular dichroism studies were conducted at 25˚C on a JASCO J-810 spectropolarimeter equipped with a Peltier temperature control unit.

4.5.6 NMR spectroscopy

The peptide sample was prepared with peptide concentrations of 2 mM in 0.6 mL of 9:1 v/v water/D$_2$O mixture in 50 mM sodium phosphate, pH 5.5. All spectra were recorded at 10 °C on a Bruker Avance III 500 MHz spectrometer equipped with a cryogenic probe. All 2D homonuclear spectra were recorded with standard pulse sequences[310]. Spectra were processed and analyzed using the programs nmrPipe[311] and XEASY[312], respectively. (See Appendix A - Supporting Information.)

4.5.7 Protease activity assays

Peptides were evaluated for ability to bind and subsequently inhibit the cysteine proteases using standard proteolytic fluorescence activity assays. Inhibition was assayed using a standard donor-quencher strategy using a previously published peptide substrates[138,158,313].

Enzyme concentration for Calpain-1 was 25 nM. Enzyme concentration for papain was 25 nM. Enzyme concentrations for cathepsin B and cathepsin L was 3 nM. Calpain and papain buffer contained 10 mM dithiothreitol (DTT), 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), and 0.015% Brij-35. Substrate concentration for calpain and papain was 0.25 µM H-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH (K$_m$ calculation in Appendix A - Supporting Information Figure A.8 and Figure A.10)
Cathepsin buffer contained 10 mM DTT, 500 mM sodium acetate (pH 5.5), and 4 mM EGTA[138,158,313]. Substrate concentration for the cathepsins was 0.25 µM Z-FR-AMC. Calpain was activated by the injection of CaCl$_2$ to a final concentration of 5 mM. Papain and cathepsin assays were activated by the addition of the substrate via a multichannel pipette. Varying concentrations of inhibitor, 1-100 µM, were used for each assay. All assays were done at a total well volume of 100 µL in 96-well plate, and each well contained a separate inhibitor concentration. Fluorescence was read in a Berthold Tri-Star fluorimeter. The excitation wavelength was 380 nm and the emission wavelength was 500 nm for H-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH. The excitation wavelength 351 nm and emission wavelength was 430 nm for Z-FR-AMC.

4.5.8 Kinetic analysis of calpain-1 by 3c

To identify inhibition type we used standard Michaelis-Menten treatment. Initial velocities (obtained from the linear segment of the progress curves) were plotted against substrate concentration[314]. Due to the linearity of the first segment of the progress curve we believe that autoproteolysis during the first 500 seconds was not substantial enough to prevent the use of simple Michaelis-Menten kinetics, i.e. loss of enzyme did not change the velocity enough to cause it to deviate from linearity and incorporation of this additional complex would severely complicate the kinetics. Velocities were determined in RFU/sec then converted to µM/sec using the conversion factor 1386 RFU/µM. The conversion factor was obtained by the total hydrolysis of the substrate H-
Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH in a known concentration by papain. To avoid weighting errors we used the values of $K_{m}^{app}$ and $V_{max}^{app}$ determined directly from the non-linear least-squares best fits of the untransformed data and put these values into the reciprocal equation:

$$\frac{1}{v} = (\frac{K_{m}}{V_{max}} \times \frac{1}{[S]}) + \frac{1}{V_{max}} [314].$$

We then plotted the resulting reciprocal velocities against the respective reciprocal substrate concentrations.

**4.5.9 Determination of IC$_{50}$ against enzymes**

IC$_{50}$ curves were generated identifying the initial rate of the enzyme at each inhibitor concentration from the respective progress curves. The conversion factor (1386 RFU/$\mu$M) was obtained by the total hydrolysis of the substrate H-E(EDANS)-PLFAER-K(DABCYL)-OH in a known concentration by papain. Initial velocities were converted from RFU/sec to $\mu$M/sec. Fractional activity was calculated by dividing the initial velocity at each inhibitor concentration by the initial velocity of the uninhibited enzyme. Data obtained up to 500 seconds was used for the initial rate calculation. The initial rate was then plotted against the log of the inhibitor concentration, and IC$_{50}$ was calculated by GraphPad Prism.
4.5.10 Activity based probe linker experiments

Experimental conditions included 10 mM dithiothreitol (DTT), 1.5 µg calpain, 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), 0.015% Brij-35, and either 1 µM or 10 µM of biotinylated probe (DCG-04, NM-01, NM-02, NM-03). Calpain was activated by the addition of calcium (3.33 µM of 50 mM CaCl₂) to a final concentration of 8.3 mM in tubes containing either 1 µM or 10 µM ABP. For the negative control, water, instead of CaCl₂, was added to the calpain solution containing 10 µM probe. Probes were allowed to bind to the calpain for 20 minutes at room temperature. The reaction was stopped by the addition of 10 µL NuPage® LDS Running Buffer (Life Technologies, Grand Island, NY). 10 µL of each labeled enzyme was loaded on a 10% Bis-Tris NuPAGE® gel (Life Technologies, Grand Island, NY) and separated via gel electrophoresis for 1.5 h, 140 V. The bands were then transferred to a PVDF membrane at 30 V for 70 min. The membrane was blocked and blotted using the Vectastain® Elite® ABC kit (Vector Laboratories, Burlingame, CA). Kodak film was exposed to the membrane and developed.

4.5.11 ABP labeling experiments

Buffer conditions for calpain and papain experiments were 10 µM dithiothreitol (DTT), 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), and 0.015% Brij-35. 1.5 µg calpain-1 or 6 µg calpain-2 (calpain-2 was not as active) was used. (For labeling experiments greater concentrations of enzyme were used for ease of visualization of the enzyme on stained gels.) Buffer conditions for cathepsin experiments were 10 µM DTT,
500 mM sodium acetate (pH 5.5), and 4 mM EGTA. 1.5 µg of each cathepsin was labeled\[97,138,313\]. Probes were allowed to bind for 20 min. at room temperature. Labeled enzymes were separated via gel electrophoresis on 10% (calpain, papain) or 12% (cathepsins) Bis-Tris NuPAGE® gels for 1 hr, 140 V. A Typhoon Fluorescent Imager was used for FITC visualization of the probe bound enzyme. Following fluorescent scanning the gels were colloidal blue stained (calpain-1 and calpain-2) or silver stained (papain, cathepsin B, and cathepsin L) to demonstrate that the same amount of enzyme had been used in all lanes. (See Appendix A - Supporting Information).
Chapter 5

Allosteric inhibitors of calpains: reevaluating inhibition by PD150606 and LSEAL

5.1 Abstract

*Background:* The mercaptoacrylate calpain inhibitor, PD150606, has been shown by X-ray crystallography to bind to a hydrophobic groove in the enzyme's penta-EF-hand domains far away from the catalytic cleft and has been previously described as an uncompetitive inhibitor of calpains. The penta-peptide LSEAL has been reported to be an inhibitor of calpain and was predicted to bind in the same hydrophobic groove. The X-ray crystal structure of calpain-2 bound to its endogenous calpain inhibitor, calpastatin, shows that calpastatin also binds to the hydrophobic grooves in the two penta-EF-hand domains, but its inhibitory domain binds to the protease core domains and blocks the active site cleft directly.

*Methods:* The mechanisms of inhibition by PD150606 and LSEAL were investigated using steady-state kinetics of cleavage of a fluorogenic substrate by calpain-2 and the protease core of calpain1, as well as by examining the inhibition of casein hydrolysis by calpain and the autoproteolysis of calpain.
Results: PD150606 inhibits both full-length calpain-2 and the protease core of calpain-1 with an apparent noncompetitive kinetic model. The penta-peptide LSEAL failed to inhibit either whole calpain or its protease core.

Conclusions: PD150606 cannot inhibit cleavage by calpain-2 of small substrates via binding to the penta-EF-hand domain.

General significance: PD150606 is often described as a calpain-specific inhibitor due to its ability to target the penta-EF-hand domains of calpain, but we show that it must be acting at a site on the protease core domain instead.

5.2 Introduction

Calpains are a family of intracellular cysteine proteases that are cooperatively activated by Ca\(^{2+}\). They are responsible for converting localized Ca\(^{2+}\) signals into a cellular response through specific and limited cleavage of proteolytic targets. The cellular roles of calpains during signal transduction include reorganization of the cytoskeleton, participation in cell cycle regulation, and apoptosis\cite{39,200}. Over-activation of calpains contributes to pathological effects in conditions such as: heart attack, stroke, neurodegeneration, cancer, Alzheimer’s disease, and muscular dystrophy\cite{62,200-202}. Selective inhibitors of calpains could, therefore, have investigative and therapeutic applications in several diseases.
Calpains-1 and -2, the most extensively studied isoforms, are heterodimers of a distinct catalytic 80-kDa subunit and a common regulatory 28-kDa subunit [39-41, 205, 206]. During activation they bind up to ten Ca$^{2+}$, four to each of the two C-terminal penta-EF-hand (PEF) domains [97, 98] and one to each of the two protease core domains [89]. Binding by these ions causes a realignment of the domains to form the active site cleft and a more compact enzyme.

It is the calcium-bound form of the enzyme that is recognized by the endogenous calpain-specific inhibitor, calpastatin. This intrinsically unstructured 70-kDa protein has a domain L of unknown function, followed by four independently active inhibitory domains 1 through 4, each containing subdomains A, B, and C [124]. Recent crystal structures of calcium-bound calpain inhibited by calpastatin have revealed the mode of inhibition [97, 98]. Terminal subdomains A and C of the inhibitor form amphipathic $\alpha$-helices on binding to hydrophobic clefts in the two PEF domains. Subdomain B is responsible for inhibition of calpain and binds across and occludes the protease core active site.

The ongoing development of calpain inhibitors has focused mainly on peptidomimetic compounds containing electrophilic warheads targeted to the active site cleft. However, the cleft is quite similar in cathepsins and other cysteine proteases, making it difficult to achieve calpain specificity. Allosteric inhibitors of calpains have the
potential advantage of binding regions of the enzyme removed from the active site that are unique to calpain. Two types of allosteric calpain inhibitors have been reported: mercaptoacrylates, such as PD150606[171] and other variants[172], and peptidic inhibitors, such as the penta-peptide LSEAL[143].

The crystal structure of the inhibitor PD150606 bound to the calpain small subunit PEF domain implied an allosteric mechanism of inhibition through binding to a hydrophobic pocket distant from the enzyme’s active site[91,123]. Ca$^{2+}$-binding to the calpain small subunit PEF domain[90] results in the opening of a small hydrophobic pocket near EF-hand 1 into which PD150606 was observed to bind. Binding of PD150606 resulted in the movement of a single residue (Gln173)[91,123]. The calpastatin subdomain C binds to the same region, but primarily in the cleft between helices E1 and F1, and results in a more significant shift of the E1 helix[97,123]. The penta-peptide LSEAL was characterized as a calpain inhibitor and was noted to contain a conserved motif found in the PEF-binding regions of calpastatin[143]. NMR studies have shown the LSEAL peptide binds hydrophobic clefts in a manner similar to calpastatin subdomains A and C[161].

The recently solved mechanism of action of the calpastatin inhibitor raises questions about how small molecules binding to pre-formed hydrophobic clefts in the PEF domains confer inhibition of the enzyme[97,98]. Calpastatin uses these clefts solely as anchor points to help direct the inhibitory region to the active site cleft. Here we have
investigated the inhibition of calpain by PD150606 and LSEAL and show that they are unlikely to work by an allosteric mechanism that involves binding to the PEF domains.

5.3 Materials and Methods

5.3.1 Materials

Active and inactive (C105S) rat recombinant calpain-2[88], and calpain-1 protease core[89] were expressed and purified as previously described. Porcine erythrocyte calpain-1 was purchased from Calbiochem. The calpain substrate (EDANS)-EPLFAERK-(DABCYL) and the penta-peptide N-acetyl-LSEAL-amide were synthesized by Biomer Technology. Inhibitors were purchased from Calbiochem (E64, leupeptin, and calpeptin) or Sigma-Aldrich (PD150606), with the exception of SNJ-1945, which was graciously donated by Senju Pharmaceutical Co. Ltd. (Kobe, Japan).

5.3.2 (EDANS)-EPLFAERK-(DABCYL) hydrolysis assay.

The rate of cleavage of the (EDANS)-EPLFAERK-(DABCYL) fluorescent substrate was monitored in a 0.5 mL cuvette using a Perkin Elmer LS55 fluorescence spectrometer with $\lambda_{\text{ex}} = 335$ nm, $\lambda_{\text{em}} = 500$ nm and slit widths of 10 nm. Triplicate readings of reaction profiles were obtained at a sampling interval of 0.1 s at a
temperature of 25 °C. The cuvette contained 100 nM calpain-1 protease core or 10 nM calpain-2, 5 µM substrate, and either 50 µM PD150606 or 100 µM LSEAL peptide (each dissolved in DMSO at stock concentrations of 50 mM) in 10 mM HEPES pH 7.4, 10 mM DTT and sufficient DMSO to bring the total concentration of DMSO to 0.4% for all assays. After incubating the above compounds for 5 min, the reaction was initiated with the addition of CaCl$_2$ to 4 mM in a final volume of 0.5 mL. For enzyme kinetic analysis with PD150606, the hydrolysis assays were performed under the same conditions but with substrate concentrations in the range of 1-100 µM and PD150606 concentrations in the range of 0-50 µM. The initial reaction rates were determined by fitting a straight line to the linear portion of the progress curves (first 10 s after calcium addition). The resulting rate data were examined with Michaelis-Menten plots for the analysis of enzyme inhibition. $K_m$ and $V_{max}$ values for the reactions were obtained from direct fits of the Michaelis-Menten equation to the data using the non-linear least squares fitting function of gnuplot (version 4.6) [224]. All reactions were corrected for the inner filter effect as previously described by measuring the fluorescence of increasing amounts of substrate with a fixed concentration of free EDANS fluor and thereby calculating a correction factor[225].

5.3.3 Calpain autolysis/proteolysis assays.

Autolysis of 0.5 mg/mL (5 µM) calpain-2 was performed at a temperature of 25 °C in 10 mM HEPES pH 7.4, 10 mM DTT and initiated with the addition CaCl$_2$ to 100 mM in a final volume of 400 µL. Autolysis was performed in the absence and presence of 0.1
mM, 1 mM, and 3 mM PD150606. Immediately, and 20 min after calcium addition, aliquots (40 µL) were removed and mixed with 20 µL of SDS sample buffer and 0.5 M EDTA to stop the reaction. Samples were then analyzed by SDS-PAGE in 10% polyacrylamide gels. Proteolysis of 0.5 mg/mL C105S calpain-2 by 0.5 mg/mL (13 µM) calpain-1 protease core was performed and analyzed in the same manner.

5.3.4 Casein hydrolysis assay.

Hydrolysis of 5 mg/mL casein by 0.5 mg/mL (5 µM) calpain-2 and 0.4 mg/mL (10 µM) calpain-1 protease core was performed as previously described[315] for 30 min at 25 °C in the presence of 10 µM leupeptin, calpeptin, SNJ-1945, PD150606, or LSEAL peptide. The reactions were performed with 5 mM CaCl₂ or 5 mM EDTA in 50 mM Tris-HCl pH 8.0 and initiated by the addition of enzyme. Inhibited enzyme activities were calculated relative to control reactions that contained the same final DMSO concentrations (2%) as the inhibitor assays.

5.3.5 Thermal shift in a differential light scattering assay.

Temperature-induced aggregation was measured using differential static light scattering in a 384-well format (StarGazer, Harbinger Biotechnology) as previously described[316]. A solution containing 0.4 mg/mL of either C105S calpain-2 or calpain-1 protease core was prepared in 100 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, and 1 mM CaCl₂. Individual wells contained DMSO as a vehicle control, the compound PD150606, or the penta-peptide LSEAL. Inhibitors were tested at concentrations ranging
from 50 µM to 1 mM. Each measurement was done in triplicate and averaged to give the
temperature of aggregation ($T_{agg}$) and the change in temperature of aggregation
compared to the control values ($ΔT_{agg}$).

5.4 Results

5.4.1 PD150606 inhibits the protease core of calpain-1.

The compound PD150606 was analyzed for inhibitory activity against both full-
length calpain-2 and the calpain-1 protease core domains using the peptide FRET
substrate (EDANS)-EPLFAERK-(DABCYL). The rate of cleavage of the fluorogenic
substrate by calpain-2 in the absence of inhibitor decreased as a function of time (Figure
5.1A) mainly due to calpain autoproteolysis. A similar trajectory was seen in the
presence of 50 µM PD150606, but the activity level was reduced to about 30% of the
uninhibited reaction. Similar inhibition by PD150606 was also observed with full-length
calpain-1 (not shown), as previous studies have observed[171]. The rate of cleavage of
the fluorogenic substrate by calpain-1 protease core, which was linear with time because
the core is resistant to autoproteolysis, was reduced to 50% of the uninhibited reaction
by the presence of 50 µM PD150606 (Figure 5.1B). In comparison to PD150606,
leupeptin at 10 µM showed a greater degree of inhibition of both the full-length calpain-2
and the calpain-1 protease core.
Figure 5.1: Inhibition of calpain-2 and calpain-1 protease core. The lines represent the average of the triplicates of fluorescence-based substrate cleavage assays that were performed with (EDANS)-EPLFAERK-(DABCYL), and either calpain-2 (A) or calpain-1 protease core (B) in the absence of inhibitors (solid line), and in the presence of 50 µM PD150606 (dash-dot line), 100 µM LSEAL peptide (dotted line) or 10 µM leupeptin (dashed line).
5.4.2 LSEAL inhibits neither the protease core of calpain-1 nor full-length calpain-2.

When the LSEAL penta-peptide was tested for inhibition of both full-length calpain-2 and the calpain-1 protease core, no significant difference in the rate of fluorogenic substrate cleavage was observed between untreated and treated reactions (Figure 5.1 A and B).

5.4.3 PD150606 inhibits calpain autoproteolysis by a direct effect on the protease core.

The natural substrates of calpain are proteins rather than small peptides, and the autoproteolysis of the enzyme is an effective way to test this activity. In addition, the autoproteolysis assay is not subject to the fluorescence quenching seen with the FRET substrate and this allowed us to test much higher concentrations of PD150606. As previously observed[89], when calpain-2 is activated by Ca$^{2+}$, the large 80 kDa subunit is autoproteolyzed to produce the 38 kDa protease core and the large subunit PEF domain, which co-migrated with the small subunit PEF domain (see the 20-min time point in Figure 5.2A). The addition of 3 mM PD150606 completely inhibited autoproteolysis (Figure 5.2B). In the presence of 1 mM PD150606 partial autoproteolysis was evident.
Figure 5.2: The effect of PD150606 inhibition on calpain-2 autoproteolysis and calpain-1 protease core cleavage of inactive C105S calpain-2. Calpain-2 autoproteolysis was done in the absence (A) or presence (B) of PD150606. Calpain-1 protease core cleavage of inactive C105S calpain-2 was carried out in the absence (C) or presence (D) of PD150606. Samples were taken initially and after 20 min, as indicated. All reactions were initiated with CaCl$_2$. Reactions were sampled and terminated by the addition of EDTA and 3x SDS sample buffer, and analyzed by SDS-PAGE. The two most intense bands of the molecular weight marker have masses of 66 kDa and 27 kDa, as indicated.
with the appearance of a 60-kDa band that represents incomplete release of the protease core from the neighbouring C2L domain. At 0.1 mM PD150606 the level of autoproteolysis after 20 min was only marginally less than in the complete absence of inhibitor. To test if PD150606 was inhibiting autoproteolysis allosterically by binding to a site on the PEF domains or by direct action on the protease core we used a variation of the autoproteolysis assay where the substrate was the active site-inactivated calpain-2 (C105S mutant) and the enzyme was the calpain-1 protease core. In the control reaction the calpain-1 protease core cleaved the inactive C105S mutant of calpain-2 to produce fragments similar to those produced by calpain-2 autoproteolysis, including the 38 kDa protease core fragment, the 60 kDa product, and a slightly larger form of the large subunit PEF domain (Figure 5.2C). Both 1 mM and 3 mM PD150606 produced complete inhibition of calpain-1 protease core activity as judged by the absence of the 60-kDa and PEF large subunit products, while inhibition by 0.1 mM PD150606 was incomplete (Figure 5.2D).

5.4.4 PD150606 is a relatively weak calpain inhibitor as shown by the fixed-point casein hydrolysis assay.

Using the traditional casein hydrolysis assay for calpain, we compared the effectiveness of PD150606 on both full-length calpain-2 as well as the protease core of calpain-1 in relation to a number of other established calpain inhibitors (Figure 5.3). For full-length calpain-2, the inhibitors leupeptin (10 µM), calpeptin, (10 µM) and SNJ1945
Figure 5.3: The inhibition of calpain-2 and calpain-1 protease core in a fixed-point casein hydrolysis assay. The hydrolysis of casein by calpain-2 (dark grey) or calpain-1 protease core (light grey) was assayed in the absence or presence of inhibitors. Data from triplicate assays were corrected for controls containing EDTA and normalized to the activity of the sample containing DMSO and lacking inhibitor. Error bars represent the standard deviation.
(10 µM) reduced activity to 10% or less of the control values. In the presence of PD150606 (10 µM) activity was reduced to ~65%. When using the protease core of calpain-1 in the fixed-point casein hydrolysis assay, these inhibitors produced similar levels of inhibition to those seen with the full-length enzyme. With both forms of the enzyme PD150606 inhibition is much less effective than leupeptin, calpeptin and SNJ1945. When the LSEAL penta-peptide was tested for inhibition of full-length calpain-2 in the casein hydrolysis assay, no significant inhibition was observed (Figure 5.3).

5.4.5 PD150606 binding stabilizes full-length calpain-2 but not the calpain-1 protease core.

PD150606 and LSEAL were analyzed for binding to full-length C105S calpain-2 and calpain-1 protease core using a thermal denaturation assay. C105S calpain-2 had a $T_{agg}$ of 54.3 ± 0.2 °C (Table 5.1). This increased by over 20 °C ($\Delta T_{agg} = 20.8$ °C) when incubated with 500 µM PD150606. No such stabilization was seen with the calpain-1 protease core ($T_{agg} = 60.8 \pm 0.2$ °C), which in the same 500 µM PD150606 had a $\Delta T_{agg}$ of -0.4 °C. There was no significant stabilization of either form of calpain by the penta-peptide LSEAL even at concentrations as high as 1 mM.
Table 5.1: The effect of PD150606 and LSEAL on calpain-2 and calpain-1 protease core aggregation temperatures

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Compound</th>
<th>$T_{agg}$ (°C) enzyme alone</th>
<th>$T_{agg}$ (°C) enzyme plus compound</th>
<th>$\Delta T_{agg}$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calpain-2</td>
<td>PD150606 (500 μM)</td>
<td>54.3 ± 0.2</td>
<td>75.1 ± 0.3</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td>LSEAL (1 mM)</td>
<td>54.5 ± 0.2</td>
<td>53.5 ± 0.2</td>
<td>-1</td>
</tr>
<tr>
<td>Calpain-1 protease core</td>
<td>PD150606 (500 μM)</td>
<td>60.8 ± 0.3</td>
<td>60.4 ± 0.2</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td>LSEAL (1 mM)</td>
<td>60.4 ± 0.2</td>
<td>60.0 ± 0.2</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td>SNJ-1945 (1 mM)</td>
<td>60.4 ± 0.2</td>
<td>64.1 ± 0.1</td>
<td>-3.7</td>
</tr>
</tbody>
</table>
5.4.6 **PD150606 does not follow traditional Michaelis-Menten kinetics with full-length calpain-2.**

Using the fluorogenic substrate hydrolysis assay, the mode of inhibition of PD150606 was investigated with respect to full-length calpain-2. PD150606 binds to calpain-2 slowly in the presence of calcium but inhibition is stronger when PD150606 is pre-incubated with the calpain enzyme without calcium (Figure 5.4). This suggests that PD150606 does not follow the traditional Michaelis-Menten kinetics in enzyme inhibition, but rather exhibits slow-binding inhibition[317].

Extensive analysis was done to ascertain the mode of inhibition of PD150606 after pre-incubation with calpain, again using the fluorogenic substrate hydrolysis assay. The Michaelis-Menten plots (v vs [S]) are shown in Figure 5.5A. There appears to be a reduction in the $V_{\text{max}}$ suggesting that the mode of inhibition is not competitive. The reaction data were fit using the simultaneous non-linear regression method[318] to determine the most likely kinetic model for the inhibition (Table 5.2). The model with the lowest error estimates was the noncompetitive inhibition mechanism. It was not possible to further analyze the data in terms of the slow-binding inhibition by the method of Morrison[317] due to the fact that autoproteolysis results in no steady-state velocity (see the DMSO control curve in Figure 5.4).
Figure 5.4: Differential binding of PD150606 to apo- and Ca$^{2+}$-bound calpain-2.

Triplicate fluorescence-based substrate cleavage assays were performed with (EDANS)-EPLFAERK-(DABCYL) and calpain-2. DMSO was used as a control lacking inhibitor (solid line). The reactions were initiated by the addition of CaCl$_2$. PD150606 inhibitor was added 5 min before (dashed) or with (dotted) calpain activation by Ca$^{2+}$. The averaged plots are shown.
Figure 5.5: Enzyme kinetics curves for calpain-2 and calpain-1 protease core inhibited by PD150606. Triplicate fluorescence-based substrate cleavage assays were performed with increasing concentrations of (EDANS)-EPLFAERK-(DABCYL) and of PD150606 (0, 25, 40 and 50 μM as indicated on the graph). Progress curves for calpain-2 were analyzed using the initial linear portion to determine a reaction velocity, and this data were fit to the Michaelis-Menten (A) equation. Progress curves for calpain-1 protease core were analyzed and data were fit in a similar manner (B). $K_m$ and $V_{max}$ values for both enzymes were obtained from the fit to the Michaelis-Menten equations.
Table 5.2: Results of simultaneous non-linear regression to the Michaelis-Menten equations for different kinetic models of inhibition of calpain-2 and calpain-1 protease core by PD150606 with standard error estimates. Error estimates are also given as percentages in parentheses.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Inhibition Model</th>
<th>(k_{\text{cat}}) (s(^{-1}))</th>
<th>(K_m) (μM)</th>
<th>(K_{I1}) (μM) (^a)</th>
<th>(K_{I2}) (μM) (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calpain-2</td>
<td>Noncompetitive</td>
<td>2.9 ± 0.1 (5%)</td>
<td>39 ± 3 (7%)</td>
<td>60 ± 5 (8%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncompetitive</td>
<td>4.0 ± 0.4 (10%)</td>
<td>75 ± 9 (12%)</td>
<td></td>
<td>20 ± 3 (15%)</td>
</tr>
<tr>
<td></td>
<td>Competitive</td>
<td>2.3 ± 0.1 (6%)</td>
<td>29 ± 3 (10%)</td>
<td>40 ± 5 (12%)</td>
<td></td>
</tr>
<tr>
<td>Calpain-1 protease core</td>
<td>Mixed</td>
<td>3.1 ± 0.2 (7%)</td>
<td>44 ± 5 (11%)</td>
<td>70 ± 10 (14%)</td>
<td>44 ± 10 (23%)</td>
</tr>
<tr>
<td></td>
<td>Noncompetitive</td>
<td>0.121 ± 0.005 (4%)</td>
<td>46 ± 3 (7%)</td>
<td>330 ± 70 (21%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uncompetitive</td>
<td>0.132 ± 0.007 (5%)</td>
<td>53 ± 4 (8%)</td>
<td></td>
<td>120 ± 30 (25%)</td>
</tr>
<tr>
<td></td>
<td>Competitive</td>
<td>0.114 ± 0.004 (4%)</td>
<td>43 ± 3 (7%)</td>
<td>270 ± 80 (30%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed</td>
<td>0.129 ± 0.007 (5%)</td>
<td>51 ± 4 (8%)</td>
<td>800 ± 800 (100%)</td>
<td>150 ± 50 (33%)</td>
</tr>
</tbody>
</table>

\(^a\) \(K_{I1}\) is the dissociation constant for inhibitor binding to free enzyme
\(^b\) \(K_{I2}\) is the dissociation constant for inhibitor binding to the ES complex.
5.4.7 PD150606 appears to inhibit the protease core of calpain-1 in a noncompetitive manner.

As PD150606 inhibits the protease core of calpain-1 alone, the hydrolysis of the substrate (EDANS)-EPLFAERK-(DABCYL) was again employed for kinetic analyses with the calpain-1 protease core. The initial rate of substrate cleavage by the calpain-1 protease core was illustrated by Michaelis-Menten plots (Figure 5.5B). Again, simultaneous non-linear regression using all of the reaction data suggests a noncompetitive model of inhibition of the calpain-1 protease core by PD150606. Attempts were made to confirm the mode of inhibition by using higher concentrations of inhibitor. However, these reactions were problematic because PD150606 absorbs at the excitation wavelength of the (EDANS)-EPLFAERK-(DABCYL) substrate.

5.5 Discussion

5.5.1 Calpastatin binds to PEF domains only to anchor its inhibitory B peptide

We opened this investigation because the mechanism of inhibition of calpain by calpastatin showed that binding of the A and C peptides to the PEF domains was only responsible for the anchoring and orientation of the inhibitor. The inhibitory region of
calpastatin is the B peptide, which binds to the active site cleft and is far removed from the A and C peptides. It is important to note that peptides corresponding to the A and C subdomains of calpastatin do not inhibit calpain activity in direct proteolysis assays under calcium-saturating conditions (not shown), and truncations of calpastatin are not inhibitory without the B subdomain[134]. On the contrary, the A and C subdomains of calpastatin have been shown to be activators of calpain[319] although this is a result we have not been able to reproduce.

5.5.2 PD150606 does not induce changes in structure or dynamics of calpain

From the various calpain crystal structures, we now know that the hydrophobic clefts on the PEF domains targeted by the A and C peptides become more exposed upon calcium-activation of calpains[93,97,98]. The driving force behind calpain-activation conformational changes comes from calcium binding to the PEF and protease core domains, resulting in an overall compaction of the enzyme and aligning the active site residues. As calpastatin is an intrinsically disordered protein with some mild propensity for α-helical and β-turn formation[133], ordered structure requires binding to calcium-activated calpain and the pre-formed hydrophobic clefts. Those hydrophobic clefts only shift slightly upon binding the calpastatin A or C peptide[97,98,123]. From this perspective, it seems unlikely that a small molecule or peptide with no secondary structure would, by binding to the exposed clefts in the presence of Ca$^{2+}$, induce changes in the calpain conformation leading to inhibition. In the case of PD150606, the
overall structure of the calpain small subunit is unchanged when comparing PD150606-bound to PD150606-free structures except for the movement of a single glutamine sidechain[123].

Recently, the concept of allostery has expanded to include the idea that binding of a ligand at a distant site may cause alterations in the dynamic properties of proteins rather than simply changes in their structure[320,321]. At this point, based on the crystal structures of the small subunit PEF domains, there is no evidence to suggest that binding of PD150606 is able to alter the dynamics or the structure of the PEF domain and hence transmit those changes to the protease core to reduce calpain activity.

5.5.3 Thermal shift assays are neutral with the protease core

Thermal shift assays showed no effect for binding of PD150606 to the protease core of calpain, whereas the signal was strong for its binding to full-length calpain. This is not surprising because most inhibitors of the core do not stabilize its structure, even if covalently bound, probably because of the flexibility of the two core subdomains (PC1 and PC2) about a central pivot region[89]. For example, the alpha-ketoamide inhibitor SNJ-1945 inhibits the calpain protease core well (Figure 5.3) and forms a stable covalent complex in the active site cleft[118] and yet has a negligible stabilizing effect on the protease core in the thermal shift assay (Table 5.1). Perhaps PD150606 is able to increase the $T_{agg}$ of the full-length calpain by binding to the PEF domain but may inhibit it
only by binding to a site on the protease core. Given the lack of stabilization by SNJ-1945, then it would not be surprising that PD150606 also does not increase the $T_{agg}$ of the protease core.

**5.5.4 Mode of inhibition by PD150606 is not clear**

The mode of inhibition of PD150606 remains unclear. For the full-length calpain-2, it appears that PD150606 has some affinity for the enzyme in the absence of Ca$^{2+}$. This is inconsistent with our result of a best fit of the kinetic data to noncompetitive inhibition and to previous reports of uncompetitive inhibition. Both of these kinetic models would require binding in the presence of substrate and hence Ca$^{2+}$. The apparent affinity in the absence of Ca$^{2+}$ is also inconsistent with the structural explanation that PD150606 inhibits by binding to the hydrophobic clefts exposed on the calcium-activated PEF domains. For the calpain-1 protease core, PD150606 also appears to inhibit in a noncompetitive manner. Again this kinetic model implies that PD150606 binds only to the calcium-activated protease core with and without substrate bound to induce an inactivating conformational change. It is difficult to reconcile this kinetic model with the structure of the protease core and the active site cleft. Compared to full-length calpain the protease is core is relatively small and has no clefts other than the active site that are common to the full-length protein and that would provide an obvious binding site for PD150606.
5.5.5 LSEAL does not inhibit calpain and shows no affinity for PEF domains

LSEAL did not display any inhibition of calpain in the in vitro fluorescence-based substrate cleavage assay or the casein hydrolysis assay. This does not rule out the possibility that LSEAL can act as a calpain inhibitor in vivo. For example, the calcium-dependent clefts on calpain could conceivably govern sub-cellular localization[108,139,140] after activation by helping the enzyme to interact with hydrophobic residues on ligands and substrates[141,142,322,323]. This may explain the ability of the inhibitors discussed here to reduce cleavage of some protein substrates[123,143,144]. Studies investigating the penta-peptide LSEAL used tau and alpha-synuclein cleavage to assess inhibition[143]. In this context, LSEAL may interfere with hydrophobic interactions between enzyme and substrate. Calpastatin could possibly function in this manner as well. It may block interactions of various protein substrates with the active enzyme in addition to inhibiting proteolysis at the active site. However, we conclude that binding to hydrophobic clefts in the calpain PEF domains does not directly confer proteolytic inhibition.

5.6 Conclusion

The calpain assays presented here all showed no inhibition by the penta-peptide LSEAL on either full-length calpain-2 or the protease core of calpain-1. The assays also
show that PD150606 inhibits both the protease core of calpain-1 and full-length calpain-2. Thus it can inhibit calpain in the absence of the C2L domain and the PEF domains of the large and small subunits. While PD150606 inhibits PEF domain-containing calpains, it was also previously shown to inhibit cathepsin B and thermolysin albeit to a much lesser extent[171]. In more recent literature, PD150606 has been found to inhibit other proteases such as MMP-2[324], which have no similarity to calpains. It has also been shown to inhibit kainate-induced Ca\(^{2+}\) influx independent of its action on calpain[325]. Our results show that PD150606 has a similar inhibitory effect on both the full-length calpain-2 and the calpain-1 protease core, providing further support for a single mode of inhibition that does not require interaction with the PEF domains. We cannot completely rule out an additional allosteric effect contributing to calpain core inhibition, but it seems unnecessary and unlikely that this small compound would have a second mechanism of inhibition.

5.7 Acknowledgements

We thank Drs. Guillermo Senisterra and Masoud Vedadi from the Structural Genomics Consortium at the University of Toronto for their assistance with the StarGazer DSLS. Senju Pharmaceutical Co. Ltd. (Kobe, Japan) kindly provided the sample of SNJ-1945 inhibitor. We are grateful to Sherry Gauthier for technical assistance and to Dr. Rachel Hanna for helpful discussions. This work was funded by
the Canadian Institutes for Health Research. K.E.L. is the recipient of an R.J. Wilson Fellowship. P.L.D. holds a Canada Research Chair in Protein Engineering.
Chapter 6

General Discussion

This thesis demonstrates the potential for using the endogenous inhibitor of calpain, calpastatin, to develop novel calpain-specific inhibitors. As well it examines the feasibility of allostERIC inhibition.

Chapter 3 presented a set of novel macrocyclic and amidine-based peptides initially designed based on the β-turn of calpastatin in the calpain active site cleft. These were found to inhibit calpain in several different ways. While some work has been done by other groups to develop macrocyclic inhibitors of calpain[169], the compounds studied here were smaller with fewer bulky chemical groups. In addition, the potency of inhibition was within the range found with other peptidomimetic competitive inhibitors of calpain[149], while the specificity of these macrocyclic compounds for calpain was much greater than the series studied by Abell et al. [169]. It is clear that the novel cyclic peptides and peptidomimetic compounds shown here provide a firm basis for the development of improved calpain inhibitors.

Chapter 4 described the use of the calpain-calpastatin structures to design novel α-helical calpain inhibitors. By mimicking the interactions between calpain and the conserved residues of calpastatin subdomain B, helical inhibitors were designed and
cross-linking strategies were devised in order to lower the entropic cost of binding by allowing the helix to form prior to interaction. This resulted in a potent and selective inhibitor, much improved over the unstabilized form of the peptide. Irreversible epoxide warheads were attached to the peptides along with fluorescent-based tags, allowing for use as calpain family activity-based probes. The use of the calpain-calpastatin structure to design the stabilized helix compounds has proven to be a valid method resulting in the development of selective activity-based probes and potent inhibitors, both of which are much sought after in calpain research.

Chapter 5 involved the investigation of two calpain inhibitors previously reported to allosterically inhibit calpain through binding to hydrophobic pockets on the calpain PEF domains remote from the protease core. As a result of this study we refute any suggestion of allosteric inhibition via binding to the PEF domains. Of the two inhibitors, PD150606 was found to inhibit the calpain protease core in the absence of the PEF domains, disproving the previously described inhibition mechanism, while the LSEAL peptide was found to have no effect on either full-length calpain or the protease core construct. Allosteric inhibition of calpains might be possible, but it is unlikely that such inhibition can be achieved simply through binding the hydrophobic PEF domain clefts.
In order to monitor and assess calpain activity accurately, one must be able to determine the initial rate of substrate cleavage by calpain. This is complicated by calpain autoproteolysis but may also be difficult without a method for rapid, quantifiable determination of substrate cleavage products. A number of assays have been developed to investigate calpain activity, however many of these are endpoint assays unable to monitor calpain activation over time, and include the generic substrate casein. Due to the difficulty in obtaining close time-points, cleavage of casein by calpain has generally been used to report activities relative to control calpain conditions. One significant advantage of monitoring cleavage of casein is the reduction of autoproteolysis due to dilution of the enzyme by protein substrate. The tendency for concentrated calpain to autoproteolyse when activated by calcium can, however, be made use of as a method for monitoring calpain activity. Previous work by Chou et al. has shown that autoproteolysis is a rather fast acting process, drawing attention to the limitations of analyzing calpain activity kinetically by autoproteolysis. The amount of fully active enzyme continuously decreases until no activity remains [116]. For both of the above assay types, changes in calcium concentration or incubation time will cause calpain to be variably active and enzyme concentrations will determine the amount of intermolecular proteolysis. Thus, existing assays of enzyme activity are lacking in their ability to accurately quantify calpain activity by collecting enough data for substrate cleavage at early time points.
Previous work in our lab by Dominic Cuerrier investigated calpain’s preference for residues at and near the cleavage site in order to design a better calpain substrate[87]. The overall goals were two-fold: 1) gain knowledge about *in vivo* calpain substrates and residue preferences using small peptides that allowed sequence specificity to be determined independently of secondary structure; 2) aid in the design of peptidomimetic inhibitors and drugs by determining calpain subsite side-chain preferences. The results of this work provided a highly sensitive fluorescent calpain substrate that is cleaved more efficiently than other substrates, and can be used to monitor calpain activity. Previous fluorescent calpain substrates, such as Succ-LLVY-AMC, SLY-MCA, spectrin-based FRET substrates, and (EDANS)-TPLKSPPPSPR-(DABCYL), can be targeted by other cysteine proteases, including cathepsins. It is beneficial to have rapid turnover of substrate by calpain, because a high affinity between substrate and enzyme may help to lessen autoproteolysis early in the reaction time period, while allowing one to monitor rapid cleavage at a variety of substrate concentrations over background noise.

The fluorescent substrate resulting from this research, (EDANS)-EPLFAERK-(DABCYL), has been used to assess the potency of new calpain inhibitors during their development[117,118,326]. The initial enzyme reaction velocity can be directly measured as the increase in fluorescence over time as the substrate is cleaved by calpain. With a sensitive fluorimeter, sufficient data points can be obtained before autoproteolysis becomes problematic. A higher affinity for calpain over other cysteine
proteases makes it more likely that the substrate concentration will exceed $K_m$, a condition needed for Michaelis-Menten kinetic analyses. Cuerrier et al. had calculated the turnover rate for this calpain substrate, but we have improved on the methods presented[87]. First, corrections for the inner filter effect were not properly applied. We found, along with Liu et al. [225], that only a fraction of the fluorescence from the released EDANS group will reach the detector due to the inner filter effect from the presence of the DABCYL quenching group, either when free in solution or as part of the intact substrate. Cuerrier et al. had determined this not to be a problem at the low μM concentrations used. Inner filter effect corrections were applied for higher concentrations, where the fluorescence was linearly extrapolated based on intact substrate prior to cleavage by calpain. We have seen, however, that the inner filter effect is a much more significant problem with cleaved substrate, especially at high concentrations, and fluorescence intensities needed to be corrected appropriately.

Second, they had used only the protease core of calpain-1, whose activity is around ten times lower than that of full-length calpain-2, and whose structure is missing domains adjacent to the active site cleft[40]. Calpain-2 is favoured for *in vitro* activity assays because it can be recombinantly expressed in relatively large quantities. Additionally, assessing inhibition potency of compounds on whole calpain enzyme is physiologically more relevant. Knowing all of the above, I wanted to calculate an accurate affinity between the substrate and calpain using a calculated $K_m$ value, following which I could then calculate accurate $K_i$ values for calpain inhibitors.
IC$_{50}$ values, while useful for comparing similar inhibitors under similar conditions, depend on a specific substrate concentration and affinity. They may be easier to obtain than K$_i$ values, but the latter are representative of the affinity of inhibitor for the enzyme and are not restricted to specific reaction conditions[229,230]. This allows for better comparisons to be made between various compounds for one enzyme. In addition, calculation of just an IC$_{50}$ value does not provide information as to the mode of inhibition whereas the determination of a K$_i$ value goes hand in hand with binding mode determination (competitive, non-competitive, or mixed). In Chapter 3, kinetics analyses were performed accurately for the first time with calpain inhibitors providing clear, precise K$_i$ values for the cyclic compounds without interference from autoproteolysis. Appendix B elaborates on the determination of kinetics parameters with respect to the substrate itself, accounting properly for inner filter effect corrections. Future calpain inhibition studies following these methods will be able to not only calculate trustworthy K$_i$ values for new and existing calpain inhibitors, but also allow for determination of K$_m$ for substrate cleavage sequences and subsequent kinetic analyses. For example, our lab is trying to develop a calpain fluorescent substrate based on the CFP-YFP FRET[327] system using the PLFAER calpain cleavage sequence for applications in vivo (Christian-Scott McCartney et al., unpublished). Analyses of calpain enzyme kinetics parameters performed in this thesis will help in determining the effectiveness of this new CFP-YFP substrate over other in vivo and in vitro substrates currently being used.
6.2 Calpastatin-based peptidomimetic calpain inhibitors

The recent full-length calpain + calpastatin crystal structures have guided the structure-based design of calpain-specific inhibitors. The design of novel calpain inhibitors can be done based on the calpastatin structure and conserved residues and secondary structure in the active site cleft. Also, previous inhibition by peptidomimetic compounds can be rationalized by looking at the active site cleft with the adjacent C2L domain present, as mentioned previously for SNJ-1945[118,122].

6.2.1 Exploiting the conserved α-helix of calpastatin subdomain B

Looking at the structure of calpastatin in the calpain active site cleft, we worked to design novel calpain-specific inhibitors with Dr. Doron Greenbaum’s lab[326]. The peptide B27, which is part of calpastatin, is a highly potent inhibitor with a $K_i$ in the low nM range. However, it can potentially be degraded by peptidases in the cell. The primed side of the calpain active site cleft varies little from similar cysteine proteases, as described previously, with the exception of the region to which the conserved helix of calpastatin subdomain B binds (Figure 1.9, Figure 1.13). By manipulating this sequence, we could target an area of the active site specific to calpains while retaining conserved calpastatin residues responsible for tight interaction. Locking the helix with a cross-link was intended to lower the entropy required for binding, as the helix is pre-formed prior to
binding to calpain, which should ideally promote interaction between protein and peptide. To further increase potency, since the number of residues spanning the active site was drastically reduced compared to B27, we looked to link the peptide to an epoxide warhead for covalent reaction to the active site cysteine and irreversible inhibition. This resulted in a peptide-based inhibitor that displayed specificity for calpain over cathepsins[326].

Calpain research has had need of a reliable tool to monitor the activity of calpain in the cell. We are working to do this with the fluorescent calpain-specific substrate, but it is also possible via activity-based probes. Previous work in our lab has attempted to use inhibitors as activity-based probes of calpain[328], but lack of calpain specificity was deemed a problem to detect activity in cell extracts. The helical peptide inhibitors developed have been shown to bind active calpain with a FITC label present while remaining unreactive toward cathepsins B and L[326]. The development of calpastatin-based helical inhibitors is one successful method of using the calpain-calpastatin structure to our advantage, with clear room for future use of fluorescently tagged peptides as calpain-specific activity-based probes to aid *in vivo* calpain research.
6.2.2 Cyclic peptide mimics of calpastatin secondary structure - complications and binding site determination

A rationalization of calpastatin-based inhibitor design was presented in Chapter 3 of this thesis. Conserved residues in calpastatin are involved in a sharp β-turn required to avoid cleavage by calpain. This turn was an ideal motif to be replicated by cyclic peptides and resulting macrocycles showed fairly potent inhibition without covalent linkage. These classes of cyclic peptidomimetics, in both amidine and macrocycle forms, show potential as protease inhibitors.

The manufacture of these peptide libraries brought with it the complication of two forms of cyclic peptides being made; amidine-containing peptidomimetics and macrocycles were both synthesized from our methods (Hili:2010dc) (Zaretsky, et al., unpublished). This was an unintended result, yet both classes produced potent calpain inhibitors. One wonders, however, as to the binding location and orientation of these compounds. Initial designs intended for synthesized macrocycles to mimic the calpastatin β-turn in the active site cleft. In fact, the first compounds used calpastatin conserved residues and cycles the same size as the β-turn. Clearly from screening and obtaining hits across a variety of sequences and backbone structures, the reality is that design of calpain inhibitors based on this calpastatin structural detail is not that simple. Although some of the cycles did not inhibit in the manner originally designed, displaying non-competitive or mixed inhibition, some did show competitive inhibition. With the docking results presented earlier, it is feasible that these cyclic peptides bind to the
unprimed side of the calpain active site cleft, much in the same position of the calpastatin β-turn. If this is indeed the case, calpain-specific inhibition could be achieved through improving this manner of interacting with the C2L domain.

Docking simulations are of value where we have a structure for the inhibitor. Overwhelmingly, the compound cPGALK tended to be found at the unprimed side nested against the C2L domain. This was facilitated by NMR solution structures of cPGALK used for the docking protocols, and seemed to produce trustworthy docking poses. For other compounds, where we do not have backbone conformations for the macrocycle, docking experiments may not work. The need to specify the docking location will induce bias, as docking protocols will place the ligand in the best position within the indicated area. The grid box was set to encompass the entire active site cleft to accommodate competitively binding compounds. Due to the large size of calpain and the large number of potential binding sites outside of the active site cleft, simple docking protocols may not work in the case of non-competitive or mixed inhibition compounds.

Regardless of the mode of inhibition, more research groups are pointing out that docking results may not be accurate representations of the actual structure of ligand bound to protein. Specifically, Voet et al. made comparisons and propose that in cerebro assessment of binding modes may be on par or better than in silico results in terms of the final actual pose[329]. Binding pose prediction performance is increased when scientific experience, logical selection from structural knowledge, and manual rescoring
approaches were introduced. Many studies involving binding prediction rely on comparing the docking result to the previously solved crystal structures. For Mdmx and/or Mdm2 inhibitors, it was found that attempts to dock a class of compounds failed without previous knowledge of the binding orientation of the imidazoline ring[330]. As novel calpain inhibitors would likely lack protein-inhibitor crystal structures, it is easy to see why docking studies may not be reliable. Improving the reliability of docking protocols may come about as docking is paired with experimental data, such as STD-NMR[331] or a mix of NMR and surface plasmon resonance (SPR) techniques[332], both of which may be possibilities in the future with the calpain enzyme. Photo-crosslinking of inhibitors to calpain could also narrow down the number of docking poses under consideration. Currently, we are investigating STD-NMR as a potential technique to indicate which ligand residues are responsible for interactions with the protein, and we may be able to merge these data with our existing identified poses.

Once docking results are obtained, ranking and comparing poses is a difficult task and many times will not give clear choices. Bissantz et al. highlight that, in their trials, there was no clear relationship between docking scores and rankings, and binding energy prediction was not accurate in any case[333]. Even with the large number of docking programs and algorithms, scoring functions remain an inaccurate method of comparing poses[334], especially when analyzing across different classes of molecules. Docking programs may be able to predict the experimental protein-ligand structure; however scoring functions tend to be less successful at identifying the crystal structure
as the top pose\cite{335}. Typically, the distances comparing computational to experimental structures have 1.5-2 Å RMSD values, which seem shockingly high. If the computational structures are indeed valid, these differences underline the potential for imprecision in the original experimental crystal structures themselves\cite{336}. The problems with automated assessment of docked ligands further highlight the need for merging docking with experimental data in order to assess the quality of poses.

Docking programs are typically designed for small molecule and protein-protein docking, but less so for flexible peptides or macrocycles; a problematic feature when analyzing our cyclic peptides. Using flexible peptides as a ligand presents some difficulties in docking protocols as small molecules tend to bind small, hydrophobic clefts while peptides will generally bind to long, charged surface areas. The Glide software in the Maestro suite has been successful at docking flexible polypeptides provided modifications of the current protocols are given\cite{337}, and our cyclic peptides, especially the cyclol forms with long tails, could be analyzed in this manner in the future. In addition, Forli et al. have attempted to overcome problems associated with docking flexible ring systems in the AutoDock software suite\cite{338} however initial attempts to use this method have not produced more reliable ligand conformations than the conformational searches described in Chapter 3. Nevertheless, future software development and improvements to protocols should be followed with an eye for applications to calpain inhibitors developed in this thesis.
The issues described above become concerns especially for our work since large cyclic peptides will have a constrained cyclized backbone and flexible side chain residues. Therefore, we have found the need for a solved solution structure of the cyclic peptides in order to proceed with docking to reliably predict the ligand binding site. To add to complications, some of our cyclic peptides are non-competitively inhibiting calpain. In these cases, unless the binding pocket or the relative binding location is known, computational docking methods are difficult if not impossible. From the setbacks and complications associated with docking results, the potential unreliability of such experiments underlines the aforementioned need for full-length calpain crystallography, and future determination of ligand binding sites.

Attempts at full-length calpain-2 crystallography have proven difficult in the past, as previously discussed, and inhibitor structures of calpain have mostly depended on covalent complexes with the calpain-1 protease core. The cyclic peptide inhibitors presented in Chapter 3 do not covalently bind the enzyme, and strategizing the attachment of covalent warheads is highly dependent on competitive binding and proximity to the catalytic cysteine. While it would be extremely beneficial to have a structure of full-length calpain bound to cyclic peptide, there are arguably more realistic methods to currently determine peptide binding sites. Attempts are being made to create photo-reactive cyclic peptides through the addition of maleimide-linked benzophenone groups. Following covalent reaction to calpain by UV, tryptic digests and MALDI-TOF peptide fingerprinting is being pursued to map the attachment site onto the known
structure of active calpain. In addition, STD-NMR is being approached as a possible tool for determining key peptide residues responsible for interaction. Proper structure-guided design could then follow each of the above directions to improve potency and specificity of these calpain inhibitor leads. Allosteric binding sites, such as that of the cyclol cPGLdO, would be novel discoveries to the calpain inhibitor field.

6.3 The possibility and significance of calpain allosteric inhibition

It is feasible that allosteric inhibition of calpains could be achieved through interactions with domains neighbouring the protease core, or with regions in the core outside of the active site cleft. It is conceivable that compounds could bind to calpain in the calcium-free structure, preventing the conformational change from calcium-activation required for calpain activity. The binding locations of the non-competitive cyclic peptide inhibitors presented in Chapter 3 remain unknown. Allosteric inhibition of calpains by PD150606 and LSEAL peptide is re-examined in Chapter 4 yet the mode of inhibition by PD150606 also remains unknown. The dismissal of the PEF domain hydrophobic clefts as allosteric sites shows that calpain inhibition is a complex problem, and current compounds should not be used with the expectation that inhibition mechanisms are known. Thus, calpain inhibitor research is challenged with the identification of ligand binding sites outside of the active site cleft.
In allosteric inhibition, ligand-binding induces a conformational change in the enzyme that reduces activity. This is a theoretical possibility with the calpain enzyme because it has two subunits and the several adjacent domains that link together in a circle. It is currently thought that the surrounding domains stabilize the protease core in its active form, enhancing activity over the protease core domains alone. The latter are flexible about a hinge domain. By the same strategy the neighbouring domains rigidify the calcium-free inactive state to prevent casual activation of the enzyme. It is theoretically possible that allosterically inhibiting compounds could destabilize the tight interactions of C2L and PEF domains with the PC1 and PC2 domains, resulting in a destabilized core. Conversely, as discussed in Chapter 5, allosteric inhibition of enzymes can potentially result from upsetting protein dynamics instead of solely changing protein structure[320,321]. The case may be, as suggested by Cooper and Dryden[320], that non-competitively inhibiting compounds will bind calpain at sites that disturb a localized area of protein structure, preventing the natural protein dynamics such that the enzyme activity is lowered or prevented entirely. It is difficult but not impossible to envision a mechanism whereby calpain activity is inhibited by the prevention of side chain or backbone dynamics, or through localized unfolding of key areas. The two calcium binding sites in the protease core domains are removed from the catalytic cleft. Failure to bind Ca$^{2+}$ at either site would block the cooperative movement of the core domains to correctly align the catalytic triad. The perturbation of natural protein dynamics might be detected in future experiments using NMR or molecular dynamics simulations.
Programs used to identify possible surface binding sites produce many potential ligand binding sites on such a large enzyme (not shown). Some of these sites may be responsible for the non-competitive inhibition of calpain, but currently the calpain inhibition field remains in search of an ideal location. It is my hope that the research presented in Chapter 3, perhaps coupled with strategies for full-length calpain crystallography in Appendix D, will eventually lead to identification of a binding site for non-competitive enzyme inhibition.

6.4 Potential “activators” of calpain and inhibition of calpain autoproteolysis

Screening of cyclic peptidomimetics revealed a tendency for some peptides to “activate” calpain. This came as a surprise as the only mechanisms known to overactivate calpain are mutations in the protein sequence[74] and increased sensitivity to calcium[106,339,340]. Some activators of enzymes may bind to allosteric sites near the catalytic domains to induce a conformational change, as is the case with glucokinase activators, or by binding to regulatory subunits of the enzyme, such as with AMP-activated protein kinase activators[341]. Both cases may conceivably be possible for an enzyme such as calpain. However, Appendix C addresses the potential for “activation” of calpain mainly through inhibiting calpain autoproteolysis.
Compounds such as the macrocycle cPGLGF, identified as “activating” during screening of peptide libraries presented in Chapter 3, may inhibit enzyme autoproteolysis without affecting substrate turnover by binding to and stabilizing autoproteolytic “hotspots” previously identified such as loop regions in the C2L domain[116]. If this is indeed the case, such compounds would be useful for in vitro calpain studies, simplifying enzyme activity assays and preventing autoproteolysis, and promoting full-length enzyme stability for crystallographic structure determination. Further work needs to be done to examine the exact mechanism, including a potential docked or crystal structure, in order to further use cPGLGF as an anti-autoproteolytic tool for calpain research.

6.5 Future development of calpain inhibitors

The future of calpain inhibitor development has many avenues to explore. When the structure of the calpastatin-calpain complex was solved, it was postulated that this would drive structure-based design of calpain inhibitors. The work presented here shows that this has indeed happened and is a viable alternative to screening chemical functional groups attached to covalent warheads. Novel peptidomimetics were designed based off the calpastatin mode of inhibition, and have proven to be potent and selective for calpain. The peptidomimetics produced in this work could be further improved upon, but there remain other possibilities for calpain inhibitor design.
Fragment-based drug design is a relatively new field for the development of enzyme inhibitors and our lab approached this direction with the intent of discovering novel calpain-targeted molecules. Current traditional high-throughput screening of chemical compounds involves relatively large lead compounds, and we have previously assayed 11000 of the compounds in the Maybridge screening collection with limited success (unpublished). Complementary screening of smaller fragment libraries, on the other hand, has the potential to provide intermediate leads with options to expand the size and specificity of the molecule. Thus, starting from a fragment-like compound will help to satisfy the conditions for drug-like properties early in development: smaller overall molecular weight, lower lipophilicity, fewer aromatic rings, and fewer hydrogen-bond acceptors[342]. For these reasons, we moved to screen collections of fragments for novel calpain inhibitors.

Screening fragment libraries will sometimes involve the use of high-throughput structure determination, including high-throughput crystallography, as the technique for identification of ligand-binding[343]. The calpain enzyme presents many problems as it is too large for NMR structure determination, but also is problematic for X-ray crystallography for the reasons mentioned previously. Thus, while fragment-based drug design may be a viable alternative for novel calpain-specific inhibitors, other methods for identification of hit compounds are necessary. We have opted to use the StarGazer differential static light scattering (DSLS) assay for observing ligand stabilization of the
enzyme via a thermal shift. This technique helped to identify PD150606 as a small compound able to bind and stabilize the calpain enzyme. The StarGazer has also been used previously to identify ligands that bind proteins via stabilization[316]. Along with isothermal denaturation (ITD) and differential scanning fluorimetry (DSF), DSLS is considered a key method for identifying small molecule compounds or inhibitors for drug development[344]. We have used the StarGazer assay to screen a fragment library of 2000 compounds at the Structural Genomics Consortium in Toronto, resulting in 11 hits. These hits were then validated using the fluorescent calpain activity assay in an adapted moderate throughput capacity on a plate reading fluorimeter, as well as by STD-NMR. Of the eleven hits, three show potential as lead compounds for their ability to inhibit calpain with some specificity over cathepsin L, with IC$_{50}$ values in the range of 100-300 μM. The data presented here are so far unpublished. Further development of hits and leads from fragment compound screening will now require structure determination of the calpain-fragment complex, followed by structure-based design and addition of chemical moieties to improve potency and specificity. There is reason to think these techniques will help to identify novel classes of calpain inhibitors, especially those that may bind allosteric, calpain-specific sites on the enzyme.
References

16 MacLennan, D. H. and Wong, P. T. S. (1971) Isolation of a Calcium-Sequestering Protein from Sarcoplasmic Reticulum. Proceedings of the National ....
18 Nakamura, K., Zuppini, A., Arnaudeau, S., Lynch, J., Ahsan, I., Krause, R., Papp,


188


Evolutionary origin of a calcium-dependent protease by fusion of genes for a thiol protease and a calcium-binding protein? Nature 312, 566–570.


Preferences of Calpastatin, the Intrinsically Unstructured Protein Inhibitor of Calpain. Biochemistry, American Chemical Society 47, 6936–6945.


Comparative specificity and kinetic studies on porcine calpain I and calpain II with naturally occurring peptides and synthetic fluorogenic substrates. Journal of Biological Chemistry 259, 12489–12494.


Crocker, S. J., Smith, P. D., Jackson-Lewis, V., Lamba, W. R., Hayley, S. P.,


(2013) Prime, Schrödinger, LLC.


Williams, T. and Kelley, C. Gnuplot: an interactive plotting program.


Cheng, Y.-C. and Prusoff, W. H. (1973) Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochemical pharmacology 22, 3099–3108.


Armstrong, D. and Zidovetzki, R. Helical Wheel Projections, Riverside, CA.


Adam, G. C., Cravatt, B. F. and Sorensen, E. J. (2001) Profiling the specific
reactivity of the proteome with non-directed activity-based probes. Chemistry & Biology 8, 81–95.


upon Trans-to-Cis Photoisomerization. Chemistry & Biology 9, 391–397.


The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
Appendix A

Supplementary Information for Chapter 4: Development of α-helical calpain probes by mimicking a natural protein-protein interaction

General Information

Amino acids were purchased from Advanced ChemTech (Louisville, KY) or Chem-Impex (Wood Dale, IL). Biotinylated lysine was purchased from Anaspec (Freemont, CA). All crosslinkers and the enzyme papain were purchased from Sigma-Aldrich® (St. Louis, MO). Chemicals were purchased from Fisher Scientific (Pittsburgh, PA). Calpain-1 and cathepsin B were purchased from BioVision (Milpitas, CA). Cathepsin L was purchased from EMD Millipore (Billerica, MA). Substrates were purchased from Peptides International (Louisville, KY). Single fritted reservoirs for peptide synthesis were purchase from Biotage (Redwood City, CA). Film for imaging blots was purchased from Kodak (Rochester, NY). Bis-Tris NuPAGE® gels and a Novex® Colloidal Blue Staining Kit were purchased from Life Technologies (Grand Island, NY). A Vectastain® Elite® ABC kit for biotin blotting was purchased from Vector Laboratories (Burlingame, CA). A Bio-Rad Silver Stain Plus Kit was purchased from Bio-Rad (Hercules, CA). Peptides were synthesized on an Argonaut Quest™ 210 (Argonaut Technologies, Inc. now owned by Biotage, Redwood City, CA) or on a Symphony automated peptide synthesizer (Protein Technologies, Inc., Tuscon, AZ). Peptides were purified on an Agilent 1100 Series LC/MS or an Agilent 1200 Series LC/MS (Agilent...
Technologies, Inc., Santa Clara, CA) Hewlett Packard ChemStation software using a Vydac® C8 column (Grace, Deerfield) or a Zorbax XDB-C18 column (Agilent Technologies, Inc., Santa Clara, CA). Matrix assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectra were obtained using a Bruker Ultraflex III mass spectrometer (Billerica, MA). Electrospray ionization (ESI) mass spectra were obtained with a QTRAP® 3200 (AB SCIEX, Framingham, MA). Circular dichroism (CD) spectra were obtained with a JASCO J-810 spectropolarimeter (JASCO, Inc., Easton, MD) equipped with a Peltier temperature control unit. NMR spectra were obtained using a Bruker Avance III 500 MHz spectrometer equipped with a cryogenic probe (Billerica, MA). UV-Vis absorbance spectra were obtained using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE). Fluorescence spectra were collected with a Berthold Tri-Star multimode microplate reader (Berthold Technologies, GmbH & Co. KG, Bad Wildbad, Germany). Gels were visualized with a Typhoon Fluorescent Imager (GE Healthcare Biosciences, Pittsburgh, PA). Graphing was performed using GraphPad Prism (GraphPad Software, La Jolla, CA).
Synthetic Procedure

General peptide synthesis: All peptides were synthesized at 0.1 or 0.2 mmol scales using Chemmatrix Rinkamide Resin (substitution: 0.52 mmol/g) or CLEAR™ Amide resin (substitution: 0.46 mmol/g). Fmoc-protected amino acids (5-fold excess) were activated with 0.95 equivalents (relative to the amino acid) of HBTU in the presence of 10 equivalents of diisopropylethanolamine (DIPEA). Amino acids were coupled for 5 min at 65 °C in DMF (Quest synthesis) or 25 min at room temperature in DMF (Symphony synthesis). Fmoc deprotection was performed using 20% 4-methyl piperidine in DMF for 5 min at 65 °C (Quest Synthesis) or 2.5 min (x 2) at room temperature (Symphony Synthesis). Side chain deprotection and the simultaneous cleavage from the resin were carried out using a mixture of TFA/thioanisole/ethanedithiol/anisole (90:5:3:2, v/v) at room temperature, for 2.5 hours. The crude peptide was precipitated using cold diethyl ether and purified via reverse-phase chromatography with a C8 preparative column using buffer A (0.1% TFA in Millipore water) and buffer B’ (0.1% TFA in 60% isopropanol/30% acetonitrile /10% Millipore water). Initial HPLC conditions were 5% B’/95% A. Initial conditions were run for 5 min, followed by an increase of solution B’ to 100% at 25 min (5%/min.) at a flow rate of 5 mL/min unless otherwise indicated. The mass of all peptides was verified by MALDI-TOF or ESI-MS and purity (greater than 95%) was checked by analytical HPLC.
Table A.1: Calculated and observed model peptide masses.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Calcd m/z</th>
<th>Obsvd m/z</th>
<th>HPLC-Gradient</th>
<th>Retention Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Peptide Ac-YGGEAAREACARECAARE-NH₂[345]</td>
<td>1954.4</td>
<td>1954.1</td>
<td>0-100%</td>
<td>13.3</td>
</tr>
</tbody>
</table>

(C4 Vydac column over a gradient 0% to 100% of acetonitrile in water (0.1% TFA) over 40 minutes)
Crosslinker Screening Procedure

Preparation of stock solution: A peptide solution (0.114 mM) in NH₄HCO₃ buffer (12 mL, 50 mM, pH=8.0) was treated with TCEP (1M solution in the same NH₄HCO₃ buffer, 1.1 eq.) at room temperature for 1 h[279,280,346]. The concentration of the peptide solution was measured by $A_{280} = 1280 \text{ M}^{-1}\text{cm}^{-1}$ or by a weight based method (molecular weight was adjusted by adding one TFA salt per basic residue and by adding 10% (of the calculated molecular weight) to account for hydration after lyophilization).

Crosslinking procedure in 96-well plate: 90 mL of the stock solution was added to each well of a black round-bottomed 96-well plate (polypropylene). 10 mL of the freshly prepared alkylating agent solution (1.5 mM in anhydrous DMF, 1.5 eq.) was added to each well at room temperature and stirred for 2 h while protected from light. MALDI analysis was done to see the reaction progress and more alkylating agent was added if needed. Addition of 5% HCl to each well neutralized and subsequently quenched the reaction. 100 mL of diethyl ether was added to the organic layer to remove excess alkylating agent. The ether layer was removed by pipetting and MALDI-TOF spectra were taken of the sample in the remaining aqueous solution mixture.
High Concentration of Reactants (for “Selection of the Fittest” Reaction)

Preparation of stock solution: A peptide solution (1.2 mM) in Tris-HCl buffer (100 mM, pH = 8.0) was treated with TCEP (1M solution in the Tris-HCl buffer, 1.1 eq.) at room temperature for 1 h. The concentration of the peptide solution was measured by $A_{280} = 1280 \text{ M}^{-1}\text{cm}^{-1}$ or by a weight based method (molecular weight was adjusted by adding one TFA salt per basic residue and by adding 10% (of the calculated molecular weight) to account for hydration after lyophilization).

Crosslinking procedure in 1.5 mL centrifuge tube: This procedure is slightly modified from the procedure in Materials and Methods. 450 µL of the stock solution of peptide was added to a 1.5 mL polypropylene microcentrifuge tube. 50 mL of the freshly prepared alkylating agent solution (225 mM in anhydrous DMSO, app. 20 eq.) was added to each well at room temperature and the turbid mixture was shaken for 2 h under protection from light. The reaction was quenched by neutralization through the addition of 0.6 N HCl (10 mL) into each well. Each tube was centrifuged to remove the insoluble material and the supernatant was either purified by HPLC analysis or lyophilized.
* 1919.84 appeared to be elimination of thiol from the cysteine.

215
Figure A.1: MALDI spectra of low concentration linker with the model peptide.
Without Crosslinker

Reaction with crosslinker c14
Reaction with crosslinker c15

Reaction with crosslinker c16
Reaction with crosslinker c17

Reaction with crosslinker c18
Reaction with crosslinker c19

![Graph and molecule 19]

Reaction with crosslinker c20

![Graph and molecule c20]
Reaction with crosslinker c21

Reaction with crosslinker c22
Reaction with crosslinker c23

![HPLC profile of c23](image)

Reaction with crosslinker c24

![HPLC profile of c24](image)

Figure A.2: Crude HPLC profile (Low Concentration Reaction with the model peptide).
* denotes the major monocyclization product

Figure A.3: Crude HPLC profile – “selection of the fittest” reaction.
The peptide sample was prepared using peptide concentrations of 2 mM in 0.6 mL of a 9:1 v/v water/D$_2$O mixture in 50mM sodium phosphate, pH 5.5. All spectra were recorded at 10 °C on a Bruker Avance III 500 MHz spectrometer equipped with a cryogenic probe.

All 2D homonuclear spectra were recorded with standard pulse sequences[310]. 2D NOESY experiments were carried out with mixing times of 150 ms and 250 ms, respectively, 5483 Hz on both $t_1$ and $t_2$ dimensions with $t_{1,max} = 93$ ms and $t_{2,max} = 183$ ms, 32 scans. 2D TOCSY experiments were carried out with a mixing time of 75 ms, 5000 Hz on both $t_1$ and $t_2$ dimensions with $t_{1,max} = 93$ ms and $t_{2,max} = 205$ ms. 2D DQF-COSY experiments were carried out with 5000 Hz on both $t_1$ and $t_2$ dimensions with $t_{1,max} = 120$ ms and $t_{2,max} = 205$ ms. The $^1$H carrier frequency was always set to the water peak and chemical shifts were referenced with respect to the residual water peak at 4.90 ppm.

Spectra were processed and analyzed using the programs nmrPipe[311] and XEASY[312], respectively. Time domain data were multiplied by sine square bell window functions shifted by 60° and zero-filled once.

Using DQF-COSY, TOCSY, and NOESY sequence specific assignments were obtained following standard procedures[310].
Figure A.4: Characterization of helix formation in the peptide by NMR spectroscopy. (A) Secondary chemical shifts of α-H as a function of residue. (B) Chemical shift index (CSI) output as a function of residue. Both strongly demonstrate helix formation even in the fraying terminus.
α-H chemical shifts have a strong relationship to protein secondary structures[345,347]. Secondary chemical shifts are calculated by subtracting the experimental values from the intrinsic values. Secondary chemical shifts indicate that all 18 residues show helix formation in the peptide. Output from Chemical Shift Index (CSI) [347,348] clearly shows that 83% (15 out of 18) residues form helical structures. Two of the three non-helical residues are terminal residues. The third non-helical residue is the helix breaker glycine.
**Alanine Scanning Mutagenesis**

To identify the hot spot residues that are important to the protein-peptide interaction, the molecular modeling software package Rosetta[349] was used to calculate the binding free energy changes upon alanine mutation of each residue. The hot spots are defined as the residues that have changes in the binding free energy more than 1 kcal/mol when mutating to alanine (ΔΔG value) [350]. Modeling was begun with the 3BOW.pdb crystal structure. The target sequence was threaded into the backbone, namely E622K(204), H625E(207) and D628A(210), and then the whole peptide was repacked as the starting structure. Alanine scanning was performed for each residue sequentially and the ΔΔG value was calculated as show in Table A.2.
Table A.2: Alanine Scanning Mutagenesis.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Chain ID</th>
<th>Mutation</th>
<th>ΔΔG (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>207</td>
<td>C</td>
<td>E→A</td>
<td>-0.1</td>
</tr>
<tr>
<td>206</td>
<td>C</td>
<td>R→A</td>
<td>0.0</td>
</tr>
<tr>
<td>210</td>
<td>C</td>
<td>A→A</td>
<td>0.0</td>
</tr>
<tr>
<td>211</td>
<td>C</td>
<td>N→A</td>
<td>0.1</td>
</tr>
<tr>
<td>201</td>
<td>C</td>
<td>I→A</td>
<td>0.3</td>
</tr>
<tr>
<td>204</td>
<td>C</td>
<td>K→A</td>
<td>0.4</td>
</tr>
<tr>
<td>203</td>
<td>C</td>
<td>P→A</td>
<td>0.5</td>
</tr>
<tr>
<td>202</td>
<td>C</td>
<td>P→A</td>
<td>1.1</td>
</tr>
<tr>
<td>205</td>
<td>C</td>
<td>Y→A</td>
<td>1.5</td>
</tr>
<tr>
<td>209</td>
<td>C</td>
<td>L→A</td>
<td>1.7</td>
</tr>
<tr>
<td>208</td>
<td>C</td>
<td>L→A</td>
<td>1.9</td>
</tr>
</tbody>
</table>
Synthesis of Calpain Inhibitors

Calpain inhibitors were synthesized via solid phase peptide synthesis in the same manner as the model peptide. Helical calpain inhibitors were stabilized using the α,α'-dibromo-m-xylene crosslinker c15 and the “low concentration” linking protocol.
Table A.3: Uncrosslinked and crosslinked calpastatin fragment peptides.

<table>
<thead>
<tr>
<th>Peptide No.</th>
<th>Sequence</th>
<th>Calcd (m/z)</th>
<th>Obsvd (m/z)</th>
<th>HPLC-Gradient</th>
<th>Retention Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
<td>Ac-IPPKYRELLA-NH₂</td>
<td>1240.5</td>
<td>1240.5</td>
<td>0-100%</td>
<td>6.91</td>
</tr>
<tr>
<td>3a</td>
<td>Ac-IPCKYRCLLA-NH₂</td>
<td>1220.6</td>
<td>1220.5</td>
<td>0-100%</td>
<td>6.85</td>
</tr>
<tr>
<td>3b</td>
<td>Ac-IPPCYRECLA-NH₂</td>
<td>1205.5</td>
<td>1205.3</td>
<td>0-100%</td>
<td>7.60</td>
</tr>
<tr>
<td>3c</td>
<td>Ac-IPPKYCELLC-NH₂</td>
<td>1219.6</td>
<td>1219.4</td>
<td>0-100%</td>
<td>8.30</td>
</tr>
<tr>
<td>3a</td>
<td>Ac-IPCKYRCLLA-NH₂</td>
<td>1322.6</td>
<td>1322.6</td>
<td>0-100%</td>
<td>7.99</td>
</tr>
<tr>
<td>3b</td>
<td>Ac-IPPCYRECLA-NH₂</td>
<td>1307.5</td>
<td>1307.5</td>
<td>0-100%</td>
<td>8.60</td>
</tr>
<tr>
<td>3c</td>
<td>Ac-IPPKYCELLC-NH₂</td>
<td>1321.6</td>
<td>1321.5</td>
<td>0-100%</td>
<td>9.83</td>
</tr>
</tbody>
</table>

3a-c are all m-xylyl crosslinked. Peptides were run on an Agilent LC-MS with an Eclipse XDB-C18 column over a gradient of acetonitrile in water (0.1% HCOOH) over 20 minutes.
CD Spectra of Calpain Inhibitors without TFE

Peptide solutions were prepared at ~100 µM in 50 mM Tris-HCl (pH 7.5) without TFE. Concentrations were determined by measuring tyrosine absorbance at 276 nm with an extinction coefficient of 1400 M\(^{-1}\) cm\(^{-1}\). Scans were conducted from 260 nm to 200 nm[287,291,351]. Measurements were conducted at 20°C in 1 nm step mode with a response time of 4 seconds in a 1 mm path length quartz cuvette.
Figure A.5: CD spectrum of uncrosslinked peptide 3a-c in Tris buffer (50 mM, pH = 7.5).
Figure A.6: CD spectrum of crosslinked peptide 3a-c in Tris buffer (50 mM, pH = 7.5).
Protease Activity Assays

All peptides were evaluated for ability to bind and subsequently inhibit the cysteine protease calpain using standard proteolytic fluorescence activity assays. Inhibition was assayed using a standard donor-quencher strategy using a previously published calpain peptide substrate[138,158,313].

Calpain assay: Enzyme concentration for calpain-1 was 25 nM. The buffer contained 10 mM dithiothreitol (DTT), 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), and 0.015% Brij-35. Substrate concentration was 0.25 µM H-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH[138,158,313]. Varying concentrations of inhibitor, 0, 0.5, 1, 2, 5, 7, 10, 15, 25, 50, and 100 µM, were used for each assay. Positive controls contained no inhibitor and negative controls contained no calcium. Enzyme, buffer, substrate, and inhibitor (or DMSO in controls) were combined. Calpain was activated by the injection of CaCl₂ to a final concentration of 5 mM. All assays were done in triplicate at a total well volume of 100 µL in 96-well plate, and each well contained a separate inhibitor concentration. Fluorescence readings were taken every 13 seconds for one hour in a fluorescent plate reader. Excitation wavelength was 380 nm and the emission wavelength was 500 nm.

Papain assay: Enzyme concentration for papain was 25 nM. The buffer contained 10 mM dithiothreitol (DTT), 100 mM KCl, 2 mM DGTA, 50 mM Tris-HCl (pH 7.5), 0.015% Brij-35. Substrate concentration was 10 µM H-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH[138,158,313]. Varying concentrations of inhibitor, 0, 0.5, 1, 2, 5, 7, 10, 15, 25, 50, and 100 µM, were used for each assay. Buffer, papain, and inhibitor were all
combined first. The assay was initiated by the addition of substrate via a multichannel pipette. All assays were done at a total well volume of 100 µL in 96-well plate, and each well contained a separate inhibitor concentration. Fluorescence readings were taken every 13 seconds for one hour in fluorescent plate reader. The excitation wavelength was 380 nm and the emission wavelength was 500 nm.

**Cathepsin assay:** Enzyme concentration for cathepsin B and cathepsin L was 3 nM. The buffer contained 10 mM DTT, 500 mM sodium acetate (pH 5.5), and 4 mM EGTA, and 0.015% Brij-35[138,158,313]. Substrate concentration for both enzymes was 0.25 µM Z-FR-Amc. 0, 0.5, 1, 2, 5, 7, 10, 15, 25, 50, and 100 µM, were used for each assay. Buffer, enzyme, and inhibitor were combined. Cathepsin assays were activated by the addition of substrate via a multichannel pipette. All assays were done at a total well volume of 100 µL in 96-well plate, and each well contained a separate inhibitor concentration. Fluorescence readings were taken every 13 seconds for one hour in fluorescent plate reader. The excitation wavelength was 351 nm and emission wavelength was 430 nm.
Figure A.7: Representative calpain activity assay progress curve. Progress curves were truncated at ~500 seconds. After 500 seconds the progress curve loses linearity due autoproteolysis of the enzyme[352]. Data collected after curvature began was not used in any calculations.
**K$_m$ Determination for Calpain Substrate**

We calculated the $K_m$ for the NH$_2$-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH substrate when cleaved by calpain using standard Michaelis-Menten kinetics[314]. We identified the initial velocity of calpain at substrate concentrations, 1, 3, 5, 10, 20, and 30 µM. Velocities were determined in RFU/sec then converted to µM/sec using the conversion factor 1386 RFU/µM. The conversion factor was obtained by the total hydrolysis of the substrate NH$_2$-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH in a known concentration by papain. We then plotted velocity vs. substrate concentration and used GraphPad Prism program Michaelis-Menten (under kinetics) to determine the $K_m$. At high concentrations, >10 µM, of substrate the inner filter effect, whereby free quencher absorbed the fluorescence emission of the cleaved fluorophore, became evident. To take this quenching into consideration, the velocity at each substrate concentration was multiplied by the corresponding correction factor: 

$$\text{Corr}\% = \frac{f_{\text{EDANS}} \text{ (at each substrate concentration)}}{f_{\text{EDANS}} \text{ (in the absence of substrate)}}$$ [225].
Figure A.8: Michaelis-Menten curve for determining $K_m$ of calpain substrate.
Figure A.9: IC\textsubscript{50} Curves for enzyme assays. K\textsubscript{i} was calculated from IC\textsubscript{50} using the equation:

\[ K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}} \]

The inhibitor was tested against purified human calpain-1. The K\textsubscript{m} used for the calpain K\textsubscript{i} determination was 4.66 µM.
Kinetic Analysis of Calpain-1 by 3c

To identify inhibition type we used standard Michaelis-Menten treatment. Initial velocities were calculated from the linear segment of the progress curves then plotted against their substrate concentration[314]. Due to the linearity of the first segment of the progress curve we believe that autoproteolysis during the first 500 seconds was not substantial enough to prevent the use of simple Michaelis-Menten kinetics, i.e. loss of enzyme did not change the velocity enough to cause it to deviate from linearity and incorporation of this additional complex would severely complicate the kinetics. Velocities were determined in RFU/sec then converted to µM/sec using the conversion factor 1386 RFU/µM. The conversion factor was obtained by the total hydrolysis of the substrate NH₂-Glu(EDANS)-Pro-Leu-Phe-Ala-Glu-Arg-Lys(DABCYL)-OH in a known concentration by papain.
Figure A.10: Michaelis-Menten plot of initial velocities at different substrate and inhibitor concentrations.
Table A.4: $V_{\text{max}}^{\text{app}}$ and $K_{m}^{\text{app}}$ values obtained from the above Michaelis-Menten plot.

<table>
<thead>
<tr>
<th>[3c] (µM)</th>
<th>$V_{\text{max}}^{\text{app}}$ (µM/sec)</th>
<th>$K_{m}^{\text{app}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0029 ± 0.0003</td>
<td>4.66 ± 1.08</td>
</tr>
<tr>
<td>1</td>
<td>0.0029 ± 0.0002</td>
<td>6.47 ± 0.95</td>
</tr>
<tr>
<td>3</td>
<td>0.0029 ± 0.0003</td>
<td>7.40 ± 1.63</td>
</tr>
<tr>
<td>5</td>
<td>0.0028 ± 0.0003</td>
<td>8.11 ± 1.96</td>
</tr>
<tr>
<td>10</td>
<td>0.0030 ± 0.0004</td>
<td>11.35 ± 3.76</td>
</tr>
</tbody>
</table>
$V_{\text{max}}^{\text{app}}$ is the same at all inhibitor concentrations while $K_{\text{m}}^{\text{app}}$ increases with increasing inhibitor concentration. These results are indicative of competitive inhibition. To avoid weighting errors we used the values of $K_{\text{m}}^{\text{app}}$ and $V_{\text{max}}^{\text{app}}$ determined directly from the non-linear least-squares best fits of the untransformed data and put these values into the reciprocal equation:

$$\frac{1}{v} = \left( \frac{K_{\text{m}}}{V_{\text{max}}} \times \frac{1}{[S]} \right) + \frac{1}{V_{\text{max}}} \quad [314].$$

We then plotted the resulting reciprocal velocities against the respective reciprocal substrate concentrations.
**Figure A.11:** IC$_{50}$ Curves for enzyme assays using inhibitor 3c. $K_i$ was calculated from IC$_{50}$ using the equation:

$$K_i = \frac{IC_{50}}{\frac{[S]}{K_m}}$$

from IC$_{50}$ using the equation: [229,314,353].
Calpain Activity-Based Probe Synthesis

All probes were synthesized by standard solid phase synthesis techniques using single fritted reservoir on Rink Amide resin (0.59 mmol/g substitution). Activation of Fmoc-amino acids (5-fold excess) was achieved with O-(1H-6-chlorobenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorophosphate (HCTU) in the presence of DIPEA. The reaction solvent contains 100% N,N-dimethylforamide (DMF) (HPLC grade, Fisher). The epoxysuccinic acid was synthesized according to a procedure reported in the literature[354]. The epoxysuccinic acid was added using the same coupling procedure as the amino acids.

Biotin tag addition was done using biotinylated lysine in the peptide synthesis process. Fluorescein isothiocyanate (FITC) tag was added post cleavage to the crude peptides. FITC addition was performed post-cleavage by adding 1 eq. fluorescein isothiocyanate and 10 eq. DIPEA to the peptide dissolved in DMF and stirred under argon for 1 hr. For FITC labeled probe an allyloxycarbonyl (alloc) protected lysine was used for the non-tagged lysine. The alloc was removed using 1 eq. tetrakis(triphenyl phosphine)palladium(0) with 10 eq. 5,5-dimethyl-1,3-cyclohexane dione stirred for 2 hr.

Deprotection of side chains and cleavage of peptides from amide resin: Side chain deprotection and simultaneous cleavage from the resin were carried out using a mixture of trifluoroacetic acid (TFA)/triisopropylsilane (TIPS)/water (90:5:5, v/v) at room temperature, for 3 hours. Crude peptide was obtained by ether precipitation and purified by reverse-phase chromatography. The mass of all peptides was verified by ESI-MS.
### Table A.5: Calculated and Observed Masses for Calpain Activity Based Probes.

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
<th>Calcd (m/z)</th>
<th>Obsvd (m/z)</th>
<th>HPLC-Gradient</th>
<th>Retention Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM-01-biotin</td>
<td>Epoxide-bbiotin AAIPPKYCELLCK-biotin</td>
<td>1890.29</td>
<td>1890.1</td>
<td>0-100%</td>
<td>9.864</td>
</tr>
<tr>
<td>NM-02-biotin</td>
<td>Epoxide-bAAIPPKYCELLCK-biotin</td>
<td>1890.29</td>
<td>1890.1</td>
<td>0-100%</td>
<td>10.085</td>
</tr>
<tr>
<td>NM-03-biotin</td>
<td>Epoxide-AbPIPPKYCELLCK-biotin</td>
<td>1928.92</td>
<td>1928.2</td>
<td>0-100%</td>
<td>10.137</td>
</tr>
<tr>
<td>NM-02-FITC</td>
<td>Epoxide-bAAIPPKYCELLCK-FITC</td>
<td>2033.84</td>
<td>1016.5 (M/2)</td>
<td>0-100%</td>
<td>8.655</td>
</tr>
</tbody>
</table>
Enzyme Labeling Experiments

For enzyme labeling experiments a greater concentration of enzyme was used than for the kinetic experiments for visualization of the enzyme of gel stains.

Calpain probe linker experiments: Experimental conditions included 10 mM dithiothreitol (DTT), 1.5 µg Calpain, 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), 0.015% Brij-35, and either 1 µM or 10 µM of biotinylated probe (DCG-04 (positive control), NM-01, NM-02, NM-03) [138,158,313]. Calpain was activated by the addition of calcium (3.33 µM of 50 mM CaCl₂) to a final concentration of 8.3 mM. The negative control contained water instead of CaCl₂, in a calpain solution with 10 µM probe. Probes were allowed to bind to the calpain for 20 minutes at room temperature. The reaction was stopped by the addition of 10 µL NuPage® LDS Running Buffer. 10 µL of each labeling experiment was loaded on a 10% Bis-Tris NuPAGE® gel for 1.5 hrs, 140 V. The bands were then transferred to a PVDF membrane at 30 V for 70 min. The membrane was blocked with casein and blotted for biotin. Film was exposed for 1 hr. and developed.

Calpain labeling experiments: Experimental conditions were 10 µM dithioreitol (DTT), 1.5 µg Calpain-1 or 6 µM Calpain-2 (the calpain-2 was not as active), 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), 0.015% Brij-35, and 1, 2.5, 5, 10, or 20 µM of FITC-labeled NM-02, or 2 µM DCG-04[138,158,313]. Calpain was activated by the addition of calcium (3.33 µM of 50 mM CaCl₂) to a final concentration of 8.3 mM. The negative control was unactivated calpain tube containing 10 µM probe where water was added instead of calcium. Probes were allowed to bind to the calpain for 20 minutes at
room temperature. The reaction was stopped by the addition of 10 µL NuPage® LDS Running Buffer. 10 µL of each condition was loaded on a 10% Bis-Tris NuPAGE® gel for 1.5 hrs at 140 V. The gel was visualized on a Typhoon Fluorescent Imager. Following scanning the gel was stained with a Novex® Colloidal Blue Staining Kit.

**Papain labeling experiments:** Experimental conditions were 10 µM dithioreitol (DTT), 1.5 µg Papain, 100 mM KCl, 2 mM EGTA, 50 mM Tris-HCl (pH 7.5), 0.015% Brij-35, and 1, 2.5, 5, or 10 µM of FITC-labeled NM-02, or 2 mM DCG-04[138,158,313]. Probes were allowed to bind to the papain for 20 minutes at room temperature. The reaction was stopped by the addition of 10 µL NuPage® LDS Running Buffer. 10 µL of each condition was loaded on a 10% Bis-Tris NuPAGE® gel for 1.5 hrs at 140 V. The gel was evaluated on a Typhoon Fluorescent Imager. Following scanning the gel was stained to verify loading with a Bio-Rad Silver Stain Plus Kit.

**Cathepsin labeling experiments:** Experimental conditions were 10 µM dithioreitol (DTT), 1.5 µg Cathepsin B or Cathepsin L, 500 mM sodium acetate, 4 mM EGTA, (pH 5.5), and 1, 2.5, 5, or 10 µM of FITC-labeled NM-02, or 2 mM DCG-04[138,158,313]. Probes were allowed to bind to the enzymes for 20 minutes at room temperature. The reaction was stopped by the addition of 10 µL NuPage® LDS Running Buffer. 10 µL of each condition was loaded on a 12% Bis-Tris NuPAGE® gel for 1.5 hrs., 140 V. The gel was evaluated on a Typhoon Fluorescent Imager. Following scanning the gel was stained to verify loading with a Bio-Rad Silver Stain Plus Kit.
Appendix B

Kinetic analyses of the calpain-specific substrate

(EDANS)-EPLFAERK-(DABCYL)

In order to accurately determine the mode of inhibition via Michaelis-Menten kinetics of the cyclic peptides presented in Chapter 3, it was necessary to address three main concerns in calpain enzyme kinetics: 1) investigate and correct for the inner filter effect (IFE) at substrate concentrations used in the FRET substrate assays; 2) determine a conversion factor to move from relative fluorescent units (RFUs) read by the fluorimeter to μM of product for the Michaelis-Menten equation; 3) correct Michaelis-Menten curves for both the inner filter effect and autoproteolysis. Once all three points were satisfied, I could then assess the calpain-specific substrate (EDANS)-EPLFAERK-(DABCYL) in terms of both enzyme turnover rate as well as affinity for calpain. This information was relevant in trying to determine the inhibition kinetics and potencies presented in Chapter 3.

As presented in Chapters 3 and 5, the inner filter effect was analyzed for our substrate as per Liu et al. [225]. Liu et al. found no inner filter effect from self-absorption of EDANS fluorescence released upon substrate cleavage. However, our concentrations of substrate used and thus concentration of free EDANS are theoretically much higher,
while the model of fluorimeter used here is more sensitive. Additionally, assays are carried out in a cuvette as opposed to a plate reader. For our conditions, increasing concentrations of free EDANS in solution resulted in increasing levels of fluorescence by a proportionally linear relationship (Figure B.1), indicating that at the detectable limits of free EDANS fluorophore there was no evidence for the inner filter effect by absorption.

Although free EDANS alone does not seem to significantly affect absorption of fluorescence, we must also account for the absorption by DABCYL groups in solution. The DABCYL quenching group is released with cleavage of substrate; however, it is also present as part of the intact substrate. Thus, we need to account for IFE at high substrate concentrations. The method used by Liu et al. was again used here[225], whereby the fluorescence from free EDANS fluorophore was compared to that of substrate solution spiked with 0.25 μM free EDANS in order to obtain a correction factor. Measurements were obtained in triplicate and then averaged. Table B.1 contains fluorescence values for substrate and spiked substrate solutions, at increasing substrate concentrations. The correction factors are calculated based on the fact that only with substrate present will the fluorescence from EDANS be somewhat quenched. The correction factor can then be applied to each fluorescence value during calpain activity assays, based on the substrate concentration.

Indeed one can see that there is a drastic reduction in the fluorescence detected above substrate concentrations of 10 μM, well into the concentrations used by Cuerrier et al. [87] and those needed to fit data to Michaelis-Menten kinetics equations. Going
Figure B.1: Fluorescence of increasing concentrations of free EDANS fluorophore in solution. Linear regression was performed with $R^2 = 0.9969$. 
Table B.1: Inner filter effect corrections for (EDANS)-EPLFAERK-(DABCYL) calpain substrate.

<table>
<thead>
<tr>
<th>[S] (μM)</th>
<th>f (S)</th>
<th>f (S + EDANS)</th>
<th>f (EDANS)</th>
<th>correction factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5.26</td>
<td>754.51</td>
<td>749.25</td>
<td>1.00</td>
</tr>
<tr>
<td>0.1</td>
<td>10.49</td>
<td>753.97</td>
<td>743.48</td>
<td>0.99</td>
</tr>
<tr>
<td>0.2</td>
<td>16.15</td>
<td>743.13</td>
<td>726.98</td>
<td>0.97</td>
</tr>
<tr>
<td>0.5</td>
<td>29.94</td>
<td>755.66</td>
<td>725.72</td>
<td>0.97</td>
</tr>
<tr>
<td>1</td>
<td>48.75</td>
<td>777.21</td>
<td>728.46</td>
<td>0.97</td>
</tr>
<tr>
<td>2</td>
<td>72.12</td>
<td>777.27</td>
<td>705.15</td>
<td>0.94</td>
</tr>
<tr>
<td>5</td>
<td>163.81</td>
<td>794.99</td>
<td>631.19</td>
<td>0.84</td>
</tr>
<tr>
<td>10</td>
<td>316.06</td>
<td>881.33</td>
<td>565.27</td>
<td>0.75</td>
</tr>
<tr>
<td>20</td>
<td>383.50</td>
<td>862.84</td>
<td>479.34</td>
<td>0.64</td>
</tr>
<tr>
<td>50</td>
<td>480.32</td>
<td>799.69</td>
<td>319.37</td>
<td>0.43</td>
</tr>
<tr>
<td>100</td>
<td>420.86</td>
<td>611.58</td>
<td>190.72</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Note: $\lambda_{ex} = 335$ nm, $\lambda_{em} = 500$, slit widths = 10 nm.
forward, we can use these correction factors allowing for more accurate fluorescence values and thus more accurate reaction velocities. One problem, however, was that at high enough concentrations the substrate was no longer soluble in solution. Because the concentration of substrate does not meet or exceed the $K_m$ value, calculations of $K_m$, $V_{\text{max}}$, and $k_{\text{cat}}$ will have higher associated error.

In order to convert from RFU values to $\mu$M of product formed, we used the calpain-1 protease core, thereby avoiding autoproteolysis, to cleave a set concentration of substrate to completion. From an average of triplicate reactions, it was determined that each $\mu$M of product/substrate corresponds to 1358 RFU for this fluorimeter. This value was used to calculate reaction velocities in $\mu$M/s for the kinetics presented here and in Chapters 3 and 5.

To determine the kinetics of cleavage of the (EDANS)-EPLFAERK-(DABCYL) fluorescent substrate by calpain-2, fluorescence was monitored as in Chapters 3 and 5, with substrate concentrations in the range of 1-100 $\mu$M. The initial reaction rates were determined by fitting a straight line to the linear portion of the progress curves (first 10 s after calcium addition). With the high sensitivity of the fluorimeter, the first 10 s provided enough data points to record the initial rate of reaction while avoiding the complications of autoproteolysis. The resulting rate data were examined with Michaelis-Menten plots for the analysis of enzyme inhibition. $K_m$ and $V_{\text{max}}$ values for the reactions were obtained from direct fits of the Michaelis-Menten equation to the data using the non-linear least
squares fitting function of gnuplot (version 4.6) [224], in a similar method as Chapters 3 and 5, and the rate of substrate turnover (k\text{cat}) subsequently calculated. For the cleavage of the calpain-specific FRET substrate $K_m = 41.4 \pm 2.3 \, \mu M$, and $k_{\text{cat}} = 2.94 \pm 0.07 \, s^{-1}$.

The kinetics presented here will allow for the accurate comparison of this substrate to others, including novel substrates being developed by Christian-Scott McCartney in our lab (unpublished) as discussed previously. The methods to calculate these values hold valid for any studies involving calpain kinetics. In addition, we can now acknowledge that the turnover rate calculated by Cuerrier et al. is not valid for reasons elaborated upon earlier, although they are close to the values obtained here. The turnover rate for this substrate remains much higher than those for other substrates, as compared by Cuerrier et al. [87], and while kinetics values have been recalculated, the (EDANS)-EPLFAERK-(DABCYL) substrate remains the most preferred for in vitro calpain activity assays.
Appendix C

The potential for “activation” of calpain via inhibition of autoproteolysis

During the analysis of the data presented in Chapter 3, it was found that a few cyclic compounds seemed to “activate” calpain through somehow causing an increase in the initial reaction velocity of cleavage of substrate, or more likely through inhibition of autoproteolysis resulting in a longer enzymatic life-time. Compounds with the potential to activate the enzymes are interesting as they could be useful for in vivo studies on the roles of calpains, while molecules that inhibit autoproteolysis could be equally useful for in vitro studies where autoproteolysis is a complication, such as kinetics and crystallographic structure determination.

As discussed previously, studies on calpain autoproteolysis suggest the process is drastic and fast-acting, reducing the supply of active enzyme[116]. Indeed, the process can be monitored during cleavage of FRET substrates. Loss of enzyme activity by autoproteolysis is evident by a plateau in fluorescence, as the amount of active enzyme is reduced over time. When screening the cyclic peptide compounds in Chapter 3, it was observed that this plateau was delayed with certain compounds and calpain activity increased with others. I then proceeded to follow through on investigations as to how the delay in autoproteolysis may be occurring with the compound cPGLGF.
Cleavage of the FRET substrate (EDANS)-EPLFAERK-(DABCYL) by calpain was monitored as in Chapters 2 and 3. The solution contained 10 nM calpain-2, 5 µM substrate, and 20, 50, and 100 µM cPGLGF and/or 10, and 20 nM B27 peptide. Buffer solution was 10 mM HEPES pH 7.4 and 10 mM DTT. Control assays were performed with an equivalent amount of DMSO since cPGLGF is dissolved in DMSO at stock concentrations of 50 mM. The reaction was initiated with the addition of CaCl$_2$ to 4 mM in a final volume of 0.5 mL.

Calpain was mixed first with increasing concentrations of cPGLGF to observe if this was a dose-dependent reaction. With increasing concentrations of cPGLGF there is an increasing delay of autoproteolysis coupled with activation of the enzyme, especially at concentrations of 50 µM and higher (Figure C.1A). I began to question the mechanism of this activation of enzyme/inhibition of autoproteolysis. We know from previous work that the C2L domain of calpain is a hotspot for autoproteolysis[116]. If cPGLGF could bind and stabilize these targeted regions this could prevent cleavage via autoproteolysis, and potentially result in “activation” of the enzyme. Since structural determination of the binding site of cPGLGF would be difficult, I chose to use the B27 inhibitory peptide to attempt to out-compete the cyclic compound. Figure C.1B shows the progress curves for cleavage of the FRET substrate in the presence of B27 alone, cPGLGF alone, and B27 in combination with cPGLGF. Increasing concentrations of B27 seems to remove the protective, activating effect of cPGLGF. In order to quantify this
Figure C.1: Activity of calpain-2 in the presence of cPGLGF and B27 peptides.

Cleavage of 5 μM (EDANS)-EPLFAERK-(DABCYL) by 10 nM calpain-2 was monitored by fluorescence, with the reaction initiated by 4 mM CaCl₂. Shown are the progress curves for the reactions in the presence of (A) cPGLGF, and (B) cPGLGF and B27, each at concentrations shown. Each reaction was run in triplicate with the averaged values displayed.
effect and compare activity and autoproteolysis, I looked to use decay equations to tease out the rate of autoproteolysis and the rate of product formation.

Relating autoproteolysis to product formation, we can pull apart the rate of the forward reaction and the rate of autoproteolysis from our progress curves (product formation over time). The remaining concentration of enzyme in solution is a result of second-order decay with a rate constant for autoproteolysis, represented by the following equation,

\[ E_t = \frac{E_0}{1 + k_p E_0 t} \]

where \( k_p \) represents the rate of inactivation/autoproteolysis.

Product formation will depend on the reaction velocity and is thus affected by the rate of autoproteolysis as follows,

\[ v_t = \frac{v_0}{1 + k_p v_0 t} \]

where \( v_0 \) is the reaction velocity at \( t=0 \).

Thus the product concentration as a function of time, based on autoproteolysis alone, is represented as follows,

\[ [P] = \frac{\ln(k_p v_0 t)}{k_p} \]
However, the rate of the reaction will also be affected by lowering concentrations of available substrate, and so taking into account autoproteolysis as well gives the equation below,

\[ [P] = S_0 \left( 1 - e^{(-k_2E_0t)} \right) \]

where \( S_0 \) and \( E_0 \) are initial substrate and enzyme concentrations and \( k_2 \) is the rate of forward reaction (\( k_{\text{cat}} \)).

Finally, combining the progress curve equation with that of enzyme decay we obtain the following,

\[ [P] = S_0 \left( 1 - e^{\left( \frac{-k_2E_0t}{1 + k_pE_0t} \right)} \right) \]

When I fit the data for the progress curves to the above equation, I obtained the rate of autoproteolysis and the rate of product formation for each condition in order to accurately and quantifiably compare the effect of each compound on calpain activity. These data can be seen in Table C.1. As can be seen, cPGLGF did not drastically change the rate of product formation. However, the addition of B27 inhibited the enzyme and competed with cPGLGF for binding. Thus, while cPGLGF drastically reduced the rate of autoproteolysis, this effect was reversed by the addition of B27 peptide. B27 seemingly increased the rate of autoproteolysis, but this can be attributed to slow, tight-binding inhibition.
Table C.1: Rate constants comparing product formation and autoproteolysis during calpain “activation”

<table>
<thead>
<tr>
<th>Rate constants</th>
<th>DMSO control</th>
<th>100 μM cPGLGF</th>
<th>10 nM B27</th>
<th>100 μM cPGLGF + 10 nM B27</th>
<th>100 μM cPGLGF + 20 nM B27</th>
</tr>
</thead>
<tbody>
<tr>
<td>product formation ($k_d$)</td>
<td>0.07</td>
<td>0.06</td>
<td>0.02</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>autoproteolysis ($k_p$)</td>
<td>0.38</td>
<td>0.07</td>
<td>1.5</td>
<td>0.15</td>
<td>0.26</td>
</tr>
</tbody>
</table>
The data presented above suggest that cPGLGF binds to the calpain enzyme in a site adjacent to the active site cleft also occupied by the B27 peptide. We can conclude that this would likely be part of the C2L domain, in such a location that it will stabilize the enzyme against autoproteolysis but far enough away to not interfere with the cleavage of small substrates.

We now may have a better idea as to the mechanism for prevention of autoproteolysis by cPGLGF, however other questions remain. Do the other “activating” peptides delay autoproteolysis function in the same manner? How might other compounds increase calpain activity independent of autoproteolysis? It is clear that further investigations are needed into these compounds, both by monitoring calpain activity as well as structural determination.
Appendix D

Crystallography of whole calpain bound to inhibitors

Our group has previously explored the calpain active site cleft through the use of peptide and inhibitor libraries, and through crystallographic studies using just the protease core of calpain-1. The latter avoided problems associated with aggregation and precipitation of full-length calpain [87,89,96,117,119]. While early inhibitors only interacted with the unprimed side of the active site (subsites S1 and S2 primarily), studies in our lab have shown that extensive unprimed-side extensions are important in calpain specificity[87]. The crystal structure of the full-length calpain-2 in complex with calpastatin showed that the C2L domain, missing from the protease core structures, contributes to the S3 subsite[97,98]. In addition, inhibitors could interact with the C2L domain in the unprimed side of the active site cleft if they are long enough. For example, the structure of the α-ketoamide-containing inhibitor SNJ-1945 bound to the calpain-1 protease core shows that the long ether chain on the unprimed side of the inhibitor seems to occupy two different conformations, one of which could potentially interact with the C2L domain when modeled bound to whole enzyme[118]. Another interesting point to note is the very different inhibition between the protease core and full-length calpain; IC50 values are 2.93 ± 0.27 μM for the calpain-1 protease core[122] and 0.051 μM for full-length calpain-2[118]. This could indicate that the long ether chain may indeed interact with the C2L domain as modeled, and aid in inhibition potency and specificity.
While the use of the protease core will remain important for crystallography with inhibitors in order to explore their possible interactions with the C2L domain and exploit this feature for calpain specificity, it is necessary to pursue full-length calpain crystallography.

The thesis written by Rachel Hanna covers current research into self-aggregation of full-length calpain and how to prevent this[355]. While self-aggregation and autoproteolysis are the main problems plaguing full-length calpain crystallography, it is possible to minimize the contribution of each. Subunit-dissociation, previously described as a potential problem for crystallography, is not seen under the conditions we use[116]. However, obtaining a structure with a covalently bound inhibitor introduces complications. When active enzyme is used, calcium-induced aggregation becomes a problem. Autoproteolysis can be minimized through the use of inhibitors and a specific order of addition of reagents. Following the research done by Hanna et al., we have been able to use peptides corresponding to the subdomains A and C from calpastatin to minimize aggregation and precipitation, paving the way for crystallization of full-length calpains. While NaCl has previously been shown to help prevent aggregation[111], the use of A and C peptides also allows us to use non-NaCl containing solutions, greatly increasing the number of useable crystallization conditions. I have been working to crystallize full-length calpain-2 using the A and C peptides to hinder self-aggregation. Despite extensive efforts, crystals have not been obtained of a protein-inhibitor complex or for an enzymatically inactive mutant in the absence of inhibitor. Importantly, not all
drops and conditions produce calcium-activated aggregation and precipitation.

Therefore, future studies using these methods may yet allow for structures to be solved of the full-length enzyme, illuminating possible interaction of inhibitors with the C2L domain in the unprimed side of the calpain active site cleft.

To date, I have attempted to crystallize multiple constructs of calpain, including calpain-1 and -2, most in the presence of inhibitor. The enzymes and the inhibitors, if applicable, are listed with their respective attempted crystallization screens as in Table D.1. Omitted are any expansions around potential hits that did not lead to protein crystals as detected by X-ray diffraction.

It is worth emphasizing that calpastatin does not alter the structure of calpain-2, but rather it recognizes and binds to the active enzyme. The structure of the calcium-bound 21K homodimer [90,91] is very similar to that of the 21K homodimer bound to calpastatin subdomain C[123], and the protease core domains of the enzyme when bound to calpastatin[97,98] is almost identical to the protease core crystallized on its own with calcium present[89]. In contrast, the active form of calpain causes three alpha-helical structures to form in what is otherwise an intrinsically unstructured calpastatin polypeptide[133]. Thus, using the A and C peptides to prevent aggregation should not affect the native structure of active calpain nor should they prevent inhibitors from binding.
Table D.1: Attempted crystallization screens of calpain and calpain-inhibitor complexes.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Inhibitor</th>
<th>A + C peptides?</th>
<th>Protein concentration (in solution before drop)</th>
<th>Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>rat calpain-2 C105S mutant¹</td>
<td>N/A</td>
<td>no</td>
<td>10 mg/mL</td>
<td>PEGs I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no</td>
<td>4 mg/mL</td>
<td>PEGs I, Classics L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no</td>
<td>13 mg/mL</td>
<td>PEGs I, JCSG+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>yes</td>
<td>13 mg/mL</td>
<td>JCSG+, PEGs II</td>
</tr>
<tr>
<td>rat calpain-2 WT¹</td>
<td>SNJ-1945</td>
<td>no</td>
<td>10 mg/mL</td>
<td>Classics L, PEGs I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>yes</td>
<td>10 mg/mL</td>
<td>MPD, Microlytic crystal formers (CombiScreen, Smart Screen, Pure PEGs), Cations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no</td>
<td>13 mg/mL</td>
<td>PEGs I, JCSG+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>yes</td>
<td>13 mg/mL</td>
<td>PEGs I, JCSG+, PEGs II</td>
</tr>
<tr>
<td></td>
<td>WR-18</td>
<td>no</td>
<td>13 mg/mL</td>
<td>PEGs I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>yes</td>
<td>10 mg/mL</td>
<td>JCSG+, PEGs I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>yes</td>
<td>10 mg/mL</td>
<td>Crystal formers (SmartScreen, Pure PEGs), Cations</td>
</tr>
<tr>
<td>Protein Type</td>
<td>Inhibitors</td>
<td>Concentration</td>
<td>Solution</td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>------------</td>
<td>---------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>rat calpain-2 C205S (with Type-3 fused to 21K)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>N/A, yes, 6 mg/mL</td>
<td>PACT, PEGs I, PEGs II</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>yes, 3 mg/mL</td>
<td>PEGs I, PEGs II</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no, 6 mg/mL</td>
<td>PACT</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no, 10 mg/mL</td>
<td>PACT</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no, 3 mg/mL</td>
<td>PACT</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>human calpain-2 ΔG1&lt;sup&gt;3&lt;/sup&gt;</td>
<td>SNJ-1945, yes, 5 mg/mL</td>
<td>PEGs II</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>yes, 5 mg/mL</td>
<td>PACT, PEGs I</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E64, yes, 5 mg/mL</td>
<td>PEGs I, PEGs II</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td></td>
<td>leupeptin, yes, 5 mg/mL</td>
<td>PEGs I, PEGs II</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

1 [88]
2 [356]
3 [357]

Note that 5 mM CaCl<sub>2</sub> was added to all conditions, and if applicable, inhibitors were at 2 mM while A and C peptides were at 0.5 mM in the protein solution.
Appendix E

**Determination of the inhibitor binding site via photoreactive labeling**

Attempts at full-length calpain-2 crystallography have proved difficult in the past, and protein structures of calpain have mostly depended on covalent complexes with the calpain-1 protease core. The cyclic peptide inhibitors discussed in Chapter 3 do not covalently bind the enzyme, but strategizing the attachment of covalent warheads is a possibility in the case of competitively inhibiting compounds. While it would be extremely beneficial to have full-length calpain bound to cyclic peptide, attempts are being made to make photo-reactive cyclic peptides through photo-reactive amino acids or chemical moieties, and using mass spectrometry for the mapping of peptide attachment sites onto the known structure of active calpain. The determination of ligand binding sites on calpain is necessary for further structure-based drug design, and is especially important considering the struggles with whole-enzyme crystallography especially when we want to investigate the binding pocket of non-competitive inhibitors, such as amPGLdO.

For this reason, I looked to previous work done by O'Neil et al. with peptide ligands of calmodulin for inspiration[358,359]. amPGLdO was synthesized with a maleimide-linked benzophenone UV reactive group. A reaction solution was set up with 0.5 mg/mL C105S calpain-2 (~ 5 μM) and 25 or 50 μM amPGLdO in 10 mM HEPES pH 7.4, 10 mM DTT, 4 mM CaCl₂. The solution was set up in 96-well Corning polystyrene
plates and allowed to react under long-wave UV for 2.5 h. The reaction products were sent to the Protein Function Discovery Facility at Queen’s University for peptide mass fingerprint analysis digested with trypsin and subjected to MALDI-TOF. The resulting peptide masses were compared to those from a DMSO negative control reaction. I looked for a difference of approximately 880 Da between the covalently-attached amPGLdO and “native” peptides. Five peptides were found matching the molecular weight for calpain tryptic digest with covalently-linked amPGLdO, and these were mapped to the calpain-2 crystal structure (PDB: 3BOW) [97](Figure E.1).

One can see that of the tryptic digest peptides, three map to regions consistent with the noncompetitive nature of amPGLdO. Two others map to PEF domain clefts and, as described in Chapter 5, are highly unlikely to cause inhibition of calpain through amPGLdO binding. Thus, it is probable that amPGLdO binds to regions on calpain interacting with the protease core domains but remaining out of the active site cleft, and could function by either impeding the conformational change required for calcium-activation of the enzyme or . This potential novel allosteric binding location can be further supported through the use of proper control conditions. In order to determine the accuracy of mapping the binding site, a photoreactive version of the competitive inhibitor cPGALK was made and was reacted with C105S calpain-2 in a similar method as described for amPGLdO. The reactions will be analyzed by tryptic digest and MALDI-TOF peptide analysis. Should cPGALK be mapped to the active site cleft, as per the
**Figure E.1: Potential binding sites for amPGLdO on calpain-2.** The crystal structure of calpain-2 is shown with the catalytic cleft in the front (A) and back (B). Highlighted in red are peptides identified as potential but unlikely binding locations for amPGLdO. Highlighted in royal blue, navy blue, and purple are likely binding locations for amPGLdO.
enzyme inhibition mechanism, we could gain confidence in this method of binding site mapping via photoreactive labeling.