# Structural Insight into the Development of Novel Allosteric Calpain-1 Inhibitors

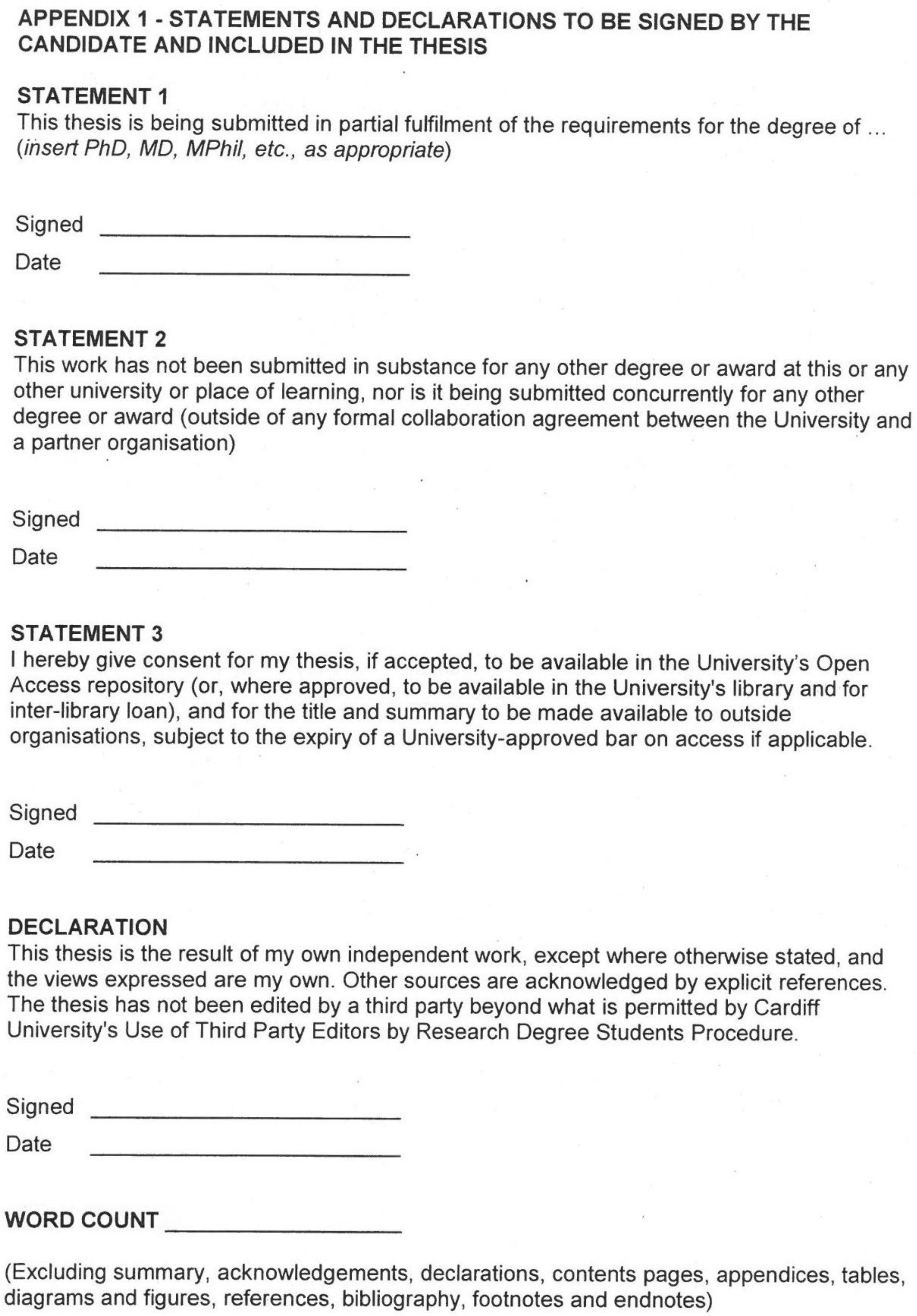
by

Joel Cresser-Brown

A thesis submitted to Cardiff University for the degree of Doctor of Philosophy

School of Chemistry
Cardiff University
April 2019

# Declaration



## APPENDIX 1 - STATEMENTS AND DECLARATIONS TO BE SIGNED BY THE CANDIDATE AND INCLUDED IN THE THESIS

### STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of … (*insert PhD, MD, MPhil, etc., as appropriate*)

Signed ____________________________
Date ____________________________

### STATEMENT 2

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is it being submitted concurrently for any other degree or award (outside of any formal collaboration agreement between the University and a partner organisation)

Signed ____________________________
Date ____________________________

### STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available in the University's Open Access repository (or, where approved, to be available in the University's library and for inter-library loan), and for the title and summary to be made available to outside organisations, subject to the expiry of a University-approved bar on access if applicable.

Signed ____________________________
Date ____________________________

### DECLARATION

This thesis is the result of my own independent work, except where otherwise stated, and the views expressed are my own. Other sources are acknowledged by explicit references. The thesis has not been edited by a third party beyond what is permitted by Cardiff University's Use of Third Party Editors by Research Degree Students Procedure.

Signed ____________________________
Date ____________________________

**WORD COUNT** ______________________

(Excluding summary, acknowledgements, declarations, contents pages, appendices, tables, diagrams and figures, references, bibliography, footnotes and endnotes)

# Acknowledgements

Firstly I would like to extend my gratitude towards my supervisor Professor Rudolf Allemann for allowing me to undertake this exciting project. Dr. David Miller I would like to thank for your support during this project and proofreading this work. Dr. Robert Mart for providing excellent experimental advice for this project. Chris and Big Rob without your good companionship, humour and wanders my time in Cardiff would not have been the same.

I would also like to thank my collaborators Professor Maurice Hallett and Dr Emma Robinson with your help in the live neutrophil imaging experiments. Dr Pierre Rizkallah, without you I would not have been able to acheive the crystal structures described in this work, broadening my horizons to the world of structural biology.

Finally I would like to thank my friends and family who have been a great help throughout the course of my studies. Mum and Dad your unremitting support have got me to where I am today. And Holly for somehow putting up with me for the duration of my PhD.

# Abstract

Calpain-1 is a calcium activated cysteine protease involved in a diverse range of physiological processes, such as enabling the rapid cell spreading of neutrophils during the chemotactic response to tissue damage. Rheumatoid arthritis is a pathological condition associated with the neutrophil spreading process, rendering calpain-1 a prized therapeutic target. Small molecule α-mercaptoacrylic acid inhibitors exhibit selectivity towards calpain-1, however differences in potency were reported for the compounds in different redox states. The dihedral bond geometry of the more potent oxidised disulfide form enabled an increased binding interaction with the calcium binding domain, PEF(S), however the less potent thiol form would be predominantly present in the reducing environment of the cell.

In this project, diselenide analogues of the α-mercaptoacrylic acid inhibitors were synthesised and evaluated as allosteric, redox stable calpain-1 inhibitors. They were shown to inhibit human calpain-1 with moderate potency, binding to the calcium binding domains PEF(S) and PEF(L) with similar potencies to the sulfur counterparts, whilst X-ray co-crystal structures of PEF(S) with the diselenides showed that the dihedral bond geometry was analogous to the disulfides. The calpain-1 active site, CysPC, was expressed and purified as a truncated catalytic protein. The diselenides inhibited CysPC with a greater potency than the full length enzyme, suggesting they are not allosteric in their mode of action. CysPC was crystallised in the presence of a diselenide inhibitor, the electron density map suggested the presence of adducts in the active site cleft was responsible for the inhibition formed by selenyl-sulfide bonds with surface exposed cysteine residues.

Computational docking experiments using the calcium binding domain PEF(S), identified ten compounds that were shown to bind to PEF(S) by a fluorescence assay. Of these compounds, three displayed an apparent allosteric mode of action by inhibiting the calpain-1 full-length enzyme but not the active site. The allosteric inhibitors did not inhibit calpain-2 up to a concentration of 100 μM indicating moderate isoform selectivity which was attributed to an interaction with the calcium binding domain PEF(L). This finding may serve as a novel target in the development of isoform selective, allosteric inhibitors.

## Table of Contents

The following standard abbreviations were used for the amino acids:

| | | |
|---|---|---|
| Alanine | Ala | A |
| Arginine | Arg | R |
| Asparagine | Asn | N |
| Aspartic acid | Asp | D |
| Cysteine | Cys | C |
| Glutamine | Gln | Q |
| Glutamic acid | Glu | E |
| Glycine | Gly | G |
| Histidine | His | H |
| Isoleucine | Ile | I |
| Leucine | Leu | L |
| Lysine | Lys | K |
| Methionine | Met | M |
| Phenylalanine | Phe | F |
| Proline | Pro | P |
| Serine | Ser | S |
| Threonine | Thr | T |
| Tryptophan | Trp | W |
| Tyrosine | Tyr | Y |
| Valine | Val | V |

## Abbreviations

| | |
|---|---|
| A | Adenosine |
| AcOH | Acetic acid |
| Acyl-CoA | Acyl-coenzyme A |
| AD | Alzheimer's disease |
| ADMET | Absoption, distribution, metabolism, excretion and toxicity |
| AMC | 7-amido-4-coumarin |
| APCI | Atmospheric pressure chemical ionisation |
| APS | Ammonium persulfate |
| ATP | Adenosine triphosphate |
| BE | Binding energy |
| BSA | Bovine serum albumin |
| C | Cytosine |
| c | Concentration |
| CAPNS1 | Calpain small subunit 1 |
| CAST | Calpastatin |
| CBSW | Calpain-type beta sandwich domain |
| CCP4 | Collaborative Computational Project 4 |
| CD | Circular dichroism |
| CNS | Central nervous system |
| COOT | Crystallographic Object-Oriented Toolkit |
| COSHH | Control of substances hazardous to health |
| Cryo-EM | Cryo electron microscopy |
| CV | Column volume |
| CysPC | Catalytic domain of calpain |
| Da | Dalton |
| DABCYL | 4-((4-(dimethylamino)phenyl)azo)benzoic Acid |
| DG | delta G, ΔG |
| $dH_2O$ | Deionised water |
| DMARDs | Diseases modifying antirheumatic drugs |
| DMF | *N,N'*-Dimethylformamide |
| DMSO | Dimethylsulfoxide |
| DNA | Deoxyribose nucleic acid |
| DS | Disulfide bonds |
| DTT | Dithiothreitol |
| *E. coli* | Escherichia coli |
| ECM | Extracellular matrix |
| EDANS | 5-((2-Aminoethyl)amino)naphthalene-1-sulfonic acid |
| EDTA | Ethylenediaminetetraacetic acid |
| EFH | EF hand |
| EI | Electron ionisation |
| ER | Endoplasmic reticulum |
| ES | Electrospray ionisation |
| ES-MS | Electrospray mass spectrometry |

| | |
|---|---|
| FAM | 6-Carboxyfluorescein |
| fMLP | N-Formylmethionine-leucyl-phenylalanine |
| FPLC | Fast protein liquid chromatography |
| FRET | Fluorescence resonance energy transfer |
| G | Guanosine |
| GAAP | Golgi anti-apoptotic proteins |
| GB1 | B1 domain of *Streptococcal* protein G |
| GC-MS | Gas chromatography–mass spectrometry |
| GPCR | G protein-coupled receptors |
| GR | Glycine rich |
| GSH | Glutathione in reduced form |
| GSSG | Glutathione in oxidised form |
| HB | Hydrogen bond |
| HBA | Hydrogen bond acceptor |
| HBD | Hydrogen bond donor |
| HEPES | (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) |
| Hfq | Host Factor for phage Q β-RNA replication |
| IA | Interface area |
| $IC_{50}$ | The half maximal inhibitory concentration |
| IL-6 | Interleukin-6 |
| IMAC | Immobilised metal affinity chromatography |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| IR | Infrared spectroscopy |
| ITC | Isothermal titration calorimetry |
| kb | kilobasepairs |
| $K_d$ | Dissociation constant |
| kDa | kilo Dalton |
| $K_i$ | Inhibition constant |
| $K_iPO_4$ | Potassium phosphate buffer |
| l | Path length |
| LB | Lysogeny broth/Luria-Bertani media |
| MALDI | Matrix Assisted Laser Desorption/Ionisation |
| MeCN/CAN | Acetonitrile |
| MeOD | Deuterated methanol |
| MeOH | Methanol |
| mL | Millilitre |
| mM | Millimolar |
| MMP-2 | Matrix metalloproteinase-2 |
| mp | Melting point |
| MRE | Mean residue ellipticity |
| mRNA | Messenger RNA |
| MW | Molecular weight |
| MWCO | Molecular weight cutoff |
| n | Number of peptide bonds |
| n-BuLi | n-Butyllithium |

| | |
|---|---|
| NADPH | Nicotinamide adenine dinucleotide phosphate |
| nM | Nanomolar |
| nm | Nanometre |
| NMR | Nuclear magnetic resonance |
| OD | Optical density |
| PAINs | Pan-assay interference compounds |
| PCR | Polymerase chain reaction |
| PDB | Protein data bank |
| PEF | Penta-EF |
| PEG | Polyethylene glycol |
| PISA | Protein Interfaces, Surfaces and Assemblies |
| pKa | Acid dissociation constant at logarithmic scale |
| pM | Picomolar |
| PV | Hydrophobic P-value |
| R factor | Residual factor |
| r.m.s.d | Root mean square deviation |
| RA | Rheumatoid arthritis |
| RCF | Relative centrifugal force |
| RNA | Ribonucleic acid |
| Ro5 | Rule of 5 |
| ROS | Reactive oxygen species |
| SAR | Structure activity relationship |
| SB | Salt bridges |
| SDS | Sodium dodecyl sulfate |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SOCE | Store operated calcium entry |
| T | Thymidine |
| TCEP | Tris(2-carboxyethyl)phosphine |
| TEMED | *N,N,N',N'*-Tetramethylenediamine |
| TEV | Tobacco Etch Virus |
| TFA | Trifluoroacetic acid |
| TNFα | Tumour necrosis factor α |
| TNS | 6-(*p*-toluidino)-2-naphthalenesulfonic acid |
| TOF | Time of flight |
| Tris | Tris(hydroxymethyl)aminomethane |
| tRNA | Transfer RNA |
| UV-Vis | Ultraviolet-visible |
| V | Volts |
| vdW | van der Waals force |
| β-ME | β-mercaptoethanol |

| | |
|---|---|
| ε | Extinction coefficient |
| θ | CD signal (millidegrees) |

| | |
|---|---|
| $\lambda_{max}$ | Maximum absorbtion wavelength |
| μM | Micromolar |
| μm | Micrometre |
| ν | Wavenumber |

## List of Figures

## List of Tables

# Chapter 1 - Introduction

# 1.1. Introduction

## 1.1.1. Rheumatoid arthritis

Rheumatoid arthritis (RA) is a widespread autoimmune disease where overstimulation of the immune system can lead to chronic progressive joint destruction and severe disability as well as secondary effects such as an increased susceptibility to cardiovascular disease (Figure 1.1.1).[1,23] RA causes the synovial membrane within diarthrodial joints to be infiltrated by white blood cells such as leukocytes and lymphocytes, this process is initiatiated by antigen-presenting cells.[4] These cells present peptide-based antigens that bind to specific receptors on T helper cells, activating macrophages and stimulating the secretion of proinflammatory cytokines such as interleukins and tumour necrosis factor α (TNFα).[5] The release of these small signaling proteins attracts synovial fibroblasts, neutrophils and chondrocytes to release highly destructive metalloproteases and reactive oxygen species into the synovial membrane, resulting in widespread damage around the joint.[6] The infiltration of the tissue by white blood cells eventually causes the synovial lining to grow and extend around the site of inflammation, which forms an invasive granulation tissue called pannus, coming from the Latin word for ‘cloth’.[5,7] The decreased viscosity of the synovial fluid thus results in a severe loss of lubrication and intense discomfort for the patient. Synovial fibroblasts can undergo long-distance cell migration, spreading RA to unaffected joints furthering the discomfort caused to the patient meaning the patient must be tended to swiftly.[8,9]

Current RA treatments range from treating discomfort with painkillers, to disease modifying antirheumatic drugs (DMARDs) that slow progression using anti-inflammatory drugs or steroids.[2,4] As of 2017 an antibody for the cytokine interleukin-6 (IL-6), tocilizumab, was approved as treatment for RA in 130 countries.[10] IL-6 must be administered intravenously and works by binding to IL-6, blocking IL-6 mediated signal transduction which results in an effective and sustained reduced inflammatory response, however there is still a preference for orally bioavailable alternatives.[10]

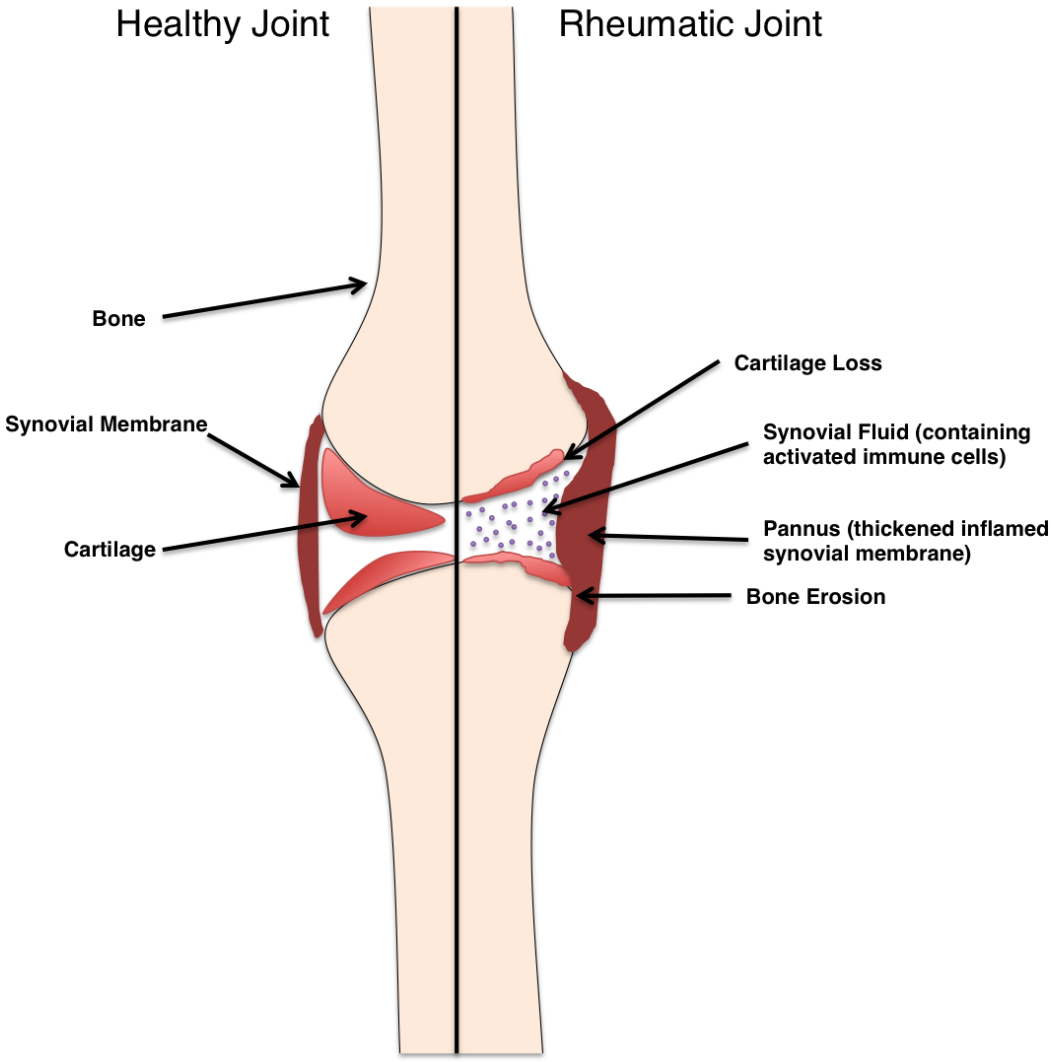


*Figure 1.1.1 - Diagram representing a healthy synovial joint (left) and the destruction caused by rheumatoid arthritis (right).*[1,4]

### 1.1.2. Neutrophils

Neutrophils are the most widespread form of white blood cells, they are spherical polymorphonuclear granulocytes 10-15 μm in diameter, granules in the cytoplasm contain toxins, metalloproteinases and reactive oxidants which are used for the destruction of foreign bodies.[11] Once a chemoattractant is detected, neutrophils respond to the signal by travelling to the site of infection with the aid of adhesion peptides called integrins.[4] The instance these phagocytes encounter a foreign body, they encapsulate it into a vacuole containing the toxic granules, releasing the contents, destroying the invading species.[12]

A site of infection or damage causes the release of chemoattractants such as interleukins, chemokines and *N*-formyl peptides produced by pathogens.[13] In order to reach a site of infection or damage, neutrophils must first travel through the walls of a blood vessel requiring them to undergo a major morphological change.[11]

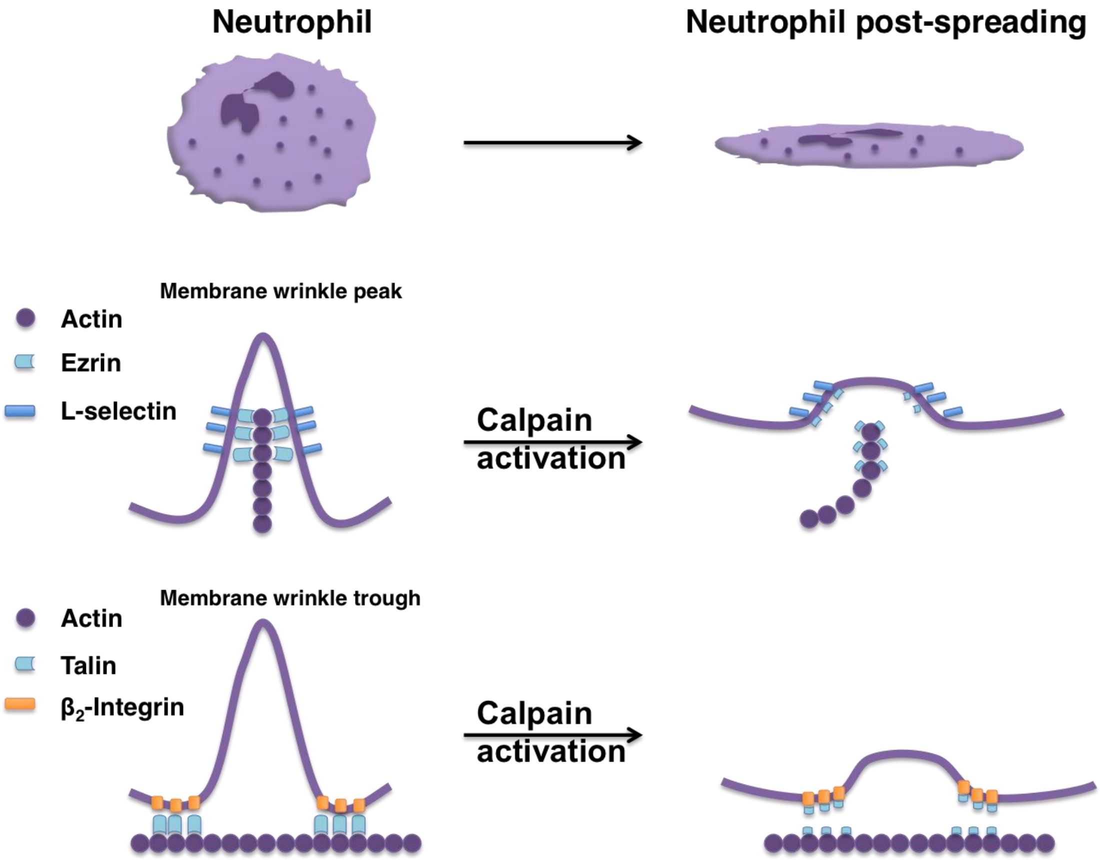


*Figure 1.1.2 - Neutrophil spreading process from spherical to flattened structure (top), cartoons of membrane proteins L-selectin/β2-integrin tethered to the actin cytoskeleton via protein bridges ezrin and talin cleaved upon calpain activation (middle and bottom).*[1,11]

Along the cell membranes of neutrophils are wrinkles held in place by membrane associated proteins L-selectin and β$_2$-integrin thethered to the actin cytoskeleton to the peaks and troughs of the wrinkles, respectively.[11] They are attached to the cortical actin cytoskeleton through two protein bridges ezrin and talin.[11] In response to chemoattractants, intracellular Ca$^{2+}$ is released causing the activation of a regulatory protease, calpain, at the plasma membrane which cleaves the two proteins ezrin and talin which unfurls the wrinkles along the membrane, allowing the cells to flatten and proceed to the site of inflammation by extravasation.[1,11,12] Without the action of the $Ca^{2+}$-activated cysteine protease, calpain, this cell spreading process cannot readily occur as the membrane remains attached to the actin cytoskeleton, which hinders the mobility of the neutrophil and limits the possible response to a chemoattractant.[11,14]

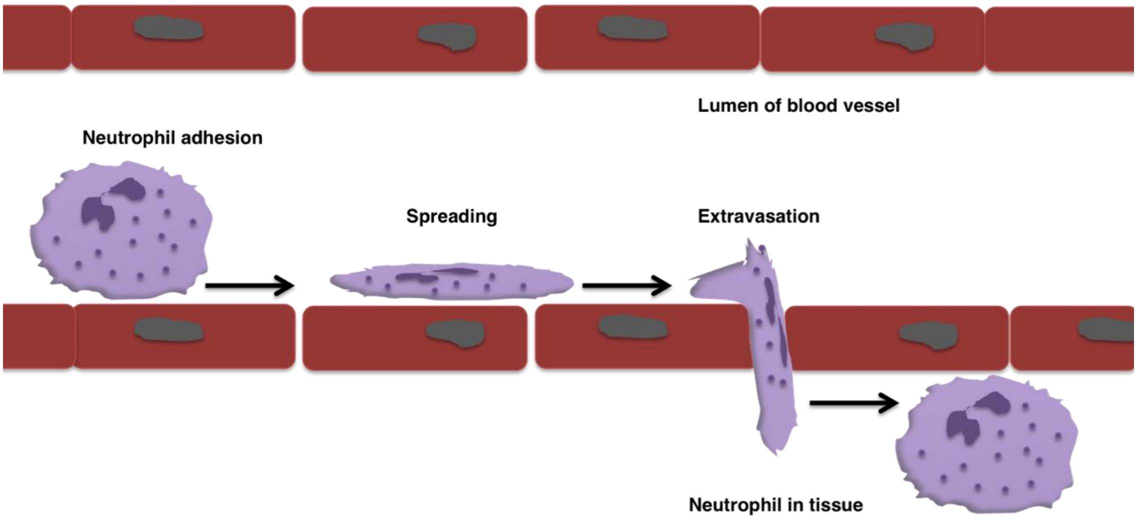


*Figure 1.1.3 - Neutrophil responding to chemotactic inflammatory response through the process of adhering to the blood vessel, spreading, and extravasation out of the blood vessel to a site of inflammation in a tissue.*[11,13,14]

The neutrophil can move forward while adhered to a surface, using actin polymerisation as a driving force on the protruding membrane of the cell, simultaneously releasing the trailing edge of the cell from the surface, allowing a net movement forwards.[15] This basic principle of cell movement is based upon the Elastic Brownian Ratchet Model of actin polymerisation, calpain activity is also important for this chemotactic repsonse of neutrophils to travel to the site of infection or inflammation.[16]

### 1.1.3. Calpain superfamily

Calpains are a family of intracellular $Ca^{2+}$ dependent heterodimeric cysteine proteases found in almost all eukaryotes.[17] They are a member of the papain superfamily that includes caspases, cathespins and papain, amongst others and their activity is strictly regulated.[17] There are 15 known isoforms, some are expressed in certain tissues whilst the two most studied are the ubiquitously expressed isoforms calpain-1 and calpain-2.[18] Proteolytic systems such as ubiquitin-proteasome generally degrade substrate proteins for the purpose of destruction, however calpains are often referred to as ‘modulator’ proteases, whereby their action transforms the function of a substrate in a controlled manner.[19] Calpain substrates play a part in a wide range of cellular processes including cytoskeletal remodeling, apoptosis and cell mobility.[20–22]

The most widely studied calpains -1 and -2 are comprised of an ~ 80 kDa large subunit (CAPN1 and CAPN2 respectively) consisting of an N-terminal α-helical domain, the proteolytic core (CysPC), an acidic phospholipid binding domain or β-sandwich like domain (C2L/CBSW) and the penta-EF hand domain that binds $Ca^{2+}$, (PEF(L)).[23] The 28 kDa small regulatory subunit (CAPNS1) is common between calpain-1 and -2 and that contains a second $Ca^{2+}$ binding domain (PEF(S)) and a glycine rich region.[17] The PEF(L) and PEF(S) domains comprise five penta-EF hand motifs, four of which bind calcium ions and the fifth links the two via an ‘EF handshake’ interaction see Figure 1.1.6. [17,23]

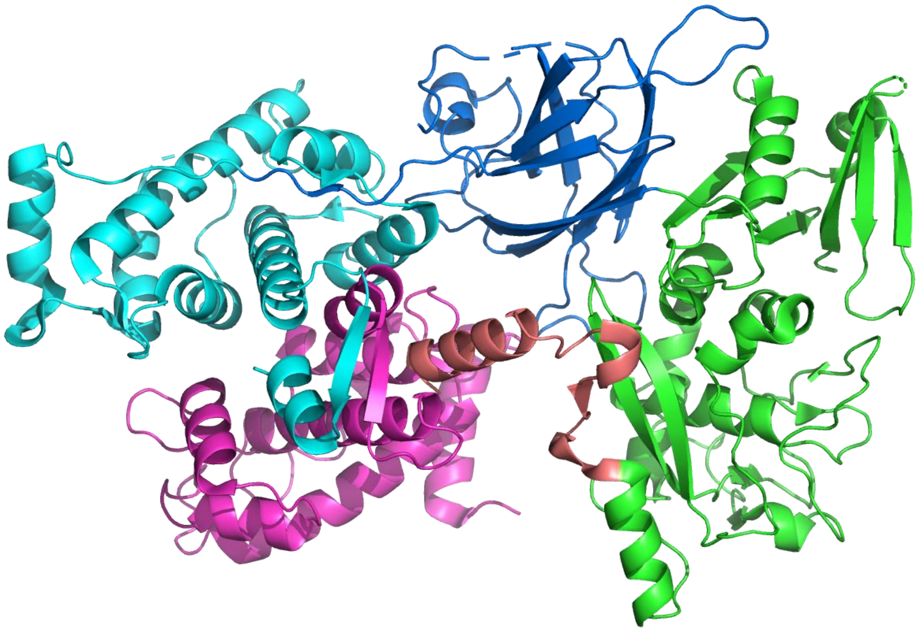


*Figure 1.1.4 - Cartoon representation of the domain structure of calpain-2 (PDB:1DF0). The large subunit anchor helix (salmon pink), CysPC (green), CBSW / C2L (blue) and PEF(L) (cyan). The small subunit GR domain (grey) and PEF(S) (magenta), PyMOL was used to render images based on the PDB coordinates file, as are all subsequent protein images shown.[24]*

### 1.1.4. Calpain proteolytic mechanism

When subjected to an increase in free cytoslic $Ca^{2+,}$ a conformational change brings together the catalytic triad comprising of a cysteine, histidine and asparagine residue situated in the active site cleft.[20] The asparagine manoeuvres the histidine residue into a favourable orientation for facile deprotonation of the cysteine residue, the thiolate then attacks the carbonyl group of the peptide substrate, the thioester is hydrolysed and in doing so, regenerates the cysteine residue.[25] Within the substrate binding site in calpain, a hydrophobic S2 site is shared with other cysteine proteases, which has complicated the design of specific calpain inhibitors.[26]

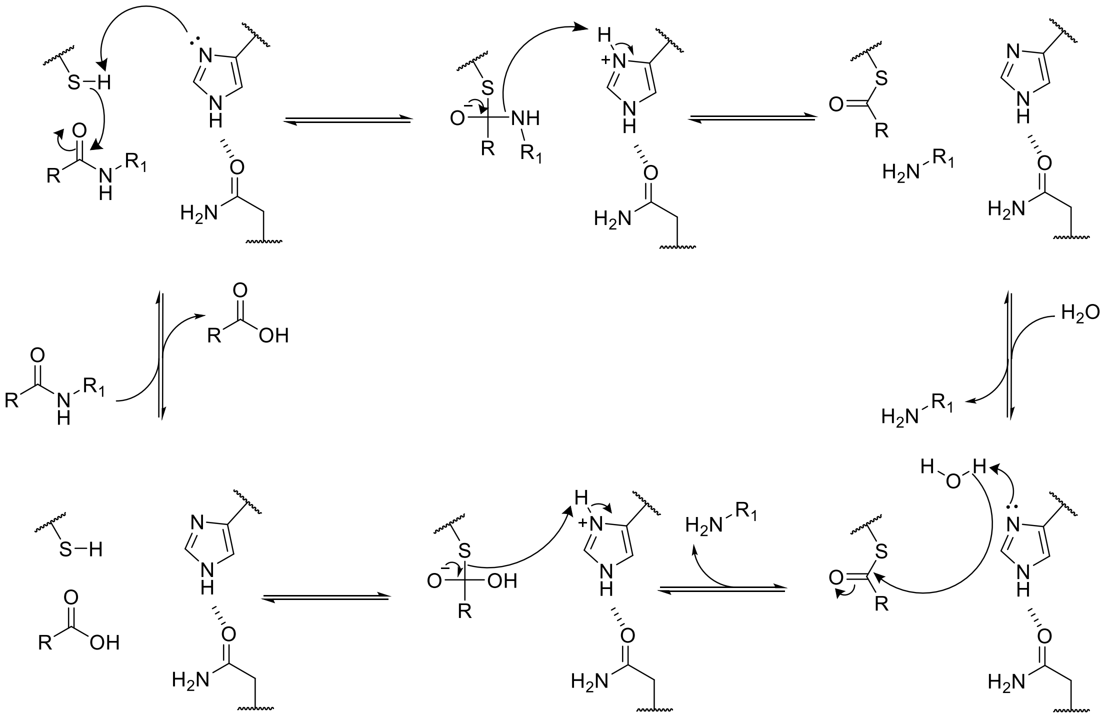


*Figure 1.1.5 – Catalytic mechanism of proteolysis by a cysteine protease, cleavage of the peptide bond by a reactive thiolate formed through deprotonation by a nearby histidine residue, which is held in a favourable conformation through hydrogen bonding to an asparagine residue.*

## 1.1.5. Calpain-1 and -2

The sequence similarity of the large subunit of calpain -1 and -2 isoforms is 62%, with CAPN1 comprising of 714 residues (81,889 Da) and CAPN2 slightly smaller with 700 residues (80,005 Da).[27] The concentration of calcium ions required for activation of calpain varies; ~ 50 μM for calpain-1 and 0.35 mM for calpain-2, also giving rise to the other commonly used names for the isoforms, μ-calpain and m-calpain, respectively.[17,28] Disruption of CAPN2 gene expression in mice prevented progression of the embryo past the blastocyst stage, while CAPN1 silenced mice still developed into normal functioning fertile animals, therefore it is clear despite sharing a small subunit and significant sequence similarity, the isoforms play hugely different roles *in vivo*.[29]

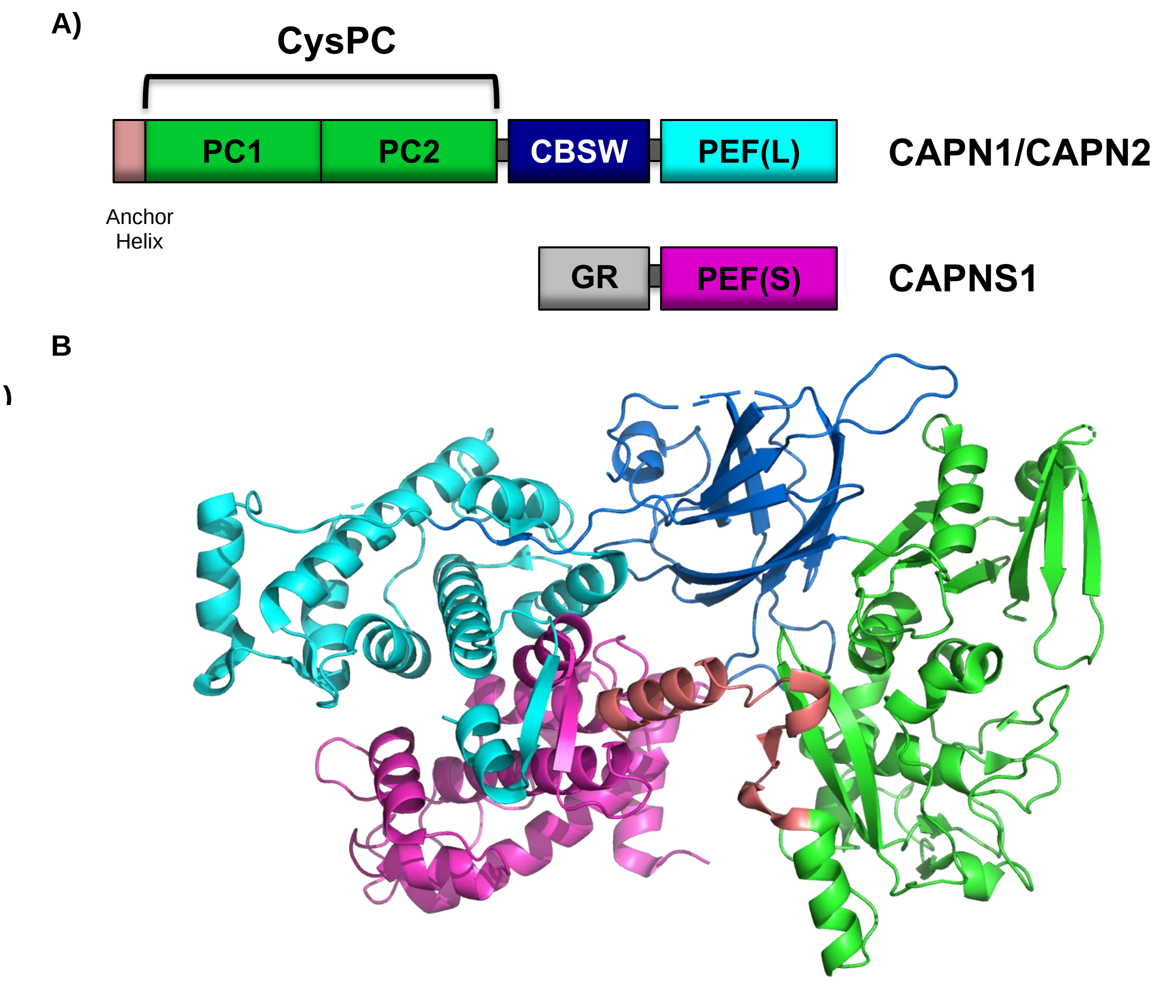


*Figure 1.1.6 –* ***A)*** *The genes representing CAPN1/2 and CAPNS1 with the associated domains respectively.* ***B)*** *Cartoon representation of the domain structure of calpain-2 represented by cartoon (PDB:1DF0). The large subunit anchor helix (salmon pink), CysPC (green), CBSW / C2L (blue) and PEF(L) (cyan). The small subunit GR domain (grey) and PEF(S) (magenta).* [24]

## 1.2. Domain Structure and Properties

To date there have been no published structures of calpain-1, although this data would be useful in the design of allosteric isoform-specific calpain inhibitors. Crystal structures have been obtained of PEF(S), calpain-1 CysPC as well as a hybrid crystal structure containing the N-terminal α-helix and PEF(L) from calpain-2 and the CysPC and calpain-type β-sandwich like domain (CBSW) of calpain-1, called µ-like-calpain.[30] Whilst solving this hybrid structure was a success, it is not a sufficient replacement for the absolute crystal structure of wild-type calpain-1 structure for mechanistic and drug design purposes. Because of this, the structural analysis of the domains in the full length protein will be based on the full length calpain-2 structure for the purpose of this project.

### 1.2.1. The N-terminal anchor α-helix

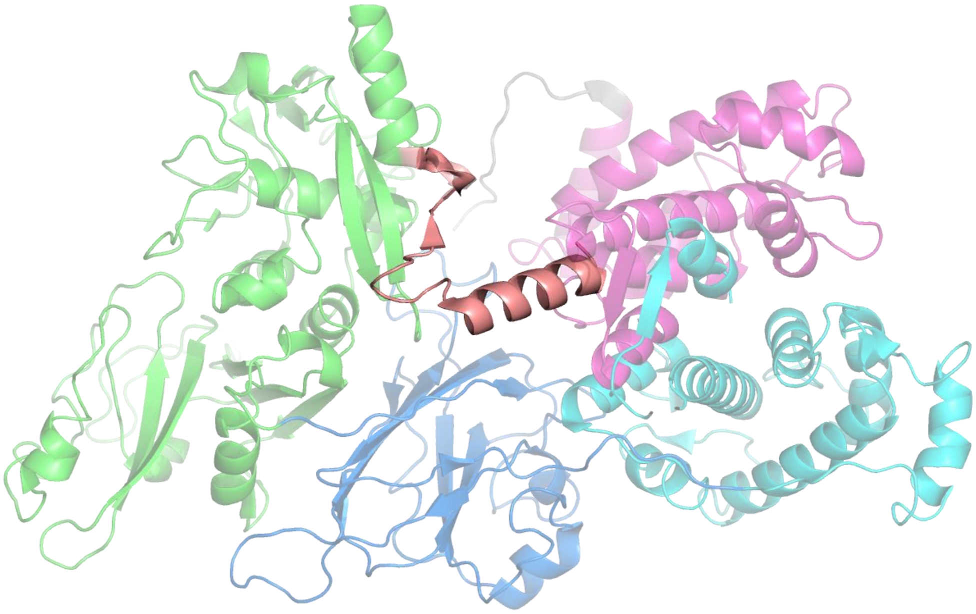


*Figure 1.2.1 – Cartoon representation of the secondary structure of the N-termainal anchor α-helix (salmon pink) within the full length calpain-2 crystal structure (PDB:1KFU).[18]*

The calpain-1 N-terminal anchor contains an amphipathic α-helical domain. Partial localization of calpain-1 has been observed in mitochondria and in this domain is a mitochondrial targeting sequence in the first 24 amino acids, not present in calpain-2.[31] Removal of this terminus blocks the mitochondrial import, whilst transferring it to another protein enables the mitochondrial import.[31] Upon activation of calpain-1, 26 residues from the N terminus are autolysed, removing the targeting sequence from the protein, although the exact function of this enzyme in the mitochondrion have not been determined, calpain is known to play significant roles in apoptosis, as discussed in section 1.3.1.[32,33]

Structural analysis of various conformations of calpain-2 shows the α-helix interacts with the PEF(S) and PEF(L) domains, stabilising the inactive conformation.[31] Upon Ca$^{2+}$ binding, the α-helix releases the protease core, allowing the active site residues to maneuver into the necessary proximity for substrate cleavage, autolysis of the anchor α-helix enables activation of calpain with a lower $Ca^{2+}$ requirement.[32,34] Following this, the heterodimer is believed to lose stability and subunit dissociation occur.[24,35,36]

### 1.2.2. CysPC proteolytic domain

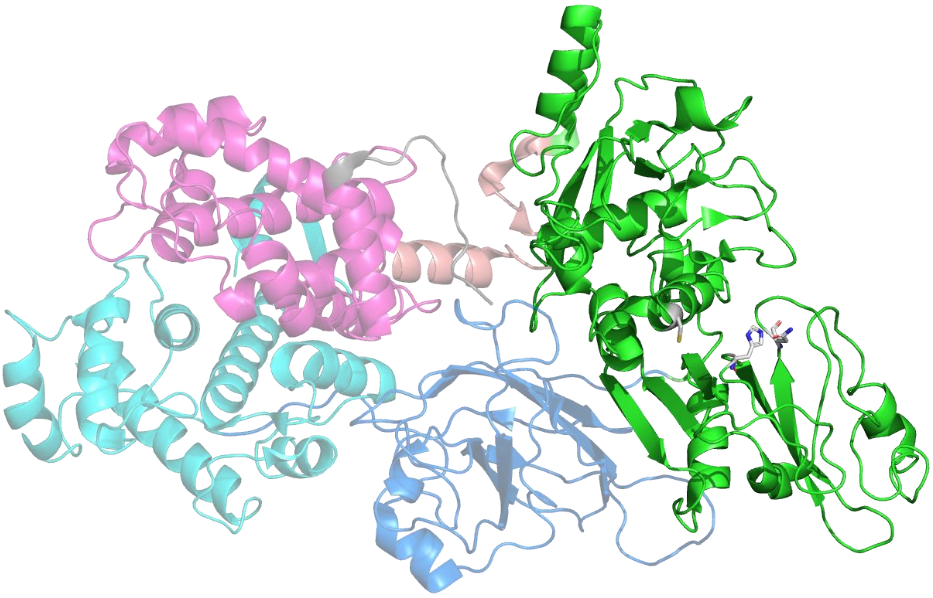


*Figure 1.2.2 - Cartoon representation of the domain structure of calpain-2. The large subunit anchor helix (salmon pink), CysPC (green), CBSW / C2L (blue) and PEF(L) (cyan). The small subunit GR domain (grey) and PEF(S) (magenta), calpain-2 active site (green) (PDB:1KFU), with the active site residues represented as grey sticks.*[18]

The proteolytic domain of calpain is comprised of two subdomains called PC1 and PC2 that form the active site cleft containing approximately 350 residues in total.[36] PC1 contains the nucleophilic cysteine (Cys) residue whilst the PC2 contains the basic histidine (His) and asparagine (Asn) residue that complete the triad.[36]

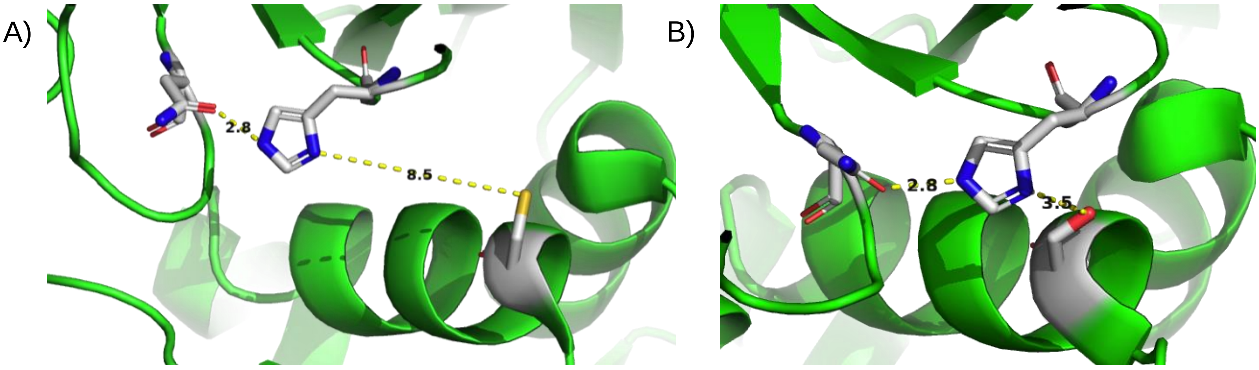


*Figure 1.2.3 -* ***A)*** *Distance between the calcium free calpain-2 active site residues Cys105, His262 and Asn286 (PDB:1KFU)* ***B)*** *Distance between the calcium bound active state calpastatin-calpain-2 C105S complex residues Ser105, His262, Asn286.* [18,37]

The active site triads are in a comparable location for calpain-1 and -2, for calpain-1 Cys 115, His 272 and Asn 296[36] whereas for calpain-2 Cys 105, His 262 and Asn 286.[38] PC1 and PC2 are both comprised of a series of α-helices and two antiparallel β-sheets, with the catalytic triad found on the interface between the two subdomains.[24] CysPC also binds two $Ca^{2+}$ ions, and is active and stable as a standalone protease, without any other calpain domains present.[39] The calpain active site shares similar secondary structure with other cysteine proteases such as papain, which subsequently has hindered the development of selective calpain inhibitors.[40]

### 1.2.3. Calpain-type β-sandwich like domain (CBSW)

This phospholipid binding domain is composed of eight anti-parallel β-sheets and links the $Ca^{2+}$ binding domains to the proteolytic core.[20] It participates in the regulation of calpain activity *via* electrostatic interactions due to the presence of a highly negatively charged acidic loop.[41] This is believed to be *via* $Ca^{2+}$ dependant phospholipid binding which facilitates the translocation of the enzyme from the cytosol to the cell membrane where the substrate cleavage reactions can to occur.[41]

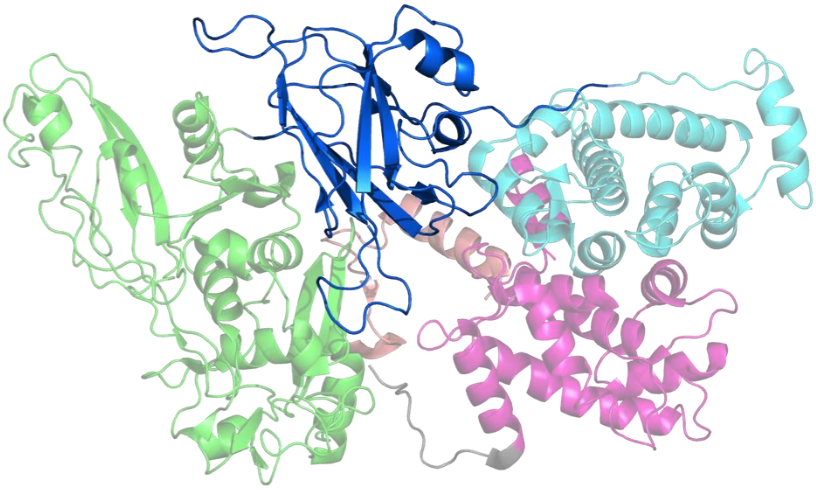


*Figure 1.2.4 – Cartoon representation of the calcium free crystal structure of calpain-2 CBSW domain (blue) (PDB:1KFU).*[18]

The notion that the CBSW domain has the facilities to binds calcium ions at a 1:1 stoichiometry, or even higher in the presence of phospholipids, is presented in the literature, however the lack of $Ca^{2+}$ ions in the calcium bound crystal structure of calpain-2, despite saturated calcium binding sites elsewhere in the structure, but not CBSW, suggests this might not be the case.[35,41,42] It is however likely that this domain plays a pivotal role in the regulation of calpain activity through interface with the active site, dominated by electrostatic interactions (Figure 1.2.5).

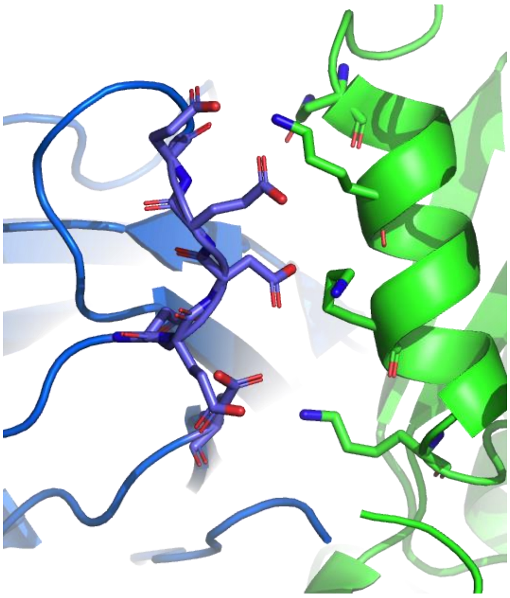


*Figure 1.2.5 - Secondary Structure of the negative CBSW loop (blue) between residues 390-405 and Glu504, and the interactions with CysPC α-helix (green) that is believed to maintain the full length complex in an inactive conformation (PDB:1KFU).*[18]

The active site cleft of calpain is believed to be held apart in an inactive conformation by the interaction of a negatively charged CBSW loop rich in aspartate and glutamate residues, with a positively charge α-helix on CysPC.[18,41] Removing the Glu504 residue by mutagenesis significantly lowered the calcium requirement for calpain activation, which was postulated to be due to a reduced electrostatic interaction between the CBSW and CysPC domains.[18,41]

### 1.2.4. PEF(L) and PEF(S)

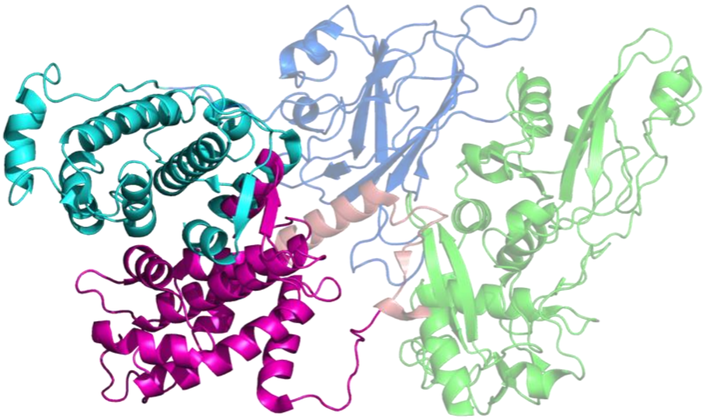


*Figure 1.2.6 - Secondary structure of PEF(L) (cyan) and PEF(S) (magenta) on the full calpain-2 structure (PDB:1KFU).*[18]

PEF(L) and PEF(S) are conserved domains, both containing five EF-hand moieties. [43] The L and S indicate the subunit each domain belongs to, with PEF(L) located at the C-terminus of the large subunit and PEF(S) is half of the small subunit.[43]

### 1.2.5. EF-hand moeity

The EF-hand motif is found in many calcium binding domains in nature, consisting of two perpendicular α-helices which are linked *via* a flexible loop that provides the residues required for binding $Ca^{2+}$.[44] The classical EF-hand is characterized by 12 residues with the pattern X•Y•Z•-Y•–X••-Z, where X, Y, Z, -X, -Y and –Z are the residues that provide the metal coordination site and circles represent the variable residues.[45] It is possible to bind magnesium in this site by forming a distinct octahedral geometry, whilst calcium ions generally bind in pentagonal bipyramid coordination geometry.[46] The binding occurs through negatively charged residues with carboxylate side chains such as glutamate and aspartate or *via* threonine and serine hydroxyl coordination, as well as carbonyls of the peptide chain in the loop.[44] The coordination sphere is then completed by water molecules. [45,47]

| EF Motif | X | • | Y | • | Z | • | -Y | • | -X | • | • | -Z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EFH-1 | L | **A** | G | D | **D** | M | **E** | V | S | A | T | **E** |
| EFH-2 | **D** | S | **D** | T | **T** | G | **K** | L | G | F | E | **E** |

*Table 1.1 - Sequence alignment of human PEF(S) canonical EFH1 with non canonical EFH2 from PEF(S) crystal structure (PDB:4PHJ).*[48]

Both canonical and non-canonical EF-hands are present in calpain-1 and -2.[20] The canonical EFH-2 uses five coordinating residues (Asp154, Asp 156, Thr158, Lys 160, Glu165) and a water to complete the coordination sphere.[48] EFH-1 is non-canonical requiring only four residues for coordination, (Ala111, Asp114, Glu116, Glu121) and an extra water replacing –X, resulting in an EFH of only 11 residues.[48]

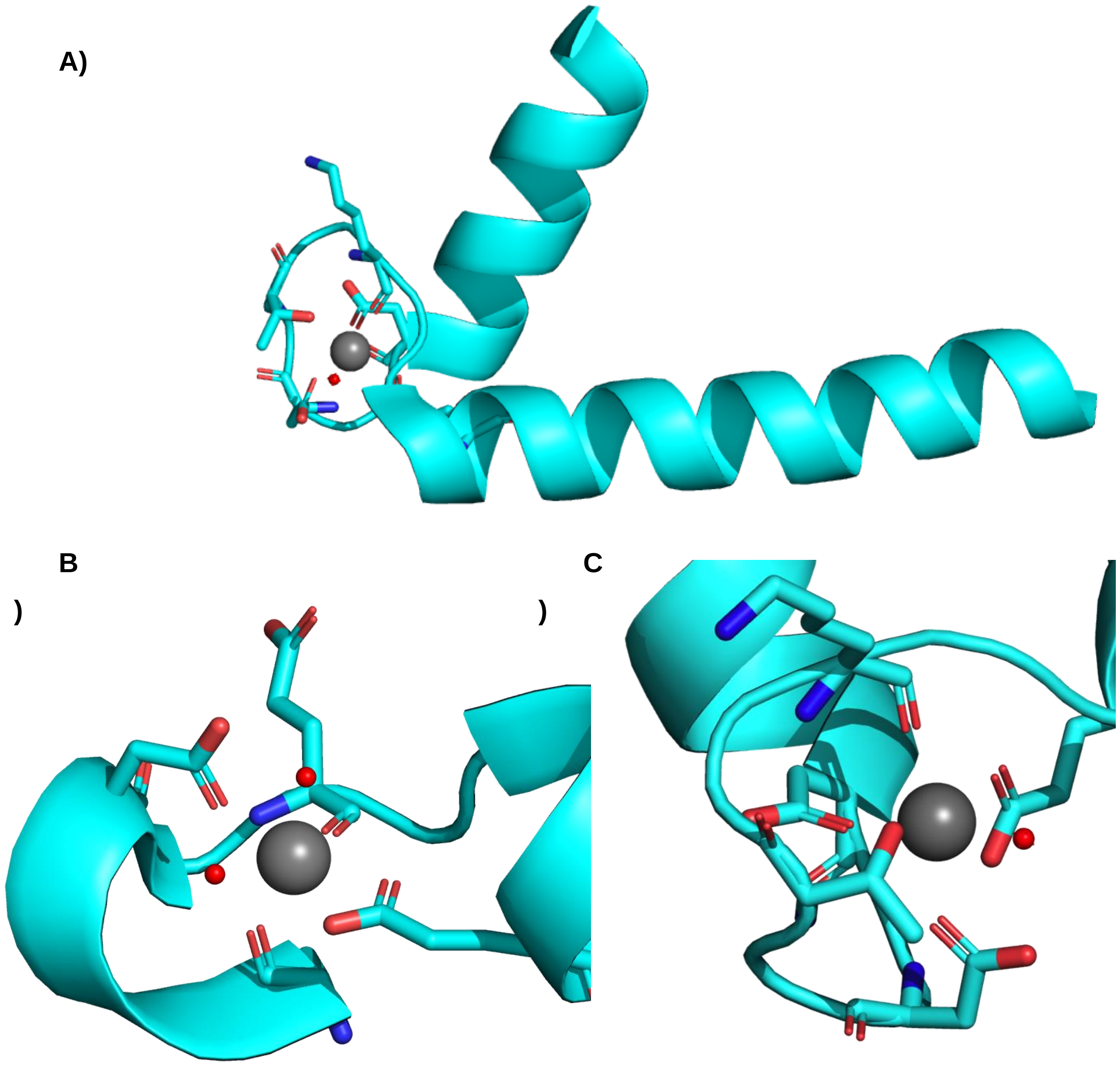


*Figure 1.2.7 -* ***A)*** *PEF(S) EFH-2 shown as cartoon with both α-helices and coordination environment* ***B)*** *Non-canonical human PEF(S) EFH2 coordination sphere around calcium ion* ***C)*** *EFH-1 of human PEF(S) showing canonical binding. Binding residues shown as sticks, waters shown as red spheres and Ca$^{2+}$ ions shown grey spheres (PDB:4PHJ).*[48]

### 1.2.6. PEF(L)

The calcium binding domain on the large subunit varies between calpain-1 and -2, the sequence similarity between PEF(L) major isoforms is only 42% therefore it is likely that variations in this domain are responsible for the differing calcium requirements of the isoforms.[49] Structurally, the domain contains eight α-helices forming a total of five EF-hands, binding two $Ca^{2+}$ ions in a canonical fashion *via* EFH-2 and EFH-3 while EFH-1 and EFH-4 bind two $Ca^{2+}$ non-canonically.[35] The only structure obtained for PEF(L) in a calcium bound form exists as a calpastatin complex, a structure for the activated form of the enzyme without calpastatin bound is still yet to be obtained.[35] Calcium addition or changes in salt concentration are known to cause subunit dissociation and protein aggregation which can complicate crystallography of the heterodimeric complex.[35,50]

### 1.2.7. EF-handshake

The fifth EFH forms an EF-handshake between PEF(L) and PEF(S), non-covalently joining together the large and small subunit of the heterodimeric calpains. This interaction mainly forms through hydrophobic interaction of aromatic phenylalanine and tryptophan residues.[51]

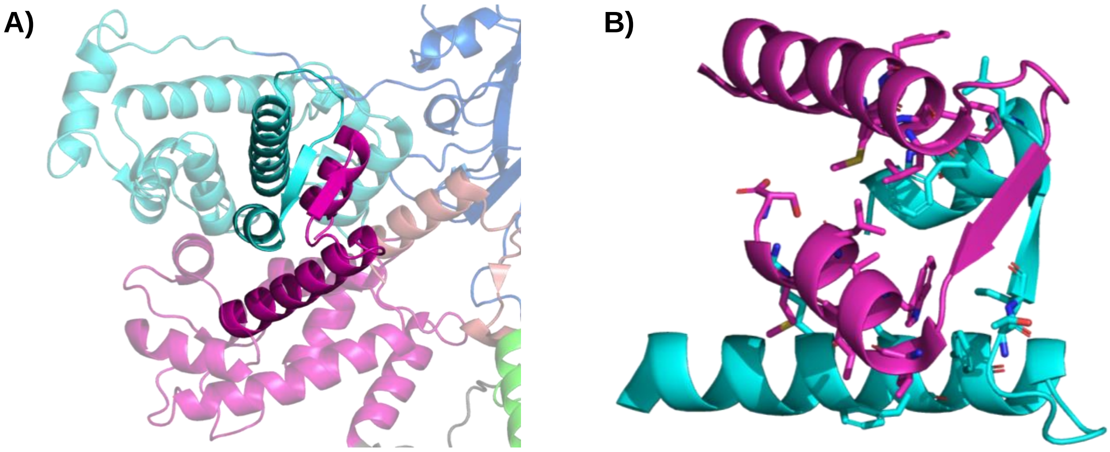


*Figure 1.2.8 -* ***A)*** *EF-handshake between the human large subunit PEF(L) domain (cyan) and small subunit PEF(S) domain (magenta) (PDB:1KFU)* ***B)*** *Zoomed secondary structure of the EFH5 interaction between PEF(L) and PEF(S), with the interface residues shown as sticks.*[18]

### 1.2.8. PEF(S)

The heterodimeric calpains -1 and -2 share an identical small subunit, also known as the regulatory subunit (CAPNS1), containing another calcium binding domain PEF(S) and a glycine rich domain.[20] PEF(S), like PEF(L) also contains five EF-hands, binding four calcium ions (two canonical, two non canonical) in an analogous fashion to PEF(L).[20] This subunit is believed to provide stability and aid calcium regulation of the large subunit, studies have shown that the small subunit can reversibly dissociate upon calcium binding which may play a part in the activation of the enzyme or be part of a signalling chain *in vivo*.[50]

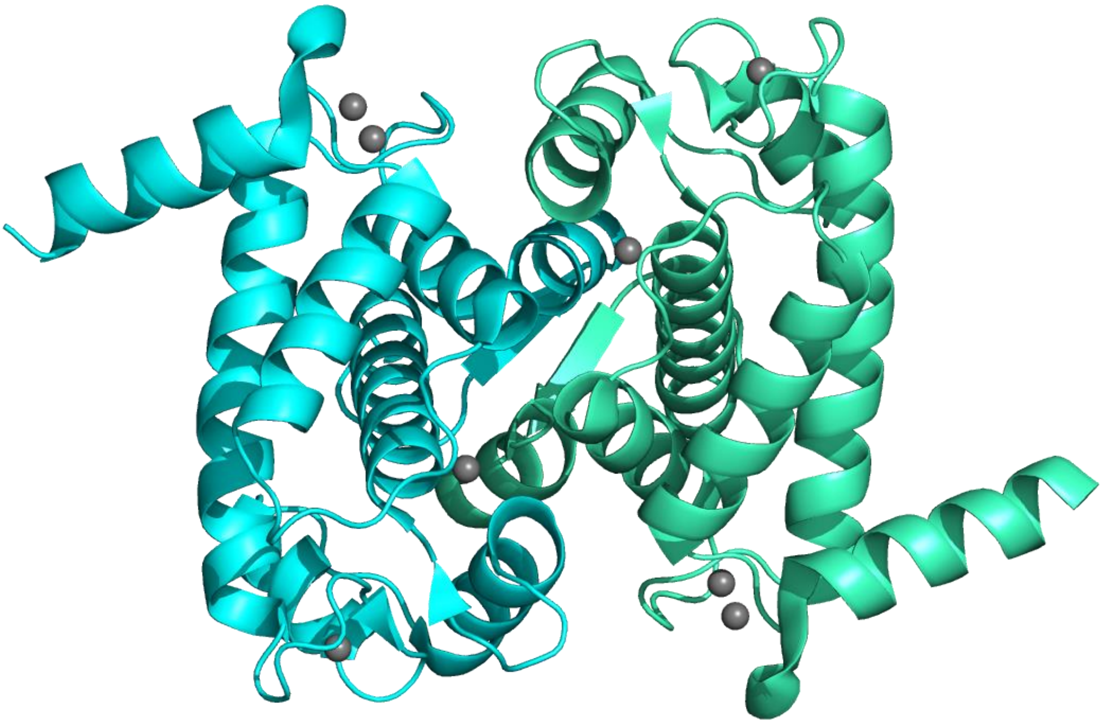


*Figure 1.2.9 - Human PEF(S) homodimeric crystal structure, chain A is shown as a cartoon in green/cyan and chain B in cyan, calcium ions highlighted as grey spheres (PDB:4PHJ).[48]*

There is only a small conformational difference between the $Ca^{2+}$ bound and free states of the small subunit, mutating the EFH residues to remove the calcium binding function in this subunit raises the calcium requirement for activation of the calpain-2.[52,53] It has been postulated that while the structural differences that occur upon calcium binding may be small, the dynamic transfer of these minute changes can be multplied to have a large overall effect on the functional properties of calpains.[52]

### 1.2.9. The glycine rich domain

The glycine rich (GR) domain contains 95 amino acids, within the first 65 residues are 40 glycines.[49] The flexbility of the glycine residue prevents this domain from being observed crystographically and as such is often truncated in recombinant constructs to improve protein expression as the GR domain was suspected to interfere with the host's translational machinery.[54] It has been suggested that this flexible hydrophobic domain, in conjunction with a rigid five proline (PEPPPP) portion, can bind phospholipids and may serve as a way of guiding calpain to the phospholipid membrane of the cell, where it is known to localise.[14,42,55]

### 1.2.10. Activation of calpain

Upon calcium binding the electrostatic switch between CBSW domain and CysPC is released, causing a large conformational change within the active site cleft allowing the residues situated on PC1 and PC2 to manoeuvre into sufficient proximity for catalysis.[18] PEF(L) and PEF(S) binding eight $Ca^{2+}$ ions in total, however these domains are postulated to play more of a regulatory role in the calpain system rather than responsible for activation as CysPC is a calcium dependent and active protease without both of these domains present.[36,56]

The resting cytoplasmic free $Ca^{2+}$ level in a cell is typically 100 nM, several orders of magnitude smaller than that required to activate calpain-1 and calpain-2 (50 *µ*M and 350 µM respectively).[57,58,59] Intracellular organelles such as the endoplasmic reticulum (ER) and the sarcoplasmic reticulum can accumulate $Ca^{2+}$ to concentrations of up to 100-500 µM.[18,55,60] Calpain is also thought to be activated in isolated parts of the cell for a specific function e.g. the cell membrane.[61,62] During $Ca^{2+}$ influx into a neutrophil, a peak concentration of 30 µM can be measured at the cell membrane often only extending 0.1 µm into the cell, this concentration is still not enough to activate either isoform of calpain, suggesting other factors are responsible for membrane localised calpain activation.[61,63]

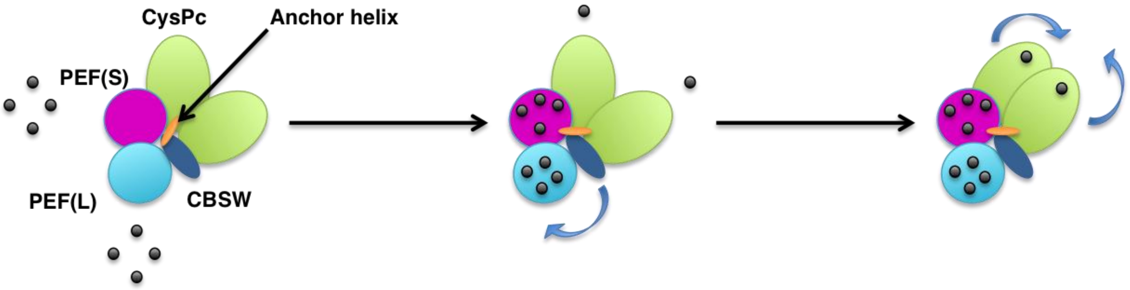


*Figure 1.2.10 - Schematic representation of the two-step mechanism for calpain activation, first binding eight Ca$^{2+}$ ions to the PEF(L) and PEF(S) domains, disrupting the electrostatic switch that the CBSW domain uses to hold PC1 and PC2 apart, followed by calcium binding to CysPc enabling the full activity of the enzyme.*[47]

The binding of an activator protein, Acetyl-coA binding protein, lowers the calcium requirement of calpain-2 from 350 µM to 10 µM, although the exact mechanism and purpose is yet to be determined *in vivo*.[64] The binding of phospholipids to CBSW and GR domains decreases the Ca$^{2+}$ requirement for calpain activation, such an effect would enable the restricted activation at certain locations, where calcium concentrations might be raised above free cytosolic levels e.g. the cell membrane.[55] Mutations to the acidic loop and 'phospholipid binding site' on the CBSW domain have been shown to reduce the calcium requirement but not to affect the lipid affinity, therefore suggesting the GR domain is likely more responsible for membrane localisation than CBSW. [14,65]

Auto proteolysis of calpain-2 from 80 kDa to two different active albeit less stable fragments 78 kDa and 75 kDa by cleavage between Ser15-Ala16 bond and Gly27-Leu28 respectively, alters the calcium activation concentration.[66] These forms have lower calcium requirements, but also short half lives triggering further auto-degradation proteolysis reactions which perhaps serves as a potential mechanism for limiting calpain activity *in vivo*.[32,34,67]

### 1.2.11. Substrate specificity

A range of proteins have been identified as calpain substrates including cytoskeletal and plasma membrane proteins, transciption factors and signal transduction proteins.[22]

Six human genes have been shown to be tissue specific, amongst these, CAPN3 is primarily present in skeletal muscle, CAPN6 in the embryonic muscles and the placenta, CAPN8 and CAPN9 in the gastrointestinal tract, CAPN11 in the testis and CAPN12 identified in the hair follicles.[68] Generally, mutations to tissue specific calpains can result in tissue specific diseases such as CAPN3 mutations resulting in limb-girdle muscular dystrophy.[53] The conventional calpains CAPN1 and CAPN2 are ubiquitously expressed throughout the body and consequently are involved in many vital celluar processes, such as the cleavage of structural membrane proteins ezrin and talin as mentioned, section 1.1.2 above, diseases are most commonly associated with dysregulation by either overactivation or inhibition.[1,62]

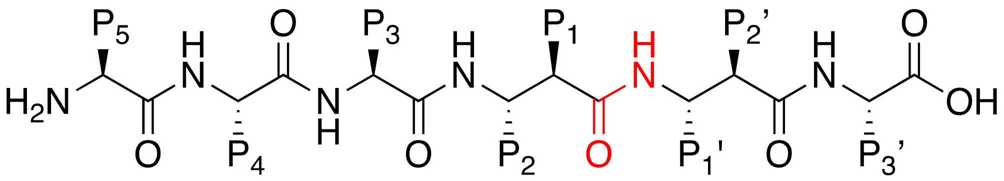


*Table 1.2 - The sequences of amino acids for optimum calpain-1 cleavage (site highlighted in red), the known substrates was from a library of 49 residue sequences,[69] the peptide library screening was from*

| **Substrate Type** | **P5** | **P4** | **P3** | **P2** | **P1** | **P1'** | **P2'** | **P3'** |
|---|---|---|---|---|---|---|---|---|
| **Known Substrates** | Gln>Lys >Arg | Gln>Ser >Pro | Lys>Trp | Leu>Thr >Val | Lys>Tyr >Arg | Ser | Pro | Pro / Lys |
| **Peptide Library** | Pro | Phe | Phe > Leu > Pro | Leu> Val | Leu=Phe | Met > Ala >Arg | Glu | Arg > Lys |

*synthetic peptide screening.[70]*

Substrate specificty is a key feature with most physiological enzymes, especially modulatory proteases and likely directs a siginificant role in the *in vivo* function. The scope of known protein substrate sequences were studied. The amino acids on the N-terminus of the cleavage site (P-site_ showing calpains prefer a much longer P-site than the C-terminus (P'site), and that around the scissile bond the P1 site shows preference for lysine, tyrosine or arginine residues the the P1' site for serine.[69] A slightly different substrate preference was observed when using a synthetic peptide library, for P1, hydrophobic leucine and phenylalanine residues and for P1' site a preference for methionine, alanine or arginine.[70] Proline residue features in the flanking regions P3, P2', and in the peptide library P5 and P3 also, this is thought to signify calpain's preference for cleaving intrinsically unstructured loop regions as proline is poor at forming α-helices and β-sheets and thus is commonly found in unstructured loops or as a 'helix breaker'.[71,72] It is likely that the observed specificity for cleaving unstructured regions in calpains is in part responsible for specific *in vivo* functions.[71]

# 1.3. Cellular functions involving calpain

## 1.3.1. Apoptosis

Apoptosis is an endogenous process whereby an external signal activates metabolic pathways that ultimately lead to programmed cell death.[73] Apoptosis is a vital mechanism for eliminating unwanted cells both during the development of a multicellular organism but also during homeostasis, especially in proliferating tissues such as the liver, prostate and adrenal cortex.[74] Calpain inhibition prevents apoptosis in motor neurons suggesting calpain activation is a requirement in apoptosis.[75] Cytoskeletal, membrane and structural proteins in the nuclear matrix such as lamins are also substrates likely degraded during apoptosis.[49,76]

## 1.3.2. Cell spreading and motility

Calpains are known to aid in the regulation of neutrophil motility *via* calcium regulated cleavage of structural membrane proteins (Section 1.1.2).[14] The communication of the intracellular cytoskeleton with the extracellular matrix (ECM) relies heavily on non covalent heterodimeric adhesion receptors known as integrins.[77] A large extracellular domain in integrins bind ECM ligands such as an endothelium whilst short cytoplasmic tails bind talin intracellularly.[77]

Dynamic adhesion and dissasembly of adhesion complexes is vital for directional movement, during this process cytoplasmic projections known as lamellipodia and filopodia extend into the extracellular environment.[55,62,78] Robles *et al.* observed $Ca^{2+}$ transients in the growth cone of filopodia cause localised activation of calpain, stabilising the filopodia, which they suggest to provide a spacial and temporal feedback mechanism from environmental cues through intracellular proteolysis initiated signalling cascades.[79]

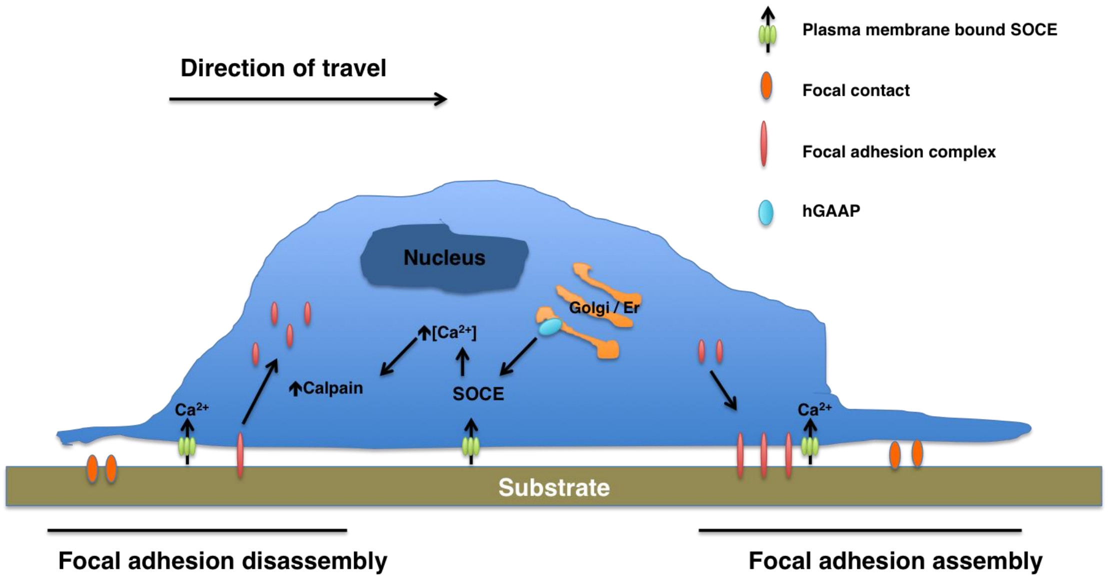


*Figure 1.3.1 - Human golgi antiapoptotic proteins (hGAAP) stimulation of ejection of $Ca^{2+}$ from ER/Golgi apparatus, opening store operated calcium entry (SOCE) channels in the plasma membrane, activating calpain near the plasma membrane resulting in turnover of focal adhesions.*[78,80]

Saraiva *et al.* postulate a mechanism whereby human golgi antiapoptotic proteins (GAAPs) stimulate a loss of $Ca^{2+}$ from the Golgi and ER, activating membrane bound store operated calcium entry (SOCE) resulting in a membrane localised increase in $Ca^{2+}$ thus increasing the turnover rate of focal adhesions.[80] This increased rate of adhesion and disassembly thus results in a more rapid cell spreading process and faster chemotactic migration.[80] The purpose of calpain in this process is to disassemble focal adhesion complexes which releases the rear of the cell, allowing the cell to progress in the direction of movement.[80]

### 1.3.3. Cell cycle progression

Calpain proteolysis is linked to several aspects of the cell cycle progression through $Ca^{2+}$ concentration transients.[81] Immunofluoresence techniques were used to monitor the relocation of calpain-2 around the cell, during interphase the protease was generally found at the membrane.[81] As mitosis progresses to the metaphase, calpain-2 relocates to the chromosomes then a perinuclear location in anaphase.[78] Injection of calpain-2 at the nucleus during interphase promoted metaphase and injection at late metaphase promoted destruction of the mitotic spindle and progression to anaphase.[81–83] It is evident that calpain plays a part in the complex disassembly of components required for the division of a cell.

## 1.4. Medical conditions linked to calpain

### 1.4.1. Cancer

Dysregulation of calpain activity can result in dramatic effects on various cancer types through reduced apoptosis, increased cell proliferation, stimulating cell migration and limiting invasiveness and thus has been highlighted as a potential anti-cancer target.[84] Overactivation of calpain-2 is involved in the migration and invasion of lung cancer cells through affecting focal adhesive properties of tumours as discussed (section 1.3.2), prostate cancer metastasis was blocked *in vitro* and *in vivo* by reducing calpain activity with inhibitors or stopping expression.[78,85,86] Treatment of cancer cells with calpain inhibitors could be a possible way of slowing the growth or mobility of a tumour, however it is stressed that such a treatment would only be used as a ‘maintenance therapy’ for prolonged survival of the patient, although none of these inhibitors are yet to be approved for therapeutic use.[84] Importantly, isoform specific inhibitors would accelerate the understanding required to tailor potential therapies to specific requirements while reducing the risk of off target effects to the patient.[84]

### 1.4.2. Alzheimer’s disease

Alzheimer’s disease (AD) is the most prevalent age-related neurodegenerative condition, representing between 60-80% of all dementia cases.[87] It is characterised by the extracellular deposition and aggregation of amyloid-β peptides that arise from proteolysis of the amyloid precursor protein. [84] Intracellular hyperphosphorylated tau neurofibrillary tangles and amyloin-β aggregates progressively and irreversibly restrict neuronal processes until a total loss of synaptic function is reached.[87,88] Overactivated calpain-1 and -2 proteases were discovered in plaques and neurofibrillary tangles in AD patients.[89] Neuronal degeneration was observed in AD mouse models through overexpression of calpain which is thought to promote amyloid-β plaque formation, inhibition of calpain was shown to alleviate the synaptic dysfunction associated of AD and thus is being explored further as a potential treatment.[90,91]

### 1.4.3. Ischemic cell death

Severe physical trauma such as cardiac arrest, stroke or injury causes a sharp increase of intracellular $Ca^{2+}$ levels resulting in rapid cell death.[92] A dysregulated rise in $Ca^{2+}$ overactivates calpain in the cell causing the uncontrolled cleavage of many substrates, ultimately resulting in apoptosis and necrotic cell death.[92] A cascade of signals exacerbates problems in cerebral ischemia whereby excess activation of glutamate receptors facilitates continual uptake of $Na^+$ and $Ca^{2+}$ into the cell causing membrane depolarisation and subsequent opening of voltage gated $Ca^{2+}$ channels.[92] Calpain inhibitors have been shown to be protective to the brain during ischemia *in vitro,* although not approved for therapeutic use and as such, current methods rely on techniques such as lowering body temperature, administering $Ca^{2+}$ channel blockers and free radical scavengers, to treat ischemia but display limited efficacy.[92–94]

## 1.5. Calpain Inhibition

### 1.5.1. Calpastatin

The endogenous inhibitor is a penta-domain protein called calpastatin (CAST), potent and selective for calpain-1 and -2, its relative molecular mass is 76 kDa and is not cell permeable.[20] Calpastatin contains four inhibitory domains (I-IV), each containing three conserved regions that bind to PEF(L), CysPC, PEF(S) *via* subdomains A, B and C respectively and each calpastatin molecule is therefore able to bind up to four heterodimeric calpains.[95] The fifth domain is a non-inhibitory L-domain suggested to aid in the inhibitory efficiency of calpastatin by guiding the inhibitory region into the active site cleft.[95] High physiological concentrations of calpastatin in tissues have been measured between 0.035 and 4.4 μM, which may help prevent the pathological implications associated with the overactivation of intracellular calpains.[96,97]

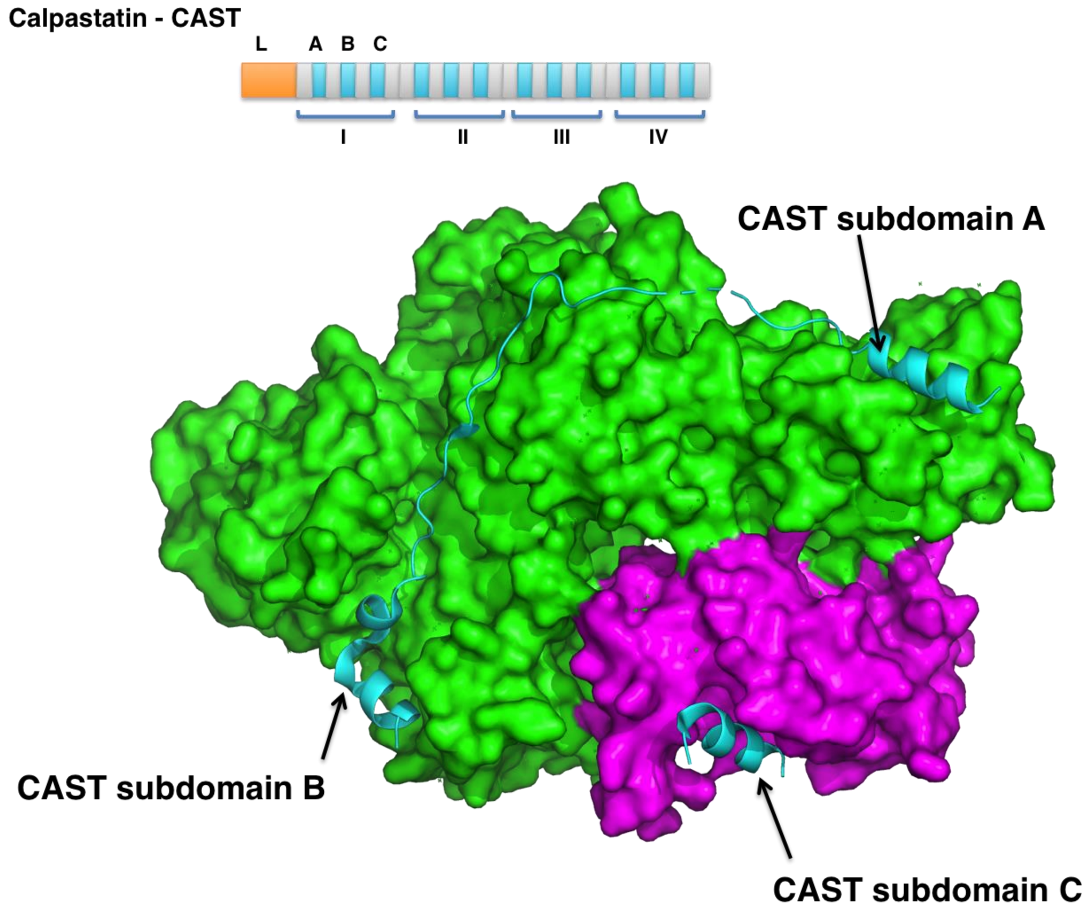


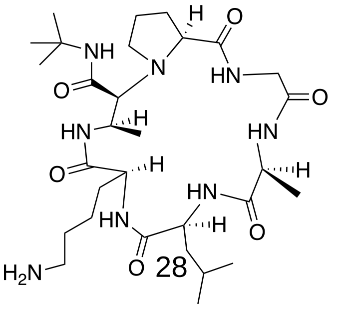
*Figure 1.5.1 – An inhibitory domain of calpastatin (cyan) binding to calpain-2 large subunit (green) and small subunit (magenta), via subdomains A on PEF(L), B over the active site cleft CysPC and C in a hydrophobic site on PEF(S). (PDB:3BOW)*[35]

Calpastatin is an intrinsically unstructured protein, displaying random coil secondary structure and three small α-helical subdomains, these subdomains have been shown to only bind to the $Ca^{2+}$-bound form of calpain, rather than triggering structural changes upon binding.[35] The data from this structure suggests that the subdomains A and C serve to anchor subdomain B which blocks the active site cleft preventing substrate binding.[35] The four inhibitory domains bind with different affinities to calpain ranging from $K_d$ values of 4.5 pM (domain I) to 4.0 nM (domain II).[98] The difference in these binding affinities *in vivo* is not known and to complicate matters the individual CAST domains can be liberated by active caspases and calpains, therefore binding studies are often carried out with the inactive serine mutant of calpain.[98]

Interestingly, when the helical portions of the subdomains for calpastatin are isolated as individual peptides, subdomain B maintained inhibitory properties.[96] Unexpectedly, subdomains A and C were significant activators of both calpain-1 and -2 by binding allosterically to PEF(L) and PEF(S).[96] A possible explanation for this is the susceptibility for calpastatin to be cleaved *in vivo* is downstream effects whereby the inhibitory region B but could be cleaved, leaving A and C intact to act as activators on calpain.[96]

### 1.5.2. Calpastatin based inhibitors

The structural and mechanistic data unveiled by the $Ca^{2+}$-bound calpain with calpastatin crystal structure has enabled the design of novel inhibitors for example the synthetic peptide, B27-WT, based upon calpastatin subdomain B (TYIEELGKREVTIPPKYRELLA), has a calpain-1 $IC_{50}$ value of 33 nM.[99] B27-WT was further developed by conjugating a membrane translocation peptide (VALLPAVLLALLAP) to the C-terminus of B27-WT to make a blood-brain barrier permeable compound, B27-HYD which was used to evaluate calpain contribution to neurological dysfunction in rats after a stroke.[99,100]



c*[PGALK]
Calpain-2
$K_i$ = 16.6 $\mu$M

*Figure 1.5.2 - Chemical structure of macrocyclic peptide inhibitor c*[PGALK].*[101]

Macrocyclic molecule c*[PGALK] was synthesised as a smaller more stable cyclic peptide based inhibitor although unselective between isoforms due to being based upon calpastatin which is not selective for calpain-1 or -2.[101] Cyclisation of peptides can decrease the entropy of binding through ligand geometry preorganisation, as well as display increased stability to proeotylic degredation.[101,102]

### 1.5.3. Covalent bond forming warhead inihbitors

Inhibitors have been developed incorporating reactive functional groups (known as warheads) such as epoxides, aldehydes and hemiacetals and a recognition sequence for the active site, providing the selectivity.[56] Warheads work by alkylating the cysteine residue of the active site, which can be disadvantageous as they can be reactive towards other cysteine proteases, resulting in undesired off-target effects.[103]

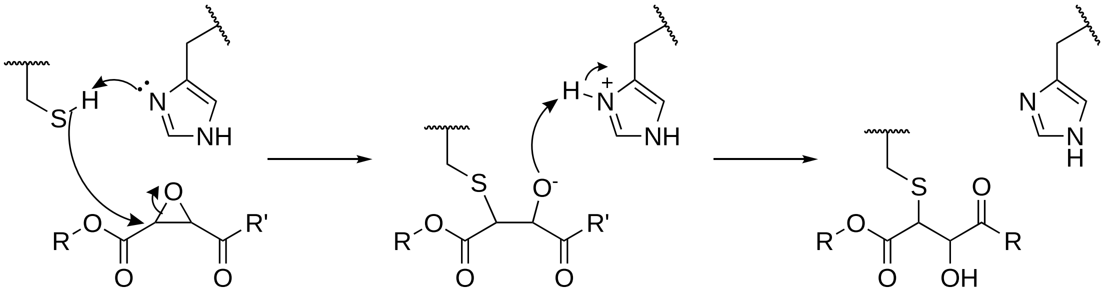


*Figure 1.5.3 – Cysteine protease epoxide warhead mechanism.*[40]

An epoxide based natural product, E64, from the fungus *Aspergillus japonicus*, was discovered to bind and block the active site of calpain, it displayed no reactivity towards enzymes of the aspartate or serine protease superfamilies, however it does inhibit other cysteine proteases.[104]

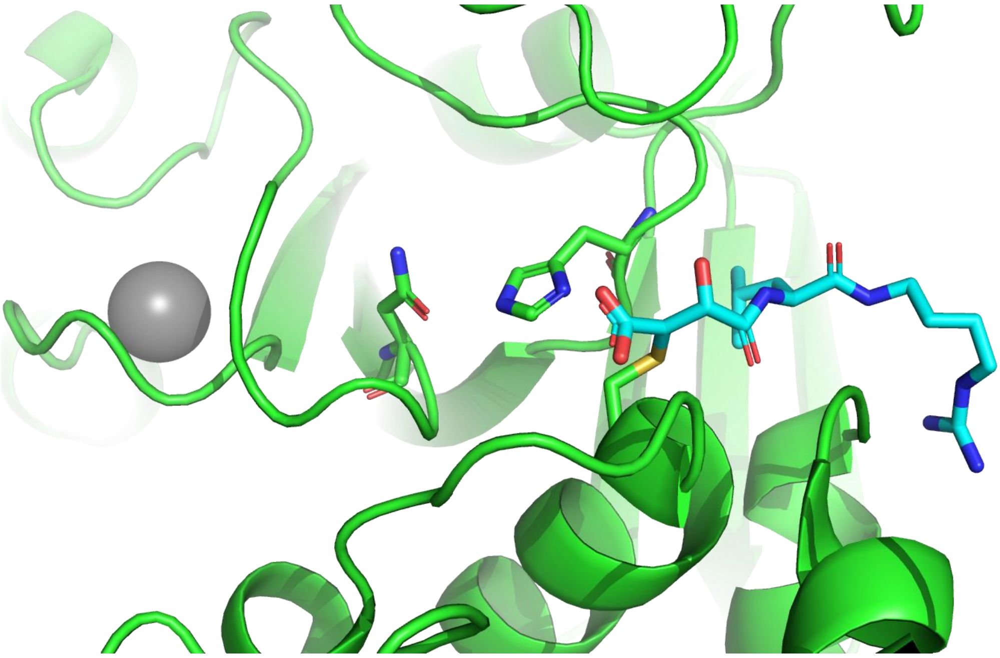
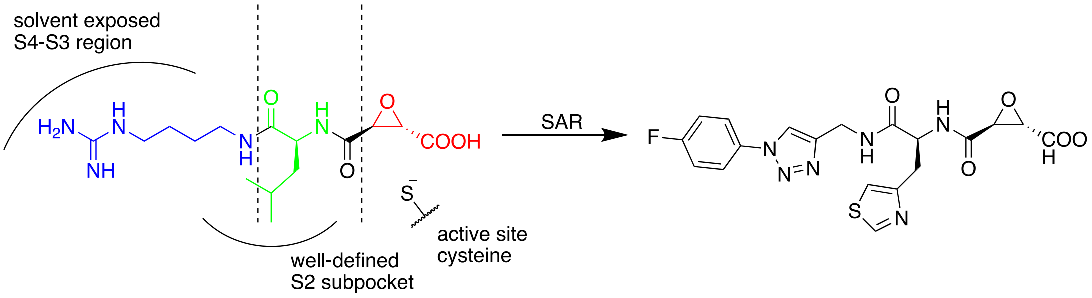


*Figure 1.5.4 - Calpain-1 active site secondary structure with catalytic residues represented as sticks (green), C115 is bound to warhead inhibitor E64 (cyan) (PDB:1TLO).*[104]

E64 was shown by crystallography to inhibit the active site CysPC without other domains present, and was not selective between isoforms or for calpain over other cysteine proteases such as papain.[104]



E64 - **1**
$IC_{50}$
Calpain-1 = 3.96 μM
Papain = 5.08 μM

**2**
$K_i$
Calpain-1 = 2.89 μM
Papain = 19.2 μM

*Figure 1.5.5 - Epoxide based calpain inhibitor, E64 (**1**) and second generation more selective inhibitor (**2**) developed using SAR from the crystal structure of CysPC.*[104,105]

An an attempt to improve the selectivity of E64 **(1)** for calpain over papain was carried out by analysing the inhibitor binding environemnt in the protein-ligand co-crystal structure.[106] Structure activity relationship (SAR) studies can aid the synthesis of potent and selective inhibitors by maximing the specific and necessary interactions between a ligand and a target protein for the desired effect.[104] In CysPC, possible hydrophobic interactions around the binding site were recognised and replacing the flexible solvent exposed region with rigid aromatic moeties guided the synthesis of **2**.[106] The resulting compound displayed an increased selectivity for CysPC over papain, validating the proof of concept for designing compounds to bind as effectively as possible by increasing the interactions to the surface environment surrounding the active site.[106]

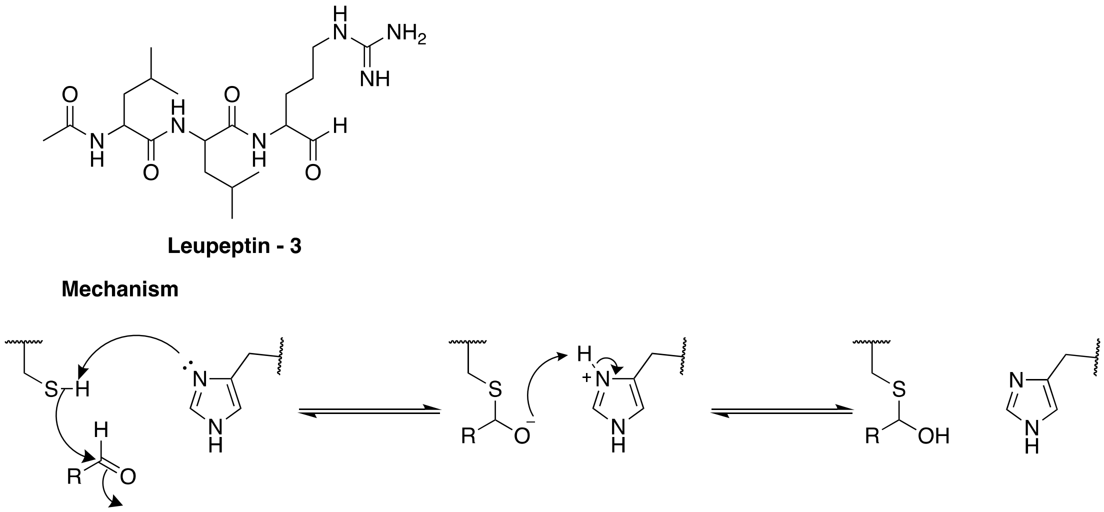


*Figure 1.5.6 - Leupeptin structure (**3**) and the mechanism of cysteine protease inhibition by a reactive aldehyde containing compound, forming a hemithioacetal moeity rendering the cysteine thiolate inactive.*[104,107]

Leupeptin (**3**) is a short acyltripeptide aldehyde first isolated in 1969 from *Streptomyces exfoliatus* and is a potent calpain-1 and -2 inhibitor ($IC_{50}$ = 0.27 and 0.38 $\mu$M, respectively).[105] It inhibits proteases by undergoing nucleophilic attack from an active site cysteine or serine residue on the reactive aldehyde moiety, forming a hemithioacetal or hemiacetal respectively, however the broad method of action means leupeptin also inhibits cathepsins A, B and D ($IC_{50}$ = 4 nM, 1 $\mu$M and 0.26 nM respectively) as well as the serine protease trypsin.[108,107] **3** was observed to bind to the calpain-1 active site by crystallography as a covalent adduct *via* the aforementioned hemithioacetal bond.[104]

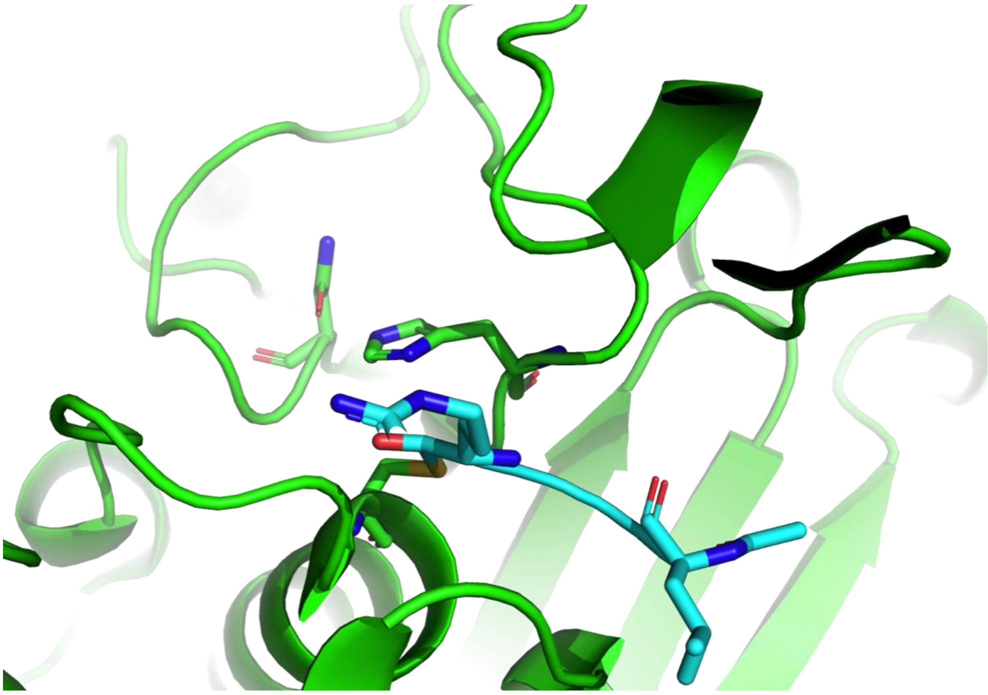
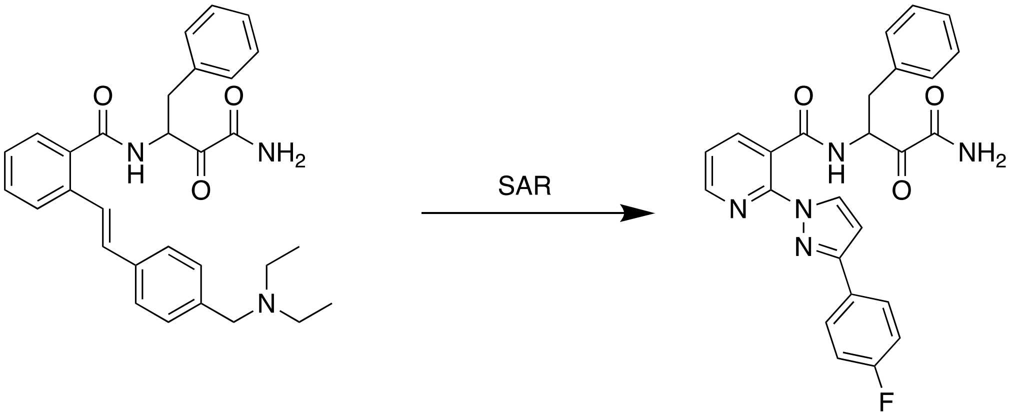


*Figure 1.5.7 - Calpain-1 active site (green) secondary structure with catalytic residues represented as sticks, bound to warhead inhibitor leupeptin (cyan) (PDB:1TL9).*[104]

Over the years many groups have attempted to produce selective, potent small molecule calpain inhibitors, however none have ever progressed far into clinical trials often prevented by poor cellular penetration, metabolic stability, lack of specifictiy and insufficient bioavailability.[1,90]



**4**
$K_i$
Calpain-1 = 56 nM
Cathepsin B = 28 nM
Cathepsin K = 1.8 nM
Cathepsin L = 137 nM
Cathepsin S 3290 nM

**5**
$K_i$
Calpain-1 = 38 nM
Selectivity Ratio (CAPN1/Cat)
Cathepsin B = 10
Cathepsin K = 18
Cathepsin L = 127
Cathepsin S = <300

*Figure 1.5.8 - Novel ketoamide based 2-(3-phenyl-1H-pyrazol-1-yl)nicotinamides as potent reversible selective calpain inhibitors, further improved by SAR.*[90]

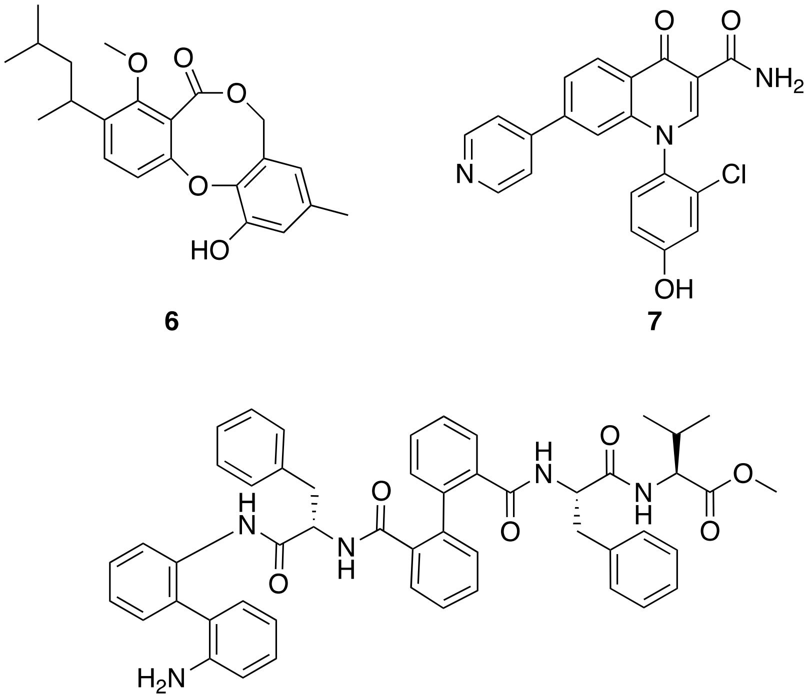
Compound **4** was identified as a efficacious calpain inhibitor to be used as a potential Alzheimers' disease treatment however was not selective against other cysteine proteases cathepsins B, K, L, S.[90] Immunosuppression can result as a side effect of inhibition of cathepsin L and S, however moderate cathepsin B and K inhibition was considered tolerable in the selectivity profile of this study.[40,90,109] SAR was used to alter the phenyl core of **4** for a nicotinamide P2 core and place five-member heterocylces in the P3-P2 position resulting in a more potent and far more selective, reversible drug.[90]

A more specific activity profile was generated for the compounds by screening against cross reactivities with G protein-coupled receptors (GPCRs), ion channels and transporters, as well as testing for CNS permeability which provided a positive overall pharmacoglocal profile. [90] **5** was then preceded into rat and dog trials and shown to prevent amyloid-β induced synaptic dysfunction, enhancing neurotransmission measured by improved cognitive performance.[88,90]

## 1.5.4. Non-covalent inhibitors

The majority of calpain inhibitors reported have been based upon peptide chemistry targetting the active site, often nondiscriminatory between serine and cysteine proteases. Small molecules that bind allosterically to calpain, targetting = domains unique to calpain whilst displaying potent isoform specific inhibition would be more advantageous to use *in vivo*.[110] Some compounds that display these characteristics have been developed.

Penicillide (**6**) is a polyketide with an $IC_{50}$ for calpain-2 of 7.1 μM, it is a reversible inhibitor and does not display a competitive mode of action therefore was suggested to bind allosterically.[111] 3-Quinolinecarboxamide derivative (**7**) displayed potent inhibition of calpain-1 ($IC_{50}$ of 0.5 μM) by blocking the active site in a similar fashion to aforementioned active site inhibitors.[112]



**6**

**7**

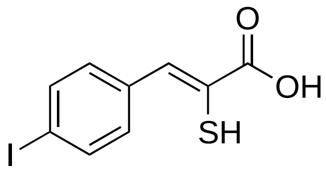
**8**

*Figure 1.5.9 - Non-covalent calpain inhibitors, penicilide (**6**) ($IC_{50}$ – 7.1 µM),[113] (**7**) 3-quinolinecarboxamide derivative ($IC_{50}$ = 0.6 µM)[112] and a potent peptide-biphenyl derivative **8** ($IC_{50}$ = 0.087 nM).[114]*

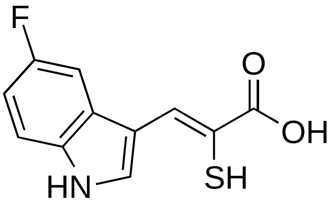
The most potent non-covalent inhibitor discovered is peptide-biphenyl hybrid compound (**8**) with an impressive $IC_{50}$ of 87 pM which was not reported to inhibit cathepsins or papain, however the mode of action is thought to be similar to EDTA, sequestering $Ca^{2+}$ ions tightly preventing the activation of calpain.[114–116]

## 1.5.5. α-Mercaptoacrylic acids

In 1996 a research programme found two compounds PD150606 (**9**) and PD151746 (**10**) that displayed modest selectively for the inhibition of calpain-1 over calpain-2.[117] PD150606 is a phenyl-based α-mercaptoacrylic acid substituted with iodine at the *para* position, PD151746 is an indole-based α-mercaptoacrylic acid substituted with a fluorine at the 5 position.[117]



PD150606 **- 9**
$K_i$ (μM)
Calpain-1 = 0.21
Calpain-2 = 0.37



PD151746 - **10**
$K_i$ (μM)
Calpain-1 = 0.26
Calpain-2 = 5.33

*Figure 1.5.10 - α-Mercaptoacrylic acid calpain inhibitors PD150606 (top) and PD151746, highlighting the difference in* $K_i$ *between the two major isoforms of caplain.*[117]

Analysis of co-crystal structures of **9** and **10** suggested that they bind allosterically to calpain-1 and -2 *via* the PEF(S) domain, in the same pocket as CAST region C.[37,117,118] Fluorescence displacement data suggests some indication of binding to PEF(L) to the displacement of a hydrophobic probe with the by **9**, which was the same for PEF(S).[117] The difference in $K_i$ values of the inhibitors on the two calpain isoforms also suggests that interaction with other domains is likely as PEF(S) is identical between the two major isoforms.[117]

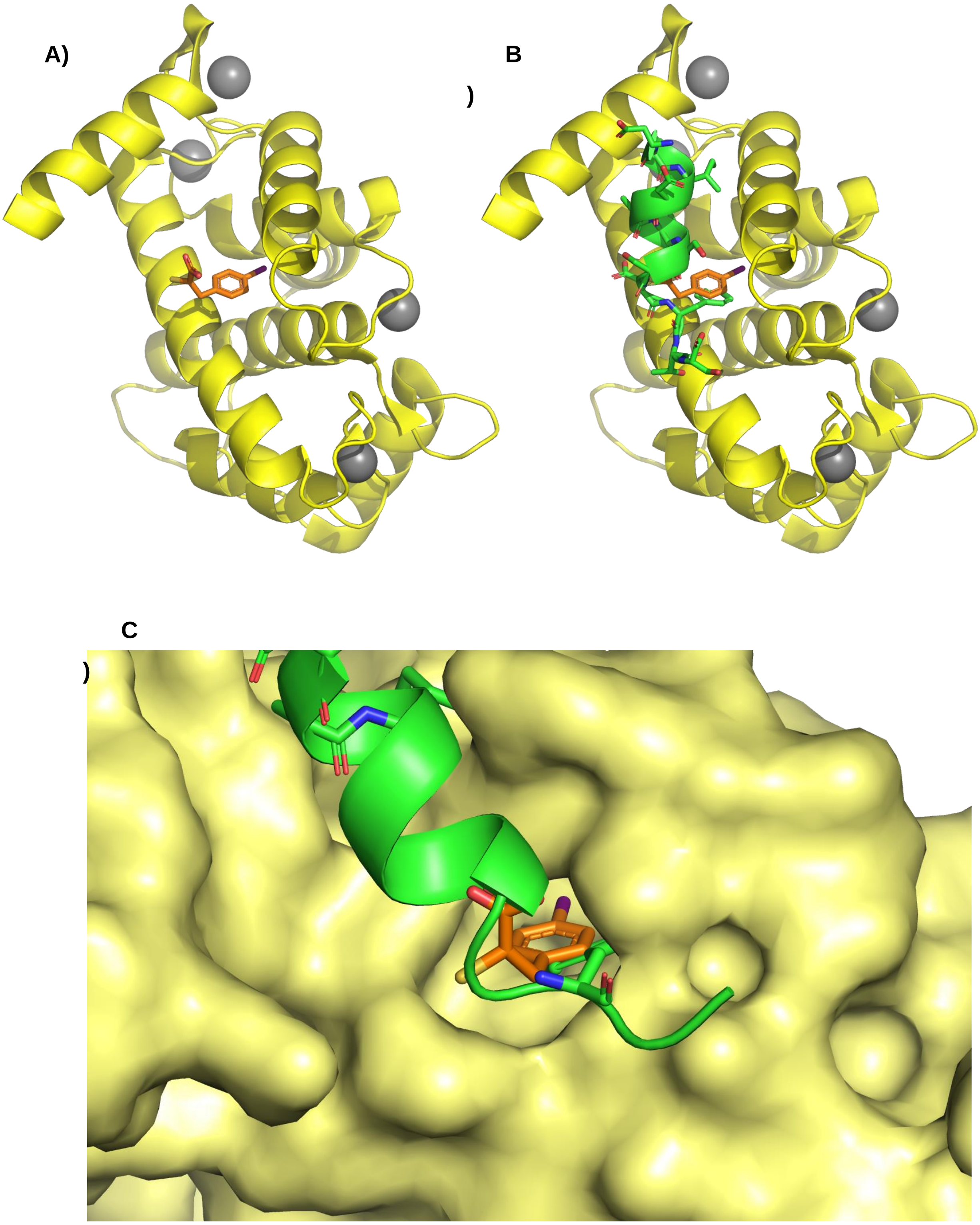


*Figure 1.5.11 – **A)** Secondary structure representation of PD150606 (orange sticks) bound to porcine PEF(S) (yellow), (PDB:1NX3), **B)** Calpastatin inhibitory region C (green) bound to porcine PEF(S) (PDB:1NX1), **C)** Both PD150606 and calpastatin inhibitory region C aligned in PyMol with PEF(S) surface representation (pale yellow).*[37]

The unique properties of these compounds encouraged further work in Cardiff to synthesise and evaluate a small library of analogues based upon the phenyl and indole α-mercaptoacrylic acids.[48] The sulfhydryl and disulfide forms were also tested and a dramatic increase in potency (up to 1000x) was determined for the oxidised disulfide form, over the reduced thiol.[63] The effect was observed by performing the assay both with and without thiol reducing agent dithiothreitol (DTT). Crystallographic data showed that the oxidised compounds were able to exhibit an increased interaction with the binding site on PEF(S).[48,63]

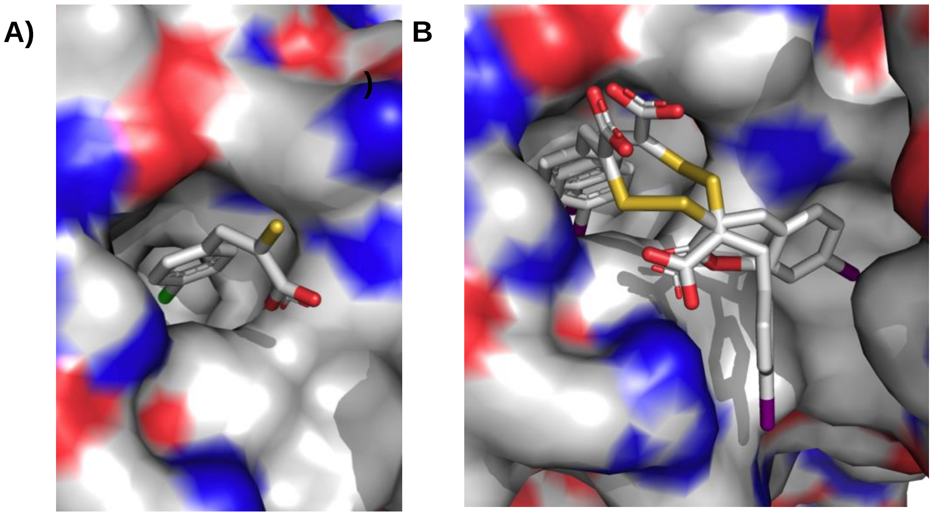


*Figure 1.5.12 - **A)** α-mercaptoacrylic acid bound in the human PEF(S) hydrophobic pocket (PDB:4PHK) **B)** Disulfide form of inhibitor shown to exhibit further binding to the PEF(S) hydrophobic pocket (PDB:5D69).*[48,63]

The second ring on the disulfide is able to interact with a second hydrophobic pocket on the surface of the protein whilst the unique geometry of the disulfide bond facilitates the bridging necessary for these two binding sites.[119] The geometry and length of the disulfide dihedral bond preoganises the ligand effectively for PEF(S). Preoganising pharmacophores reduces the entropic penalty when a drug binds to its target by reducing the conformational changes required for efficient binding, increasing the potency.[119] While the larger, oxidised compounds produced a much more potent effect *in vitro*, the reducing environment within cells would likely reduce the disulfide to the less potent sulfhydryl form when used *in vivo*.[63]

Aside from binding to PEF domains, far from the catalytic domain, other experimental data suggests **9** actually inhibits the calpain-1 protease core alone with no PEF domains present.[39] Mechanistic kinetics experiments were carried out with **9** and a penta-peptide LSEAL, which binds to the same groove in PEF(S), as a comparison.[39] **9** was shown to inhibit both full length calpain-2 and the protease core of calpain-1 *via* an apparent non-competitive model, whilst LSEAL did not inhibit, this more recent data suggests that while **9** does bind to PEF(S), the mode of inhibition is likely complex and may not completely arise from PEF(S) binding as originally thought.[39]

**9** was shown to be an inhibitor of matrix metalloproteinase-2 (MMP-2) ($IC_{50}$ = 9.3 $\mu$M), a member of the metzincin endopeptidase family that decomposes the extracellular matrix.[120] It was noted that the inhibitory properties exhibited by the α-mercaptoacrylic acid may be from an interaction with comparable calcium binding domains on MMP-2 and calpain, however was also noted that **9** readily binds zinc which is vital for MMP-2 activity therefore more likely to be the cause of inhibition.[6,121,122]

## 1.6. Aims and Objectives

Physiological roles of calpains remain poorly understood, in part due to the lack of isoform selective inhibitors. The mode of action of α-mercaptoacrylic acid inhibitors on calpain-1 is ambiguous, crystallographic evidence shows α-mercaptoacrylic acids bound to the calcium binding domain PEF(S) however studies reevaluating **9** kinetics show that it must inhibit by interaction on the active site, CysPC.[39,48] The dramatic increase in potency displayed by disulfides over the monomeric α-mercaptoacrylic acids was attributed to the unique geometry of the dihedral bond, enabling enhanced interaction with the protein.[63]

Studies highlight the stability of diselenide compounds in the reducing environment of the cell when substituting disulfide counterparts. For example, in the production of the insulin derivative, selenoinsulin, a comparable bioactivity to the native insulin was observed but with an 8-fold higher half life which was directly attributed to the ‘intrinsic stability’ of the diselenide bond.[123] A similar study exploring selenium analogues of glutathione reached a similar conclusion that was accredited to the unique chemistry and greater stability of diselenide bonds relative to disulfide bonds of the two chalcogens.[124]

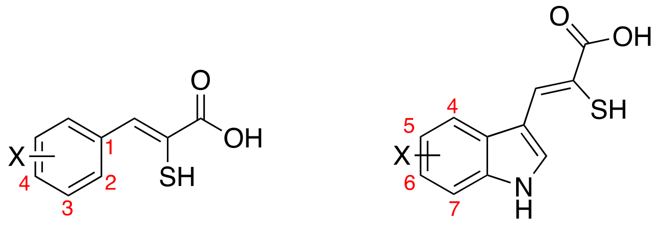
*Figure 1.6.1 - Typical disulfide and diselenide bond geometry around the dihedral angle.*[125–127]

The aim of this project was to explore the viability of diselenoacrylates as redox resistant allosteric calpain-1 inhibitors. These compounds were then to be explored by a variety of techniques such as X-ray crystallography, UV spectroscopy and enzyme kinetics experiments to determine whether they exhibit an allosteric mode of action and resistance to reduction. Promising compounds could then be used in cell imaging experiments to investigate cell permeability and physiological effects with respect to cell migration and spreading processes linked to calpain inhibition.

# Chapter 2 - Synthesis and Inhibitiory Properties of Diselenides

# 2.1. Synthesis of selenium analogues of α-mercaptoacrylic acid inhibitors

PD150606 (**9**) is a phenyl α-mercaptoacrylic acid derivative furnished with a single iodine substituent at the 4 position (Section 1.5.5), the disulfide form is readily formed in solution over several hours when exposed to oxygen.[128] When the α-mercaptocrylatic acids were used as calpain-1 inhibitors in the oxidised disulfide for, a dramatic increase in potency was observed.[128] This increase in potency was attributed to the dihedral geometry of the disulfide bond, allowing a more potent interaction with the PEF(S) hydrophobic binding site, however, under the reducing environment of the cell, was expected to be predominantly in the less potent reduced form.[128] To allow for direct comparison, the diselenide analogue of **9** was synthesised, as well as a selection of other aromatic analogues to enable a comparison between the compounds and aid in the charaterisation of these compounds as calpain-1 inhibitors. [117]



X = F, Cl, Br, I

*Figure 2.1.1 - IUPAC numbering for halogen subtituents around the phenyl and indole α-mercaptoacrylic acid aromatic rings.*

## 2.1.1. Synthesis of α-mercaptoacrylic acids

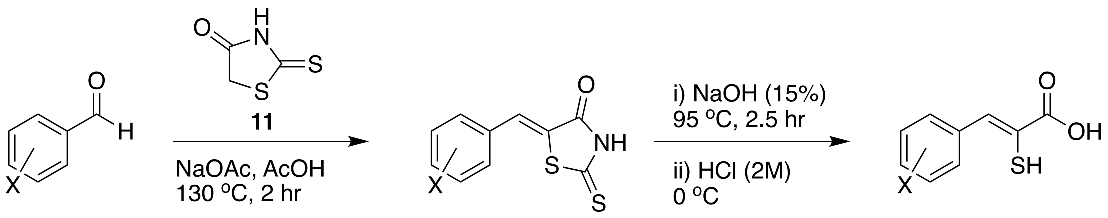
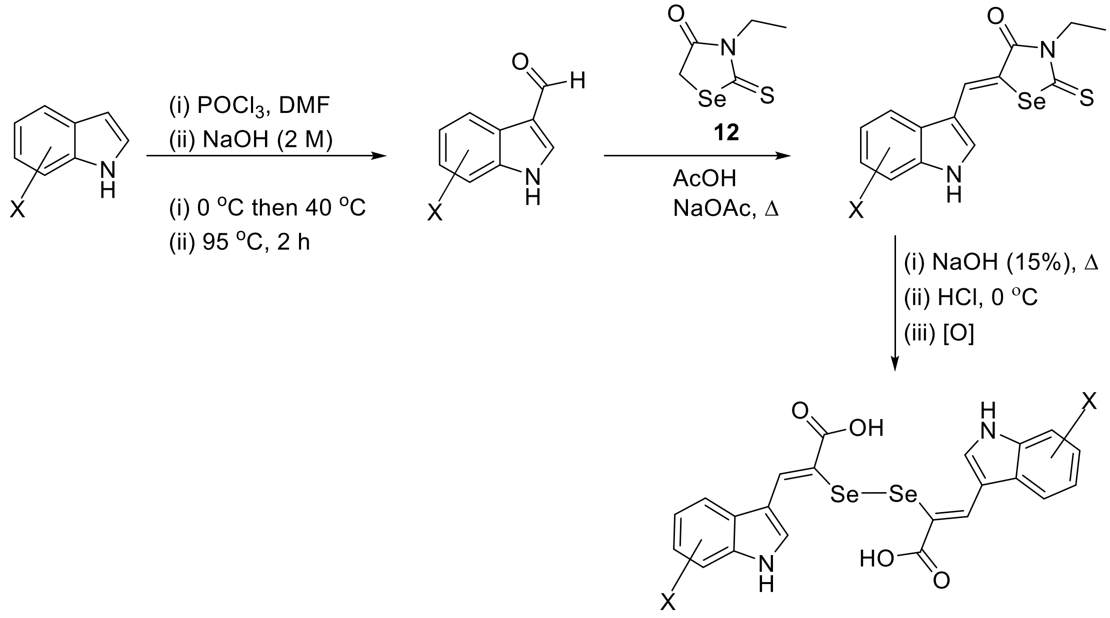


*Figure 2.1.2 - Synthetic route used in the production of α-mercaptoacrylic acids.[129]*

The synthetic route used to produce α-mercaptoacrylic acids proceeds *via* a two-step reaction from an aromatic aldehyde which reacts with rhodanine (**11**), followed by hydrolysis of the 2-thioxothiazolidin-4-one ring.[129] Rhodanines are five membered, sulfur/nitrogen heterocycles that are pharmacologically important scaffolds, possessing antibacterial, antifungal, antiviral and anti-inflammatory properties.[130] In this synthesis the rhodanine heterocycle is readily hydrolysed to yield the α-mercaptoacrylic acid moiety.

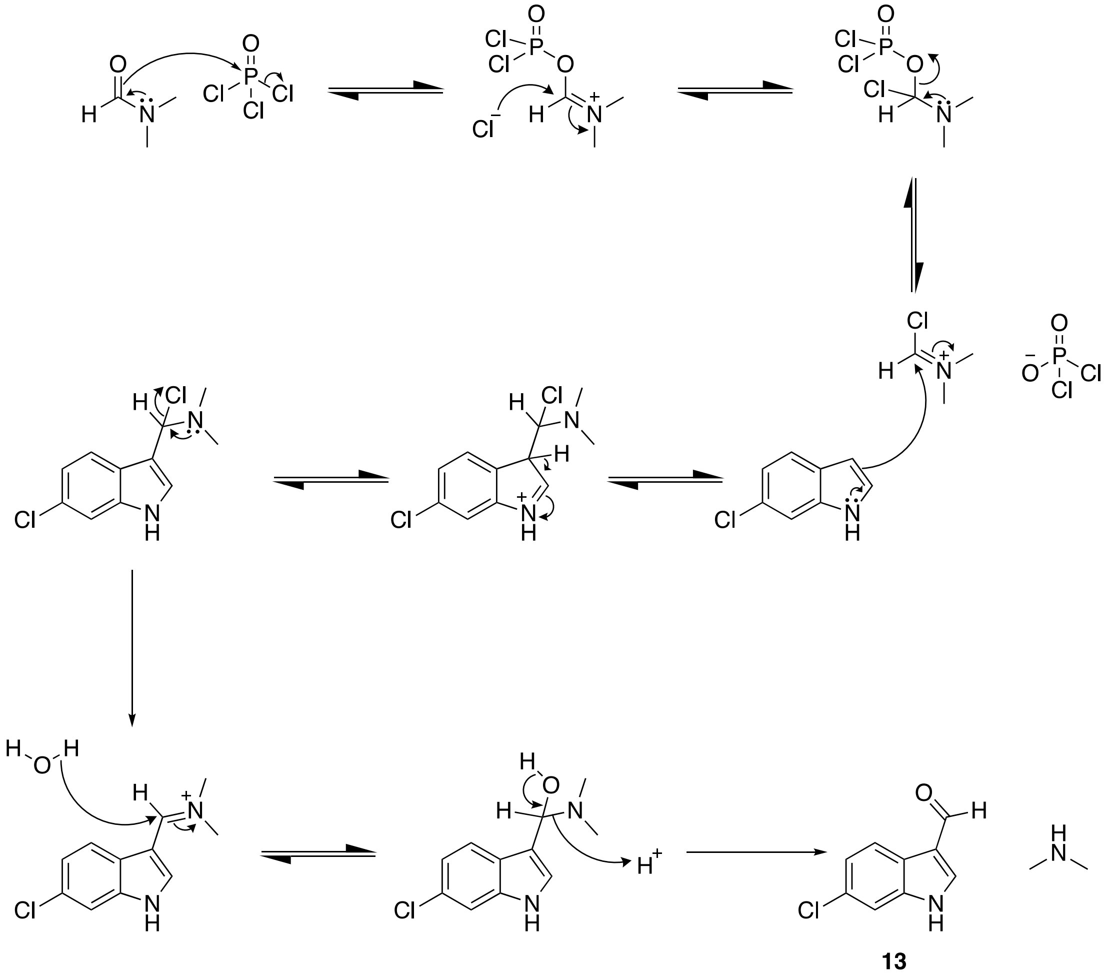
### 2.1.2. Proposed diselenide synthesis



*Scheme 2:1 – The proposed synthetic route for the diselenides via a Vilsmeier-Haack formylation, where the aldehyde was not readily commercially available, followed by Knoevenagel condensation with N-ethylselenorhodanine (**12**) and hydrolysis under basic conditions followed by oxidation with air.*

### 2.1.3. Preparation of aldehydes

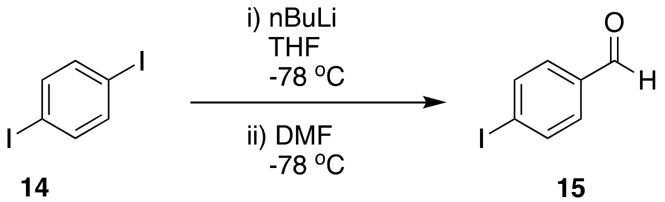
The first step of the synthesis required a substituted aromatic ring furnished with an aldehyde group. Most of these were commercially available, however for 6-chloroindole, the aldehyde was prepared *via* the Vilsmeier-Haack reaction.[131] *N*,*N*-Dimethyl formamide (DMF) reacts with phosphorus oxychloride ($POCl_3$) to produce the reactive *N*-(chloromethylene)-*N*-methylmethanimminium cation (Figure 2.1.2). Electrophilic substitution on the cation forms the α-chloro amine, which is then hydrolysed to yield the aldehyde. Upon cooling of the reaction mixture, the products readily precipitated and were isolated by filtration in high yield (89%) and purity after thorough washing with cold deionised water.



**13**

*Scheme 2:2 - The Vilsmeier-Haack reaction mechanism for forming aldehydes using DMF and $POCl_3$, using indole as an example.*[131]

The synthesis of 4-iodobenzaldehyde (**15**) was carried out using n-BuLi. The first step of this reaction is halogen-lithium exchange, followed by electrophilic quenching with DMF as a method of formylation. The hemiaminal intermediate is then hydrolysed during work-up to yield **15** as desired (Scheme 2.3).[132]

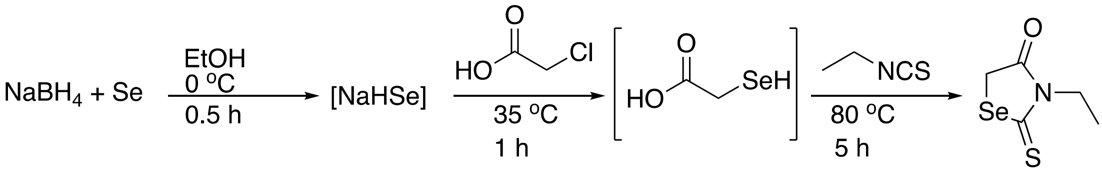


**14** **15**

*Scheme 2:3 - Synthesis of 4-iodobenzaldehyde (**14**) by halogen-lithium exchange with 1,4-diiodobenzene (**13**) followed by electrophilic quench with DMF.*

### 2.1.4. Synthesis of *N*-ethylselenorhodanine (**12**)

The synthesis of *N*-ethylselenorhodanine (**12**) was adapted from a synthesis of styryl dyes based on 3-ethyl-2-thio-selenazolidone derivatives.[133] This route is analagous to the α-mercaptoacrylic acid synthesis, however with a tertiary amine and a heterocyclic selenium atom in place of the sulfur. The reaction proceeds by forming sodium hydroselenide by reduction of selenium with sodium borohydride in thoroughly degassed ethanol, followed by the *in situ* formation of selenoglycolic acid *via* nucleophilic attack on chloroacetic acid with the hydroselenide anion. This can undergo reduction and subsequent cyclisation following the addition of isothiocyanate.[133]



**12**

*Scheme 2:4 – Proposed synthetic route to N-ethylselenorhodanine (**12**) from elemental selenium and sodium borohydride.*

The reaction was monitored by gas-chromatography mass-spectrometry (GC-MS) which separates volatile compounds by vaporising the sample and passing it through a chromatographic column in a carrier gas, usually helium.[134] The compounds are separated by their relative interaction with the stationary phase and the carrier gas, once the sample elutes from the column it then enters a MS source whereby the subsequent ionisation and analysis of the sample can provide an accurate mass and fragmentation pattern of the molecular ion.[134] This technique was used to monitor the formation of **12** by analysing a single drop of the reaction mixture in 1 mL of methanol and injecting onto the GC column.

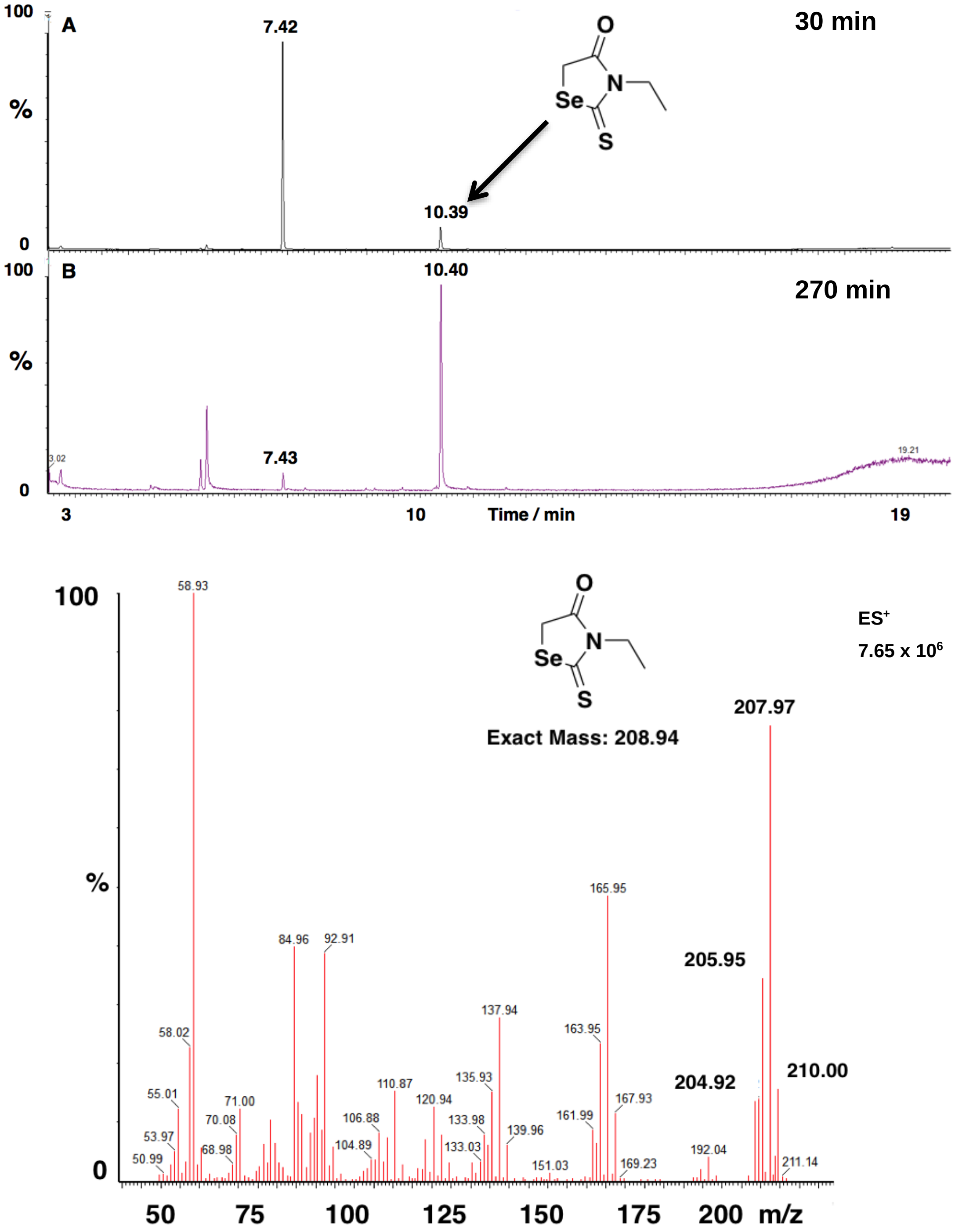


*Figure 2.1.3 – Top - N-ethylselenorhodanine (**12**) formation was monitored by gas chromatography-mass spectrometry (GC-MS) (**12** spectra) over the course of the final 5 hours with further additions of ethyl isothiocyanate added until no apparent increase of product was observed. Bottom – Mass spectra of product peak at 10.40.*

Strict anhydrous techniques were crucial for the success of the reaction, and workup of the mixture immediately upon cooling was important, quenched reactions left overnight yielded only traces of the desired product. Attempts at purifying **12** by flash chromatography resulted  in poor yields 7-19%, dark black precipiate formed on the silica as the column progressed, suggesting possible degradation of the product. Purification by vacuum distillation in Kugelrohr apparatus produced much more satisfactory results isolating the product at the reported temperature and pressure for distillation (120-125 °C/12 mm).[133]

### 2.1.5. Knoevenagel condensation

The next step was a Knoevenagel condensation which was carried out by heating **12** under reflux using ammonium acetate in acetic acid with halo substituted benzaldehydes, indole-3-carboxaldehydes or other aromatic aldehyde derivatives.[135] These reactions proceeded readily within 2 hours and upon cooling on ice, the product precipitated as a coloured solid which was filtered and washed extensively with ice cold deionised water.

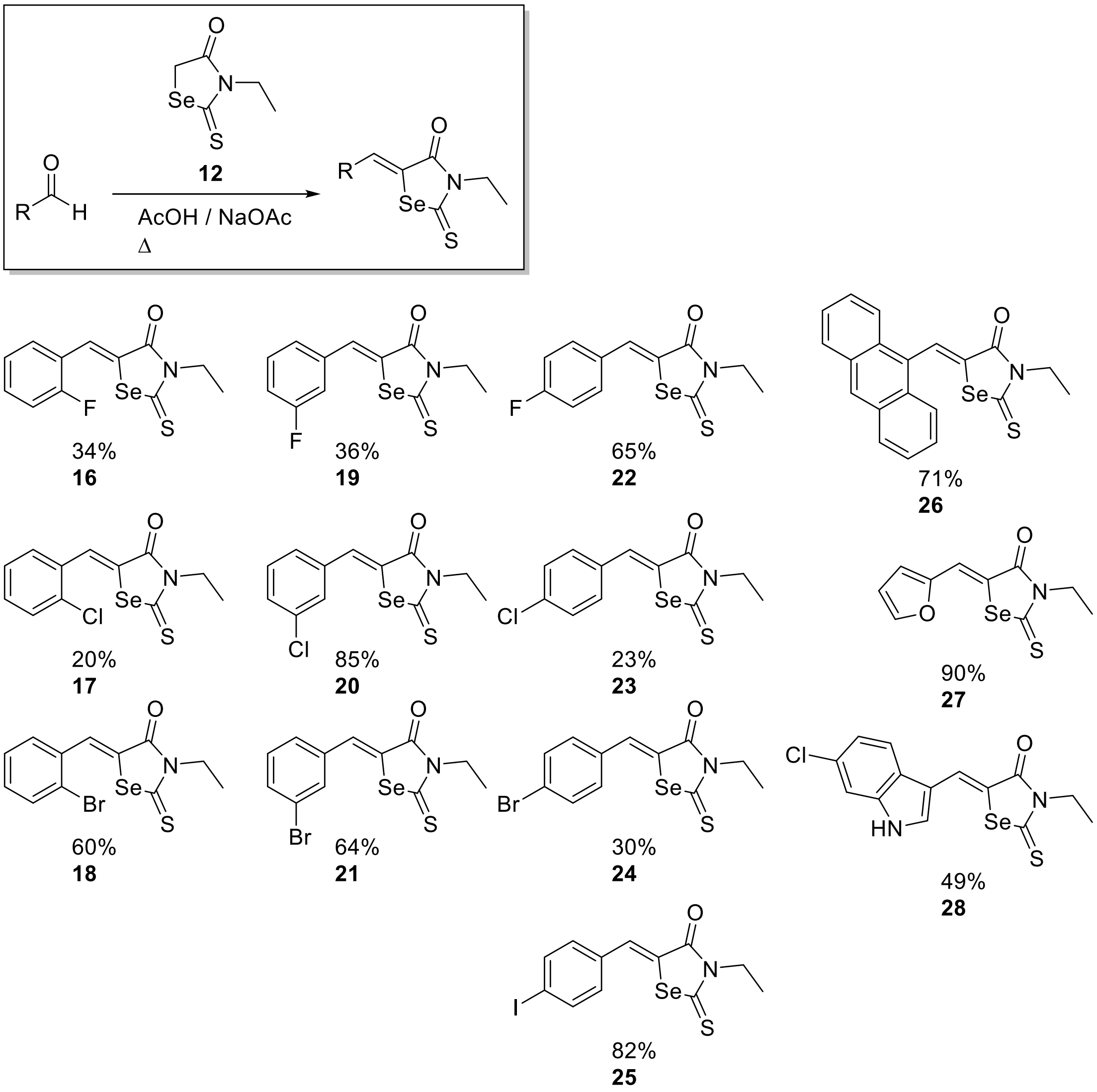
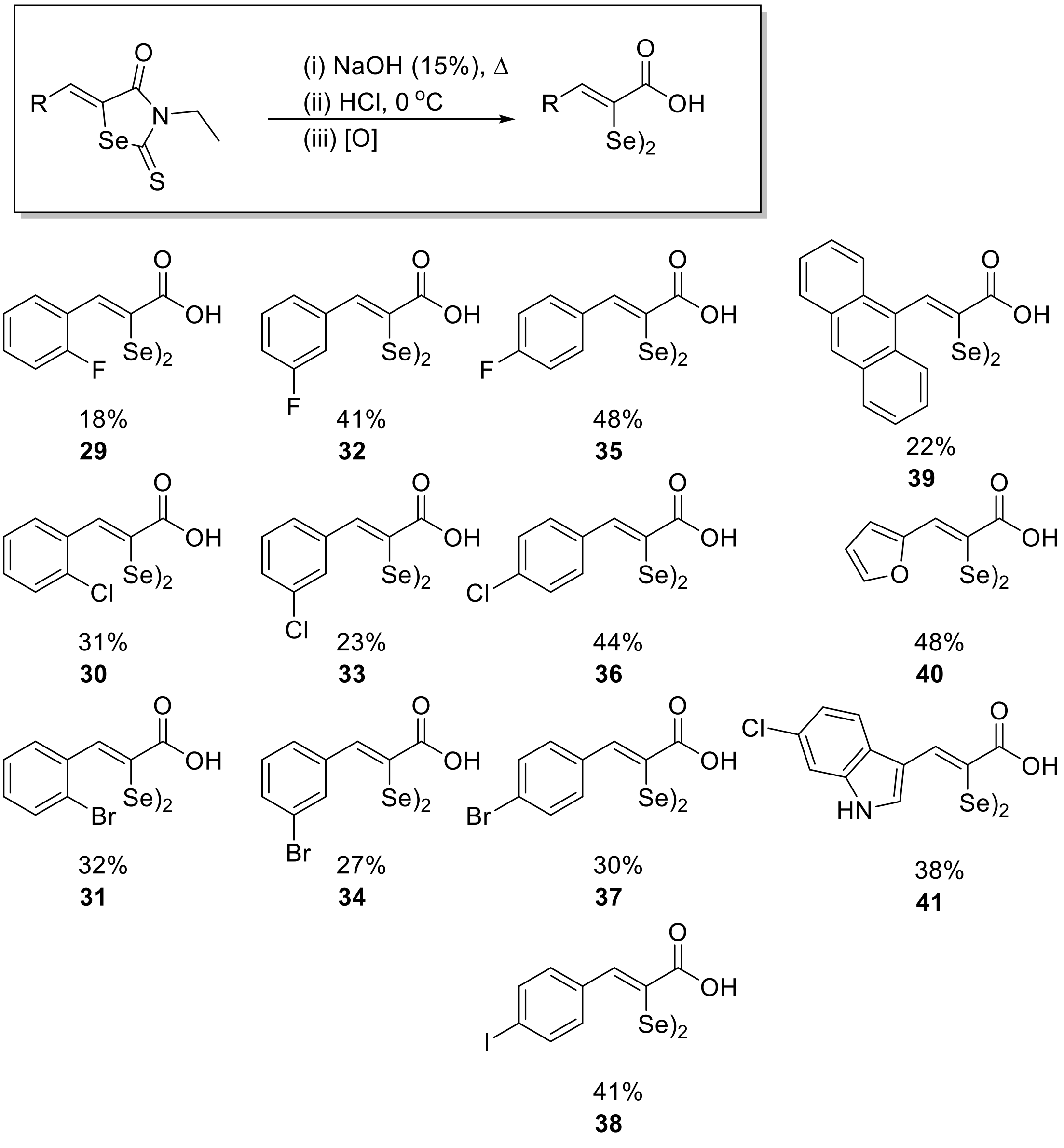


*Figure 2.1.4 - Yields of the products Knoevenagel condensation reactions of aromatic aldehydes and N-ethylselenorhodanine.*

The yields for **16-28** ranged from poor to excellent (20-85%), optimisation of reaction conditions was not carried out as the product isolated in each case was sufficient for the purposes of this project.

## 2.1.6. Hydrolysis and oxidation

The heterocyclic *N*-ethylselenorhodanine condensation products were hydrolysed by heating under reflux in $H_2O$ / methanol (2:1) with 15% NaOH for 1.5 h until all of the starting material was dissolved. The reaction was carried out in an oxygen atmosphere, providing the oxidant to form the final diselenide product. The resulting diselenides were isolated by acid-base extraction. For the less soluble indole derivative **41**, acidification of the reaction mixture and filtration of the precipitate with subsequent washes with ice cold deionised water and diethyl ether yielded the diselenide.



18%
**29**

41%
**32**

48%
**35**

22%
**39**

31%
**30**

23%
**33**

44%
**36**

48%
**40**

32%
**31**

27%
**34**

30%
**37**

38%
**41**

41%
**38**

*Figure 2.1.5. - Yields of products from sodium hydroxide hydrolysis and oxidation reactions of Knoevenagel condensation products.*

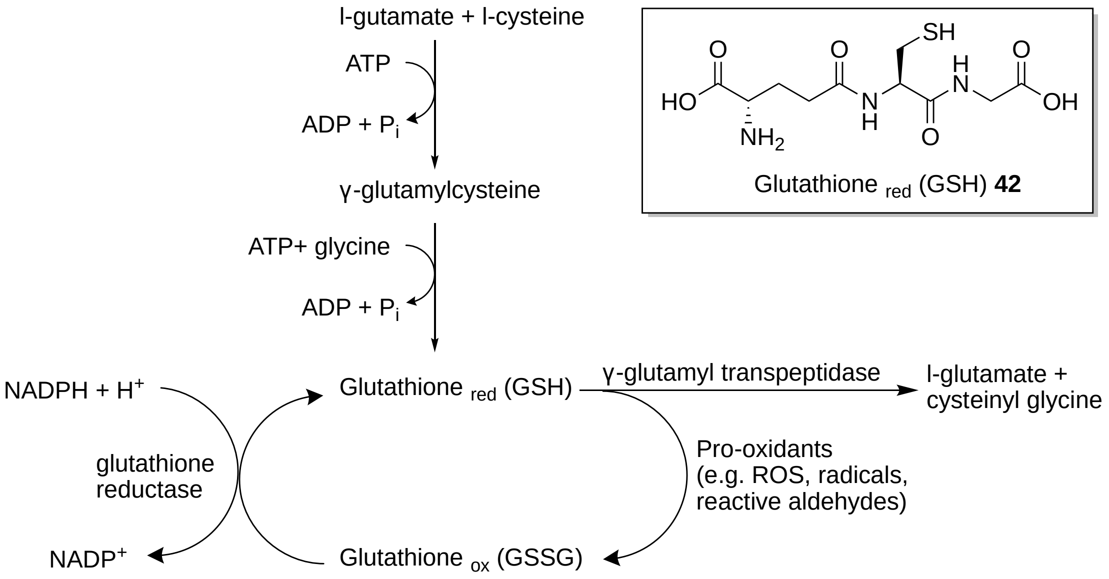
Mass spectrometry, $^1$H-NMR, $^{13}$C-NMR and IR spectroscopic analysis were used to identify the compounds **29** - **41** as the desired diselenides. The product formation was evident in the $^1$H-NMR spectra through loss of the characteristic *N*-ethyl group and the downfield C=S group in the $^{13}$C-NMR spectra. All products were isolated in 18-48% yields as the diselenide form of the product.

## 2.2. Testing for Diselenide Reduction

The geometry of the dihedral bond was identified as a key characteristic for allosteric inhibitors of calpain-1 *via* binding to the calcium binding domain, PEF(S).[48] In order for this to be maintained in the diselenide inhibitors, redox properties of the diselenides were explored by UV-Vis spectroscopy and mass spectrometry through changing the redox environment with common biological reducing agents in order to assess whether they would remain in the diselenide form within a living cell/organism.

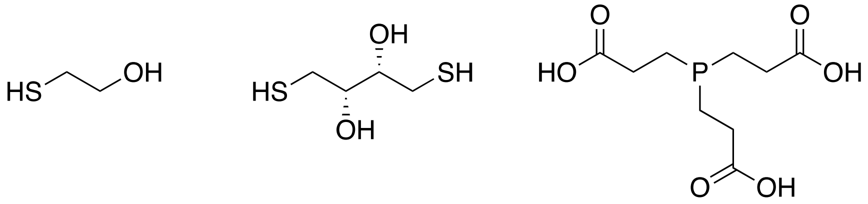
### 2.2.1. Biological reducing agents

Reducing agents are commonly used *in vitro* to simulate the cellular environment which can preserve cysteine residues in a reduced state required for enzymatic activity or prevent aggregation of proteins. This is of upmost importance in the case of cysteine proteases, where the activity is dependent on the ability of the active site thiol to implement nucleophilic attack on a substrate.[40] Thiols can undergo a variety of transformations such as oxidation to disulfides, sulfenic, sulfinic and sulfonic acids, binding metal ions, hydride transfer and radical chemistry.[136] In a cellular environment the most abundant non-protein thiol is glutathione (GSH, **42**), present in millimolar concentrations in the majority of cells.[137] It is synthesised from L-glutamate, L-cysteine and L-glycine by γ-glutamylcysteine ligase and glutathione synthetase.[138] GSH is responsible for protection from oxidative damage by reactive molecules such as free radicals and aldehydes, forming oxidised GSH disuflide (GSSG) in the process.[138] GSH can be regenerated from GSSG by glutathione reductase and NADPH, intracellular conditions of glutathione are under a constant dynamic equilibrium where a ratio of 9:1 GSH : GSSG is maintained for normal cell functioning.[139]



*Scheme 2:5 – Enzymatic synthesis, degredation and regeneration of the short peptide based reducing agent glutathione (**42)**, also highlighting how glutathione transitions between the oxidised and reduced state, protecting the cell from pro-oxidants.*[138]

In practice, small molecule reducing agents such as β-mercaptoethanol (βME, **43**), dithiothreitol (DTT, **44**) and tris(2-carboxyethyl)phosphine (TCEP, **45**) are often used in place of glutathione due to ease of handling, high solubility in water and cost.

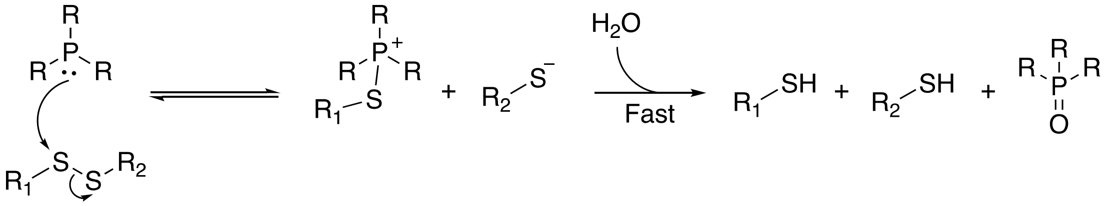


**43** **44** **45**

*Figure 2.2.1 – Frequently used reducing agents; β-mercaptoethanol **43**, dithiothreitol **44** and tris(2-carboxyethyl)phosphine **45**.*

βME (**43**) is a volatile liquid at room temperature with a reduction potential of -0.196 V and a thiol p$K_a$ of 9.61, and the reaction product of a thiol-disulfide interchange can become trapped when using a monothiol, which can be undesirable in sensitive enzymatic applications.[140] DTT (**44**) is a dimeric form of βME with a much lower reduction potential of -0.327 V due to the highly favourable formation of a 6-membered cyclic product upon oxidation.[126] Despite this, the lower thiol p$K_a$ value of

9.2 means DTT is a slower reductant where less than 1% of the thiol groups are in the reactive thiolate form at a neutral pH.[140] A more recently developed substitute for thiol based reducing agents is TCEP (**45**), which works by formation of a thiophosphonium salt, followed by rapid hydrolysis to form the thermodynamically favourable phosphine oxide and reduced thiol products.[141] The reduction potential of TCEP is -0.29 V therefore not as strong as DTT, however it is more stable in aqueous solutions, odorless, usable over a wider pH range and does not interfere with immobilized metal affinity chromatography (IMAC) for protein purification.[142–145]



*Scheme 2:6 - Mechanism of disulfide reduction by phosphines, proceeding via formation of a thiophosphonium salt and rapid hydrolysis to yield the reduced thiol products.[141]*

Isolation of organic selenols requires the strict exclusion of air and thus there is limited data available on these species, however diselenide bonds are regularly reported to be stable, absorbing light near 300 nm, resulting in a yellow solution.[146] Upon reduction, a decrease in this absorbance peak and a new UV absorbance peak between 240-250 nm has been reported to be the formation of the selenoate anion from diselenide reduction.[146] These observations confirmed the theory that DTT was in fact able to reduce the diselenide bond to the selenoate form.[146]

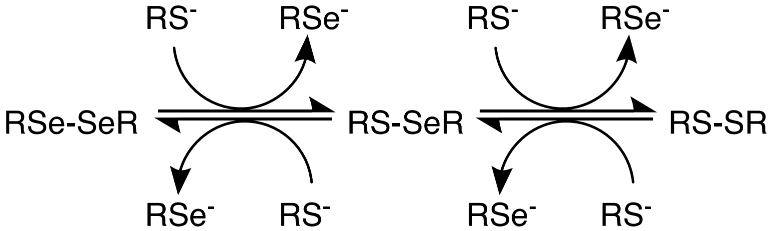
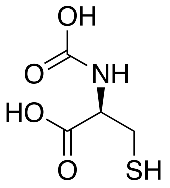
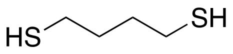


*Figure 2.2.2 - A diagram representing the dynamic exchange reaction between selenium and sulfur species.[147]*

A comparative study between selenium and sulfur for use in exchange reactions debated this and attributed the peak at 250 nm to the oxidised S-S bond, and a peak at 280 nm to the formation of a S-Se mixed species.[147] Selenoates have been also shown to form under reducing conditions using *N*-acetylcysteine (**46**) and 1,4-butanedithiol (**47**).[148]



**46**



**47**

*Figure 2.2.3 - Reducing agents N-acetylcysteine (**46**) and 1,4-butanedithiol (**47**).*

### 2.2.2. Reduction of **40** with β-ME and DTT

Titrations of common reducing agents were carried out on **40**, monitoring the UV absorbance between 200-450 nm in order to determine any response to change in redox conditions and whether these compounds displayed any resistance to reduction.

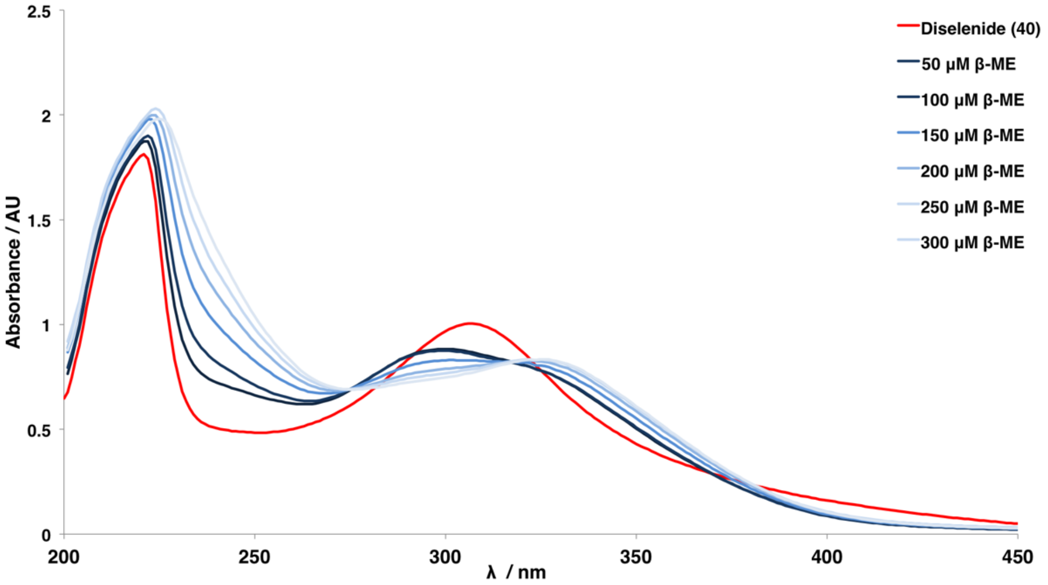


*Figure 2.2.4 – UV-vis spectra of diselenide (**40**) at various concentrations of β-ME. The peak at 306 nm was attributed to the diselenide peak and the peak at 250 nm to the selenoate anion.*

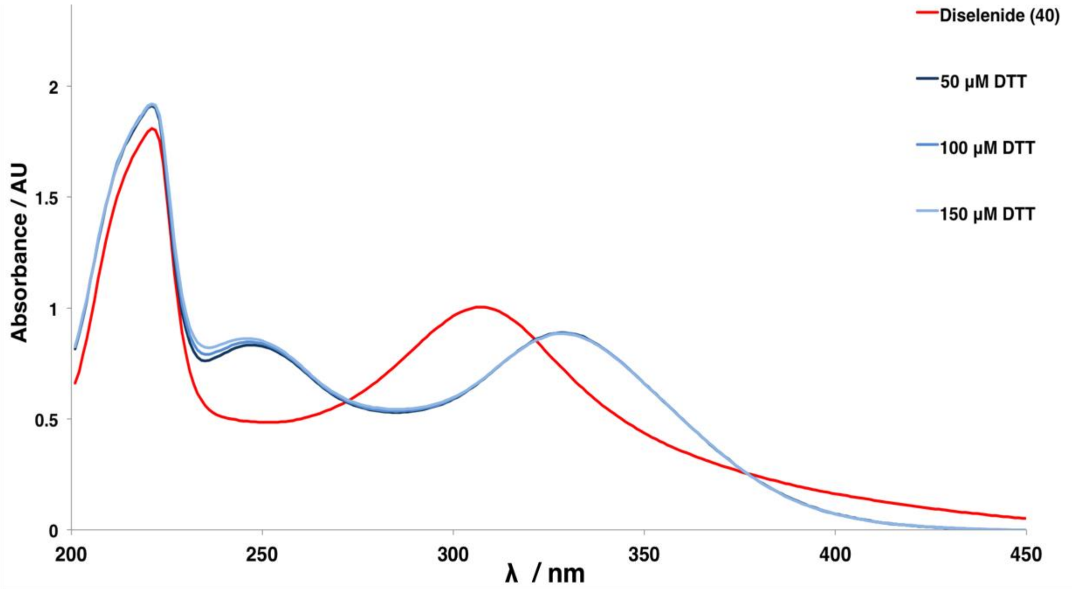


*Figure 2.2.5 - UV-vis spectra of diselenide (**40**) at various concentrations of DTT. The peak at 306 nm was attributed to the diselenide peak and the peak at 250 nm to the selenoate anion.*

Significant changes in the absorbance spectra were observed upon the addition of thiol based reducing agents βME and DTT. The peak attributed to the diselenide bond at 306 nm diminished and two new peaks became apparent at 328 and 247 nm. βME displayed a titration effect whereby each addition produced an incremental change, however after the first addition of DTT, no further significant changes were observed upon increasing the reductant concentration. This is likely due to the greater reduction potential of DTT over βME and the formation of the energetically favourable cyclic product for DTT. The literature suggests the new peak at 247 nm is likely to be the selenoate anion, indicative of diselenide reduction.[146]

### 2.2.3. Reduction of **40** with TCEP

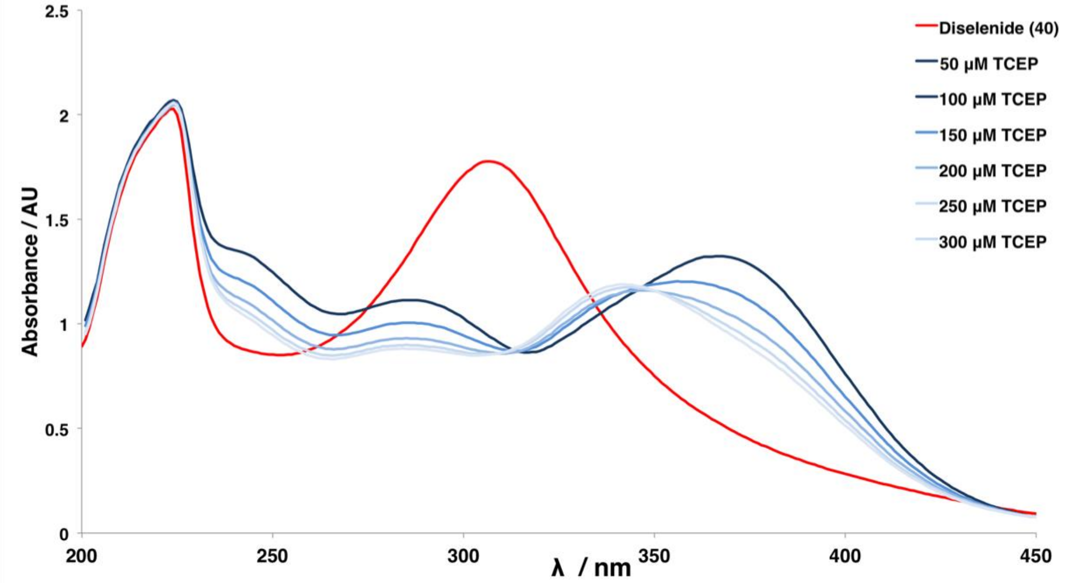
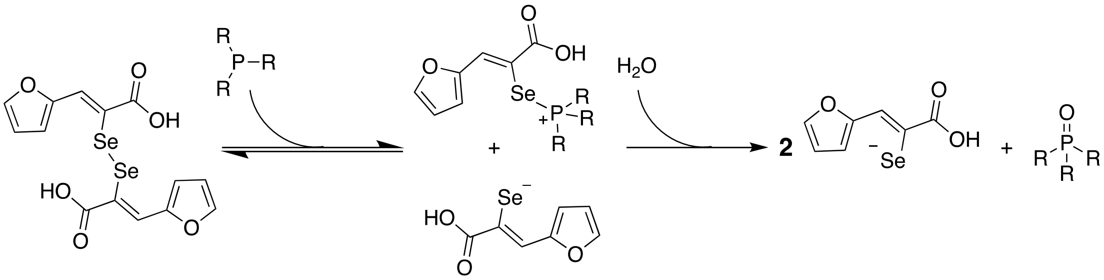


*Figure 2.2.6 - UV-vis spectra of diselenide (**40**) at various concentrations of TCEP. The peak at 306 nm was attributed to the diselenide peak and the peak at 250 nm to the selenoate anion.*

Distinct changes occurred upon the addition of TCEP to the original spectra of the diselenide compound. The intensity of the absorbance from 240-250 nm, attributed to the putative selenoate anion increased, along with the appearance of two new peaks at 285 and 348 nm. The original peak attributed to the diselenide bond at 300 nm also diminished. Whilst the literature is limited on the characterstics of P-Se bonds, applying an analogous methodology for TCEP reduction of disulfides to diselenide bonds would suggest a similar mechanism as shown (Scheme 2.8).[143]



*Scheme 2:7 - Proposed mechanism reduction of diselenide **40** by phosphine reagents.*

The UV-Vis spectra measured for the addition of TCEP to the diselenide (**40**) display more changes with increasing concentrations of TCEP than the thiol based reducing agents, which suggests that a more complex reaction possibly takes place (Figure 2.2.6).

### 2.2.4. Reduction of **40** with glutathione and cysteine

The diselenide was subject to reducing conditions by glutathione (**42**) up to 6 mM and monitored between 200-450 nm.

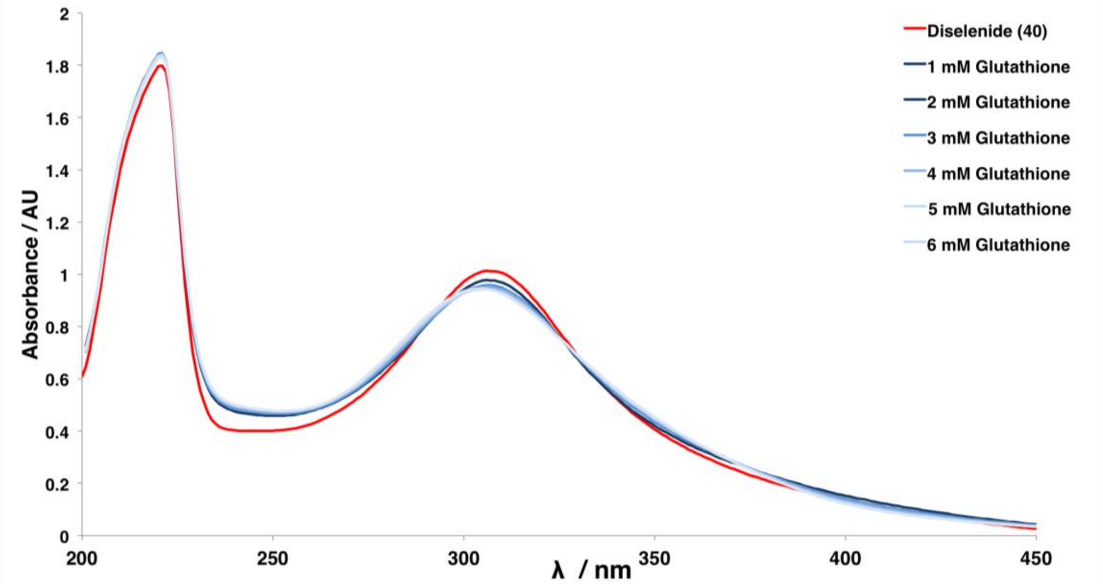


*Figure 2.2.7 - UV-vis spectra of diselenide (**40**) at various concentrations of glutathione. The peak at 306 nm was attributed to the diselenide peak and the peak at 250 nm to the selenoate anion.*

The changes in the UV-Vis spectra for **40** upon addition of glutathione were smaller compared to the other reducing agents tested. Reduction potentials for the glutathione disulfide system range from -0.17 to -0.26 V in the literature, meaning that the diselenide is not reduced as rapidly by this weaker reducing agent, which could possibly be due to the steric bulk of the agent.[137–139] Selenoglutathione is a tripeptide isolated from plants such as garlic consisting of γ-glutamic acid, selenocysteine, and glycine.[149] It has a much lower redox potential (-0.41 V) than glutathione, therefore is easily oxidised, and generally isolated as the diselenide from cells.[124,150]

Computational experiments published in 2004 predicted that nucleophilic attack on a diselenide from glutathione in the thiolate form is both kinetically and thermodynamically more favourable than on sulfur due to a lower reaction enthalpy, and the larger selenium atom being more able to form hypercoordinated intermediates.[151] More recent practical experiments however show that diselenide groups undergo exchange reactions with thiolates and the reaction rates where selenium is the electrophile can be up to 4 orders of magnitude higher than those with sulfur.[147]

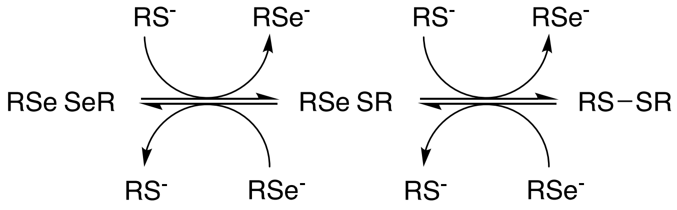


*Figure 2.2.8 – Proposed thiolate induced diselenide exchange reaction.*[147]

Reduction experiments were carried out using cysteine as a reductant, again titrating the thiol while monitoring the absorbance between 200-450 nm. The absorbance spectra showed that the diselenide peak at 300 nm decreased ,as well as an increase in absorbance between 240-250 nm, attributed to the formation of selenoate anions.[146,152]

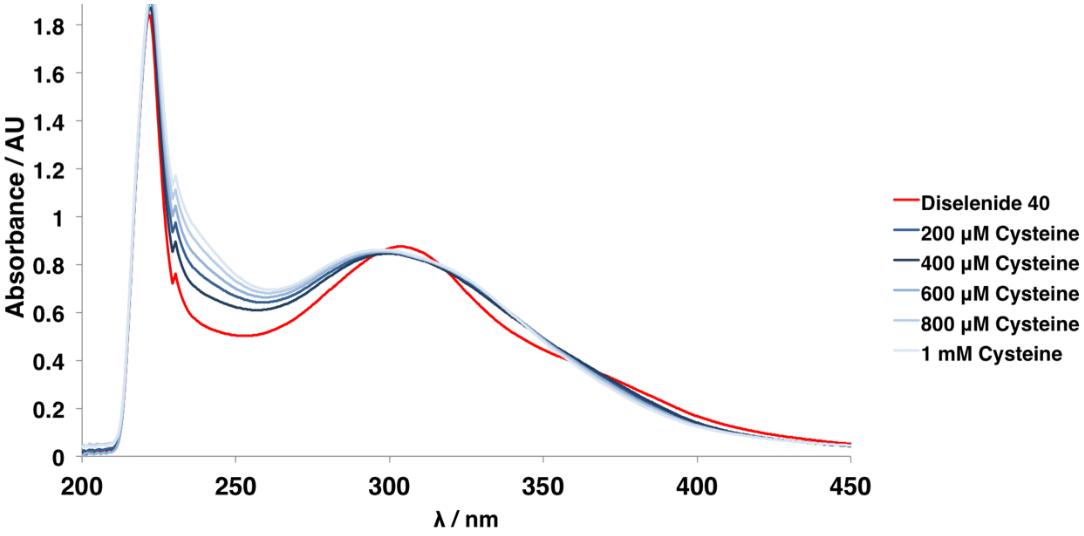


*Figure 2.2.9 - UV-vis spectra of diselenide (**40**) at various concentrations of cysteine. The peak at 306 nm was attributed to the diselenide peak and the peak at 250 nm to the selenoate anion.*

The cysteine reaction with **40** was explored by mass spectrometry to check for the presence of the possible S-Se predicted adduct. The characteristic isotope pattern of a molecule containing a single selenium atom was apparent and thus it was likely that this mixed sulfur-selenium species readily formed upon addition of a diselenide to a thiol containing compound such as cysteine.

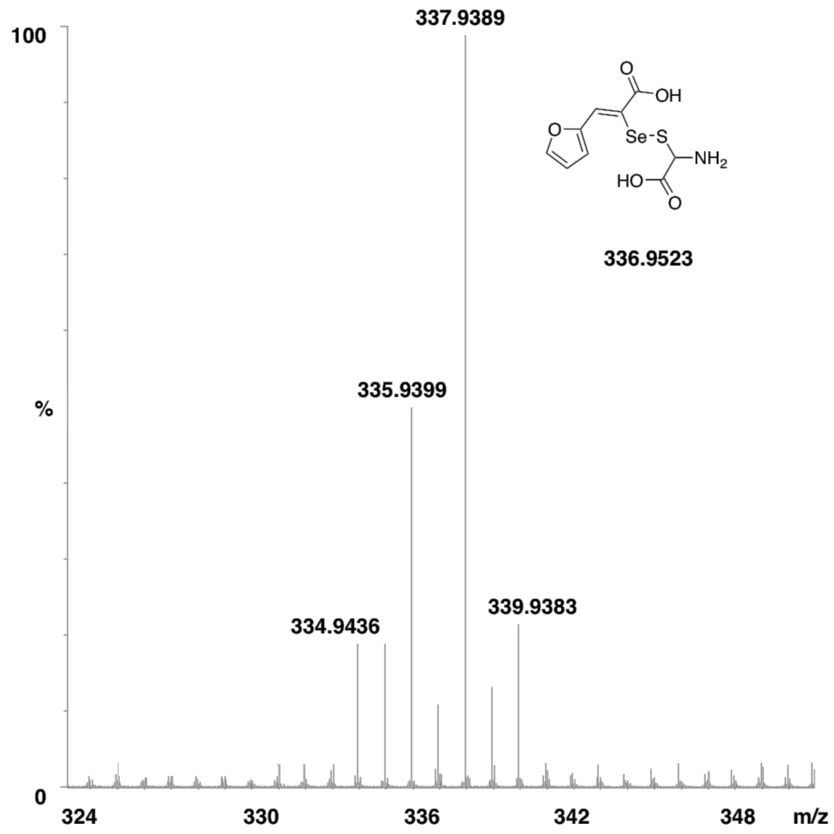


*Figure 2.2.10 –Mass of predicted sulfur-selenium metathesis product observed by HRMS ES$^+$, with a diagnostic selenium isotope pattern.*

The data presented suggests that the diselenides are not resistant to reduction from both thiol and phosphine based reducing agents. UV-Vis spectra for glutathione and cysteine displayed smaller changes than the more potent reductants tested however the presence of a selenyl-sulfide adduct apparent in the MS data would suggest that these compounds can undergo redox changes within a cell and likely form covalent bonds to cysteine residues. The implications of this UV-Vis and MS data for diselenides as redox stable calpain-1 inhibitors is that they are likely reactive to thiol based reducing agents including cysteine. It is also possible that they can react with surface exposed cysteine residues on proteins. If non-covalent binding to the PEF(S) hydrophobic site was observed alongside covalent adducts on calpain-1 a complex mode of inhibition for calpain-1 would result.

## 2.3. Testing for calpain-1 inhibition

The diselenides were tested for inhibitory action on calpain-1 by fluorescence resonance energy transfer (FRET) based assays. FRET techniques have transformed drug discovery methods, eliminating the need for gel electrophoresis and handling of radioactive material to quantify inhibitors in protease assays.[153] The standardised calpain-1 inhibition assay utilises a fluorogenic peptide from calpain-1 substrate α-spectrin ($K_M$ = 4.6 μM, $k_{cat}$ = 11 $s^{-1}$), which contains a 5-carboxyfluoroscein (FAM) fluorophore and 4-((4-(dimethylamino)phenyl)azo)benzoic acid (DABCYL) quencher, forming a FAM-DABCYL FRET pair. These moieties are situated at either end of an octapeptide ($H_2N$-K(FAM)-EVYGMMK(DABCYL)-OH), conjugated*via* lysine residues. Upon cleavage by calpain-1 between the Tyr-Gly residues, an increase in fluorescence is observed as the quenching effects diminish.

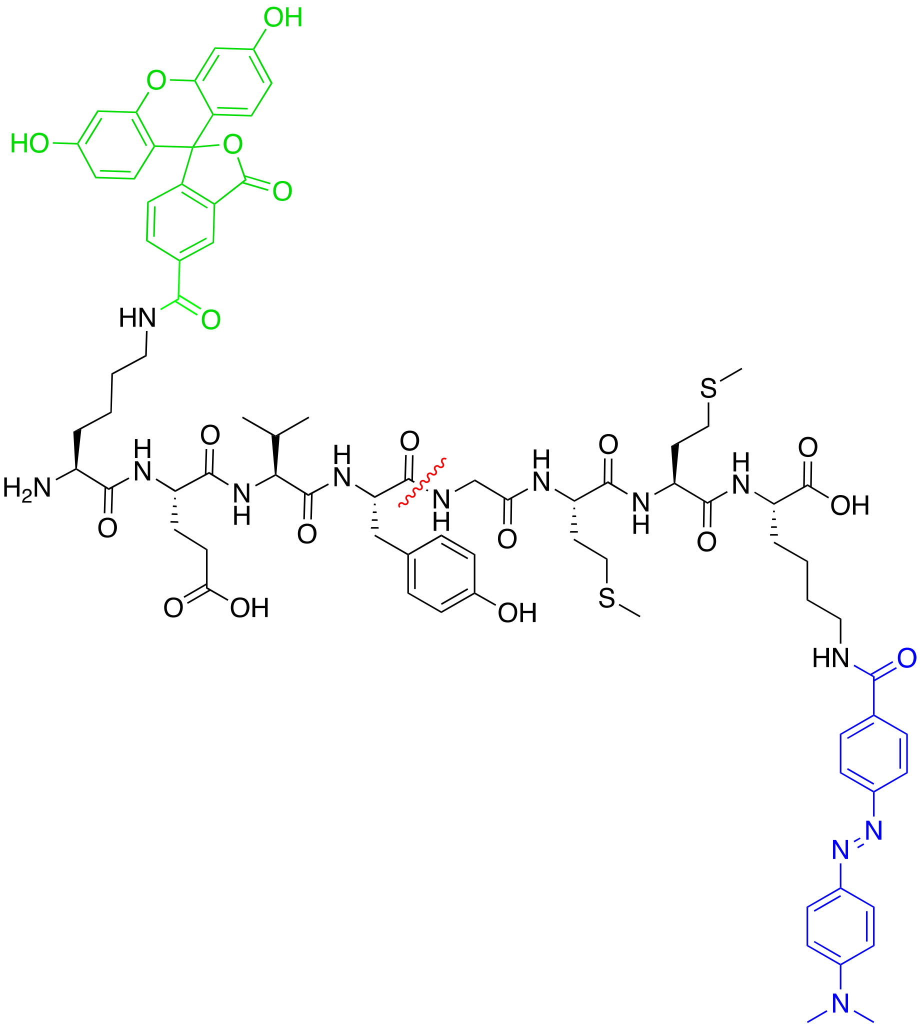


*Figure 2.3.1 - Calpain-1 substrate (H-K(FAM)-EVY/GMMK(DABCYL)-OH) containing FAM (green) – DABCYL (blue) FRET pair on a peptide based upon α-spectrin, with the cleavage site highlighted (red) between the tyrosine and glycine residues.*

The inhibition assay was carried out using human calpain-1 (Merck, 25 nM) in HEPES (10 mM), 9:1 glutathione reduced:oxidised (10 mM), EDTA (0.5 mM), bovine serum albumin (0.1%) at pH 6.8 with the FAM calpain-1 substrate (1 μM). The assay was carried out at 25 °C using an initial volume was 90 μL. The reactions were monitored in multiwell plates on a fluorescent plate reader (BMG Optistar) with an excitation band pass filter centred at 490 nm and emission detected at 520 nm. Data points were collected before the addition of 10 μL $CaCl_2$ (5 mM) to a final volume of 100 μL, to initiate the reaction.

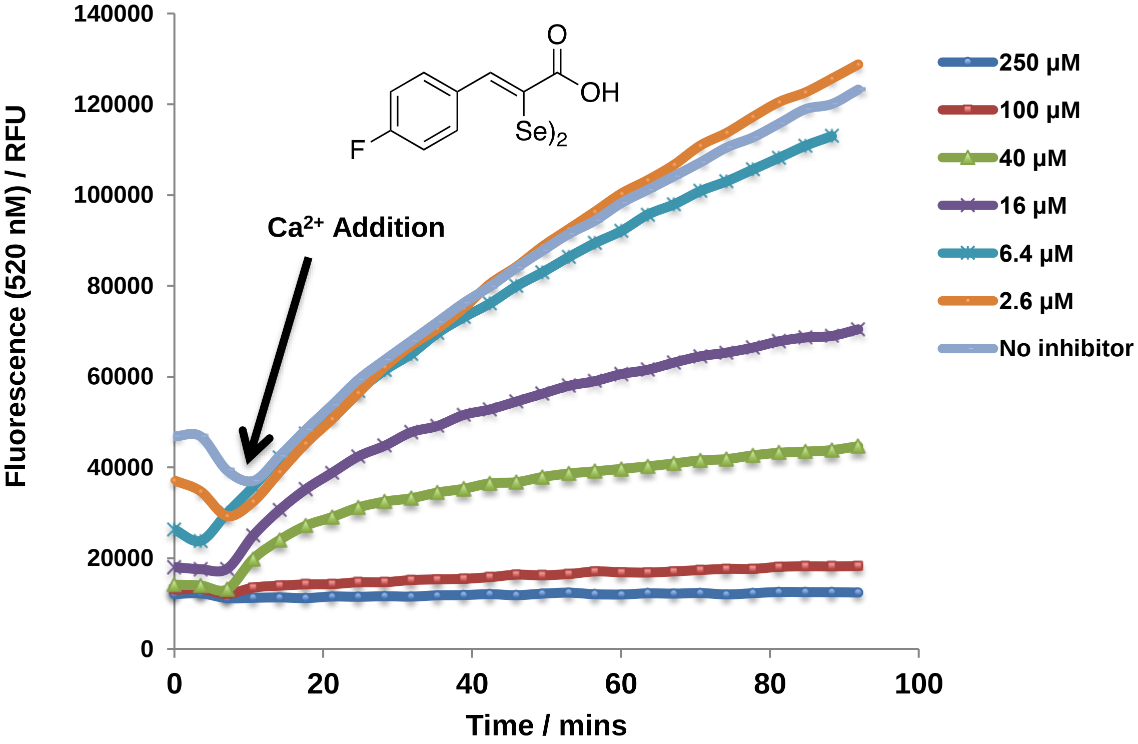


*Figure 2.3.2 - Calpain-1 inhibition assay monitoring the change in fluorescence at 520 nm at different concentrations* ***35*** *diluted from a 50 mM DMSO stock. Assays were performed with the calpain-1 FRET substrate (1 μM), calpain-1 (25 nM), HEPES (10 μM), glutathione 9:1 (reduced : oxidised), 10 (μM), EDTA (0.5 mM), BSA (0.1%) at pH 7.0. Addition of $CaCl_2$ (5 mM) initiated the reaction and the initial gradient was used for the $IC_{50}$ calculations.*

The increase of fluoresence intensity is directly proportional to the activity of calpain-1, the concentration of inhibitor required to reduce the uninhibited enzyme fluoresence intensity by half is given as the apparent $IC_{50}$ value for each respective compound. The initial rate of the reaction was taken after addition of $Ca^{2+}$. The change in fluorescence over time (Figure 2.3.2) as the concentration of **35** is increased shows that the rate of substrate cleavage is decreased. No quenching of the inhibitors was observed with the fluorogenic substrate. The curves do not approach the same final fluorescence intensity due to autolysis of calpain-1, eventually inactivating the enzyme.

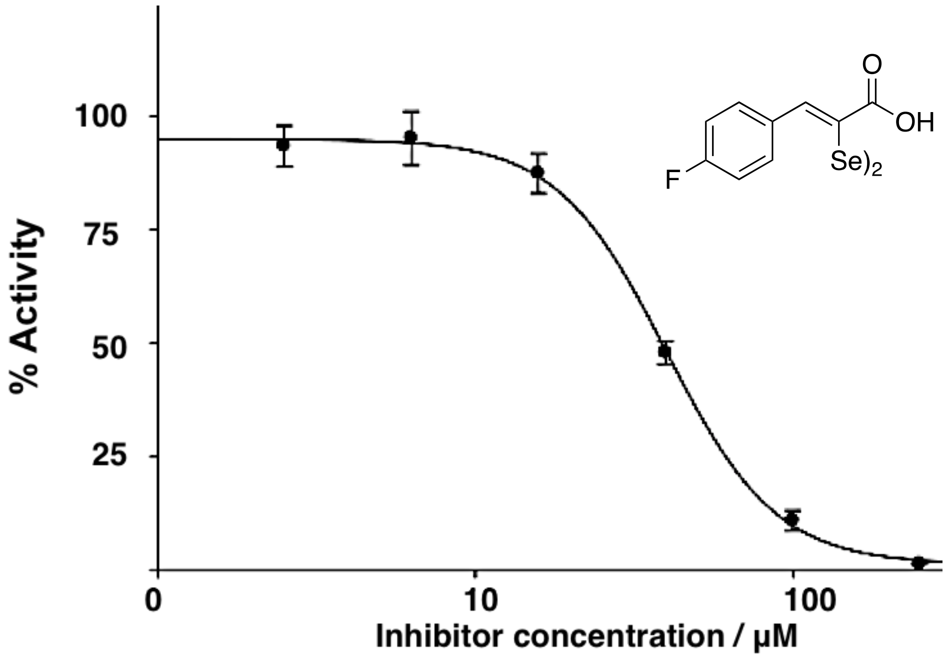


*Figure 2.3.3 – Activity for clapain-1at different concentrations of **35** in calpain-1 assay analysed by non-linear regression in Sigmaplot (Systat Software, San Jose, CA).*

The initial rate of each reaction in triplicate was analysed by non-linear regression in Sigmaplot (Systat Software, San Jose, CA) using the Ligand Binding module and the four parameter logistic equation (Equation 1).

$$y = \min + \frac{\max - min}{1 + (\frac{x}{EC50})^{-Hillslope}}$$

*Equation 1 - Four parameter logistic equation used to calculate inhibitor calpain $IC_{50}$ values of inhibitors using Sigmaplot (Systat Software, San Jose, CA).*

The data were normalised using the uninhibited enzyme as $B_{max}$ and total inhibition, no increase in fluorescence, as 0. DMSO was used to dissolve the inhibitors into stock solutions (50 mM) and its effect on the assay was monitored, at at higher concentrations DMSO displays a small inhibitory effect on calpain-1 therefore this effect was subtracted from the assay results.

The data from the assay for diselenide **35** was similar for the range of diselenides tested. The assay results revealed that the diselenides tested were moderately potent inhibtors of calpain-1, with $IC_{50}$ values in the micromolar region from 3.6 – 52.2 μM, displaying similar potency to the reduced α-mercaptoacrylic acid inhibitors.[128]

| **Compound** | **$IC_{50}$ (μM)** |
|---|---|
| PD150606 - **9** | 14.1 ± 1.6 |
| **29** | 21.9 ± 3.6 |
| **30** | 24.2 ± 1.8 |
| **31** | 11.1 ± 0.8 |
| **32** | 22.3 ± 2.7 |
| **33** | 26.6 ± 1.8 |
| **34** | 18.9 ± 0.4 |
| **35** | 10.4 ± 1.6 |
| **36** | 8.1 ± 0.7 |
| **37** | 9.8 ± 0.6 |
| **38** | 17.0 ± 3.6 |
| **39** | 3.6 ± 1.2 |
| **40** | 18.7 ± 3.0 |
| **41** | 7.2 ± 1.3 |

*Figure 2.3.4 - $IC_{50}$ values (μM) of the diselenides for inhibition of calpain-1. Assays were performed with the calpain-1 FRET substrate (1 μM), calpain-1 (25 nM), HEPES (10 μM), glutathione 9:1 (reduced : oxidised), 10 (μM), EDTA (0.5 mM), BSA (0.1%) at pH 7.0. Addition of $CaCl_2$ (5 mM) initiated the reaction; the calculated figure is the average of triplicate measurements.*

The 4-iodo selenium analogue of PD150606, **38** inhibited calpain-1 with a comparable potency to PD150606. The larger indole and anthracene based inhibitors **39** and **41** were the most potent diselenides with $IC_{50}$ values of 3.6 μM and 7.2 μM respectively. Diselenide **36** was the most potent phenyl based compound tested which in this assay was a more potent inhibitor than PD150606.

Upon oxidation, the α-mercaptoacrylic acids were reported to inhibit calpain-1 up to 1000x more potently.[128] The assays were repeated for the diselenide series of compounds with no reducing agent, allowing the inhibitors to incubate with the enzyme for 3 hours prior to calcium addition ,the same protocol was also carried out for PD150606.

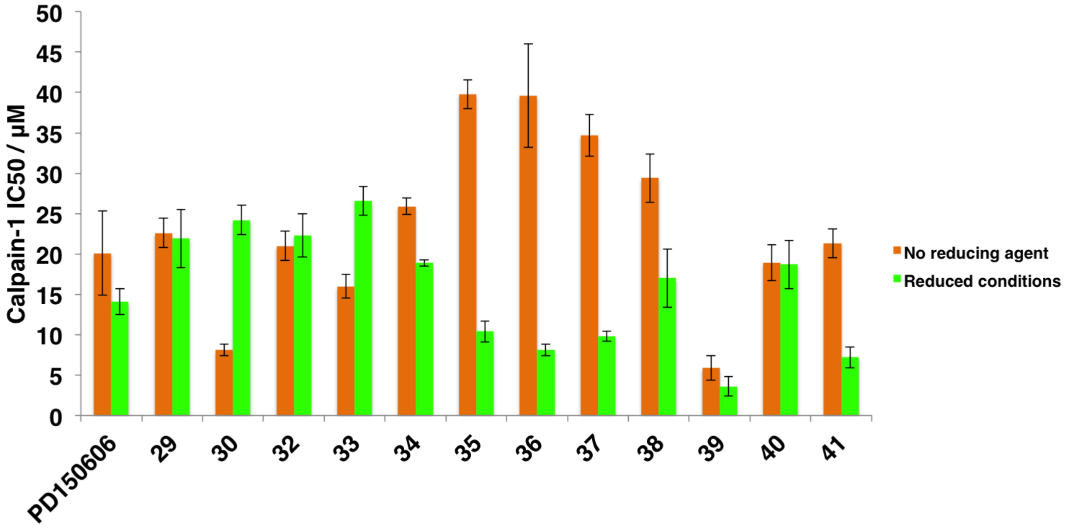


*Figure 2.3.5 - A histogram displaying calpain-1 inhibition IC$_{50}$ values for diselenide compounds under both reduced and oxidised conditions. Oxidised conditions were without reducing agent and left to incubate for 1 h before addition of CaCl$_2$.*

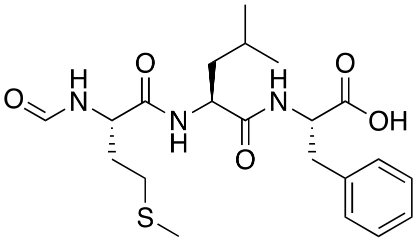
The inhibition of calpain-1 by PD150606 was reported to change from 5.0 µM to 0.0075 µM in the presence and absence of a reducing agent respectively.[65] This experiment was repeated several times with commercial enzyme and inhibitor however the effect was not observed and the data was not found to be reproducible in this work. In contrast, the data from these assays for PD150606 displayed an increase in IC$_{50}$ from 14.1 µM to 20.1 µM in the absence of a reducing agent. A four and three increase in potency was observed from oxidised to reduced conditions for **35, 36**, **37** and **41** respectively, while the chloro substituted phenyl compounds displayed a decrease in potency under reduced conditions. The cause of these changes is not immediately apparent however it is evident that these compounds are in fact redox sensitive. This has a direct effect on the potency as calpain-1 inhibitors, which complicates the mode of action.

Further data are required in order to determine whether these compounds bind to calpain-1 *via* the calcium binding domains or active site, and in light of the

MS-reduction data, whether these compounds form covalent bonds to the protein through selenylsulfide bonds.(Section 3.2, Section 4.1.8).

# 2.4. Live Cell Imaging

Neutrophils can be stimulated to rapidly spread upon an increase in intracellular calcium concentration. Addition of the short peptide *N*-formylmethionine leucyl-phenylalanine (**42**, fMLP) triggers release of $Ca^{2+}$ from intracellular stores which in turn activates calpain localised at the cell membrane, cleaving structural proteins allowing the cell to spread.[14,154]



**42**

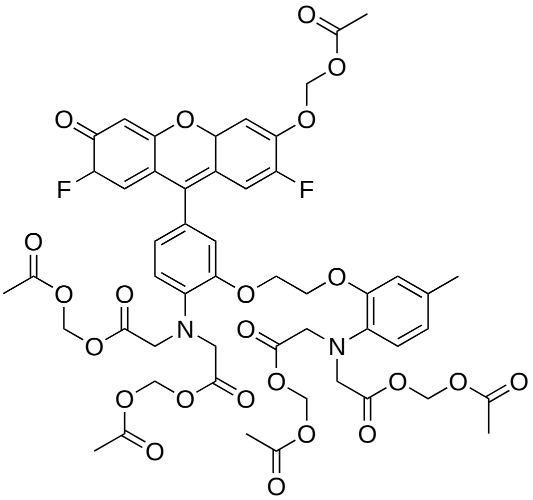
*Figure 2.4.1 - fMLP formylated peptide (**42**) used to trigger release of intracellular $Ca^{2+}$ to activate calpain for neutrophil spreading assay.*

Inhibition of the cell spreading process was used determine if the inhibitors were membrane permeable and if they also work on the target enzyme to perturb cell spreading in neutrophils, believed to be regulated by calpain.[14]

## 2.4.1. Neutrophil spreading assay

The short lifespan of neutrophils required immediate use after harvesting (Section 6.8.3). Krebs medium was used to suspend the cells (120 mM NaCl, 4.9 mM

$KH_2PO_4$, 1.2 mM $MgSO_4$, 1.2 mM $CaCl_2$, 1.3 mM KCl, 25 mM HEPES and 0.1% BSA at pH 7.4), and the inhibitor was added to a final concentration of 10 μM from a 50 mM DMSO stock solution. The solution also contained a cell permeable fluorescent indicator 4-Fluo-AM (**43**), a >100 fold in fluorescence occurs at 494/506 nm upon binding to $Ca^{2+}$.[155] This indicator was used to detect calcium released into the cell to initiate cell spreading, and as a control to ensure that the inhibitors did not simply prevent calcium release instead.



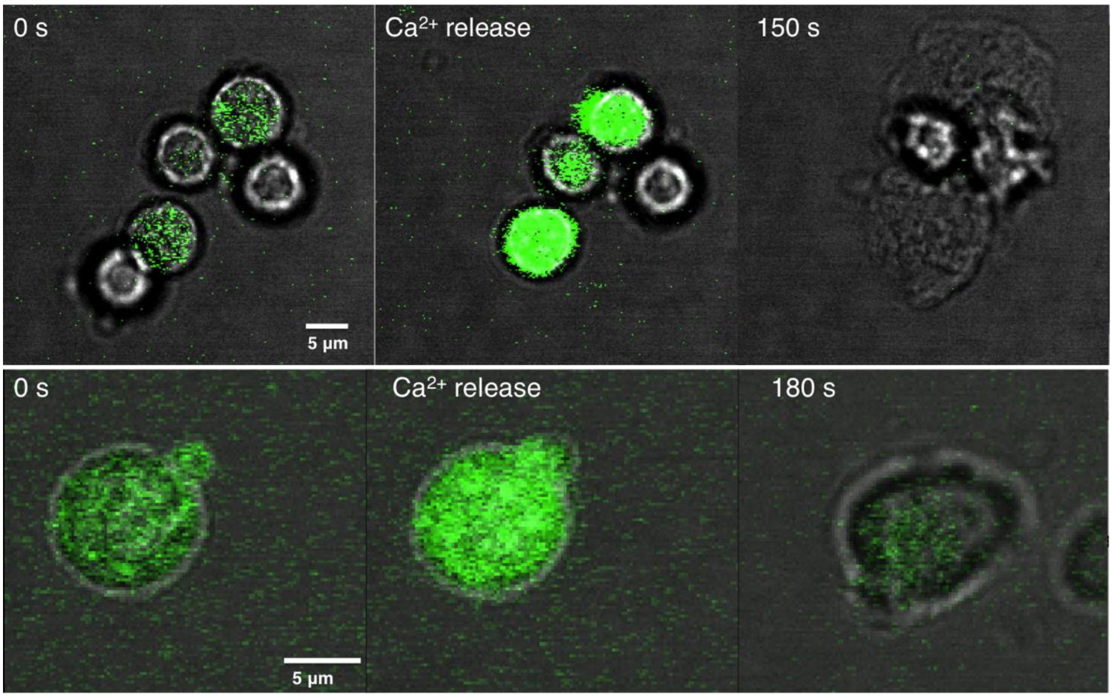
**43**

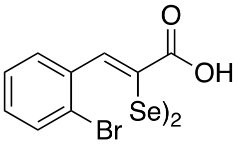
*Figure 2.4.2 - 4-FLUO-AM cell permeable fluorescent $Ca^{2+}$ indicator.*

*A microscope coverslip was prewarmed to 37 °C loaded with 100 μL cell solution containing **31**. After 2 minutes, cells that had adhered to the coverslip were monitored through a Leica confocal microscope.*

The spreading process was initiated by addition of fMLP (**42**,

Figure 2.4.1) peptide to release intracellular stores of $Ca^{2+}$, which was observed as a green flash on the microscope from the 4-Fluo-AM indicator (**43**, Figure 2.4.2). Only the cells that showed a distinct flash were monitored to ensure the calcium release and inhibitory effect could be correlated. The spreading process happened rapidly within 90 s in uninhibited cells, the neutrophils were monitored on the microscope for up to 3 minutes, the positive control contained 0.1% DMSO to ensure the solvent used to dissolve the inhibitor did not have an effect on the cells, whilst negative controls did not spread upon addition of fMLP.





**31**

*Figure 2.4.3 – Top row – control, left – neutrophils before induced spreading, centre – cells immediately after fMLP induced Ca$^{2+}$ release, right – cells 150 s after fMLP addition. Bottom row – inhibitor (**31**) (10 µM) left – neutrophils doped with (**31**), centre – cells immediated after fMLP induced Ca$^{2+}$ release, right – cells 180 s after fMLP addition. To view videos of spreading experiment, scan QR codes (left – control, right inhibitor)*

When doped with diselenide **31** (10 µM), the diameter of the neutrophils did not appreciably change, indicating that the spreading process was totally halted under the timescale used. The untreated neutrophils diameter doubled in size within 90 s following fMLP (**42**) addition (

Figure 2.4.4). This result suggests that the compounds are able to readily pass through the cell membrane and effectively inhibit calpain-1 in a live cell environment.[14,129,156]

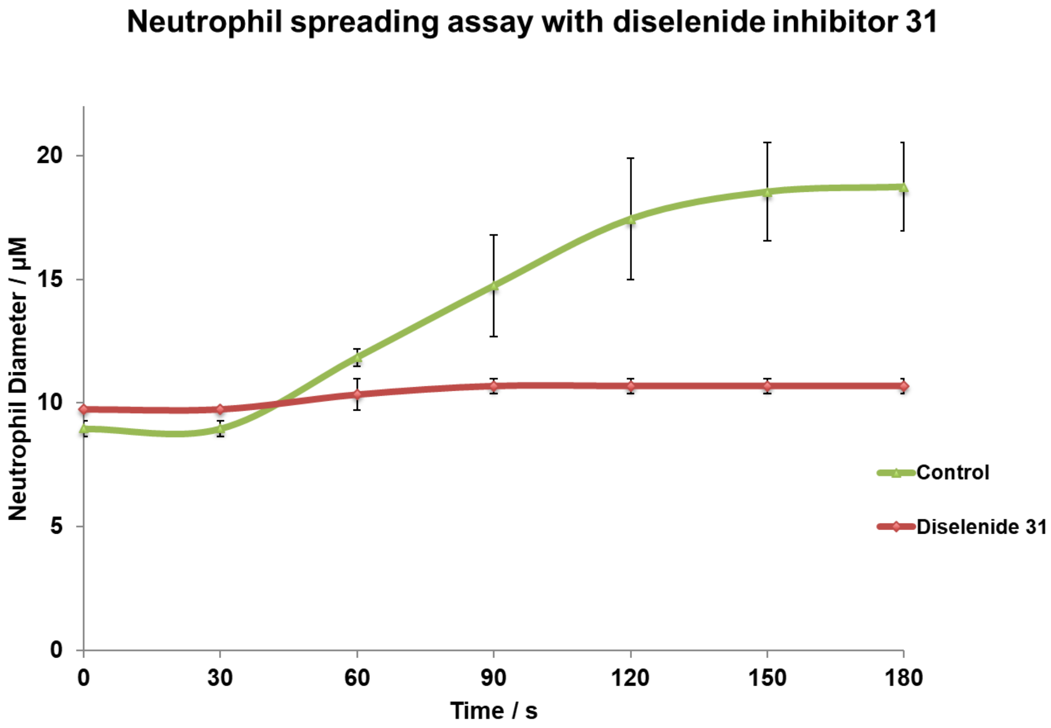


*Figure 2.4.4 - Neutrophil size change over time after addition of fMLP to initiate the rapid cell spreading process, size of neutrophils calculated using the Leica image processing software and ImageJ, at least 5 cells per experiment were measured.*

Pseudopods were observed protruding into the cell membrane after fMLP addition in the cells doped with diselenide **31** (

Figure 2.4.4). The structural proteins ezrin and talin linking the cytoskeleton to the membrane were likely not cleaved by calpain during in the inhibited spreading process, resulting in the pseudopods reaching the elastic limits of the membrane but unable to extend further.[11] This strongly suggests that calpain inhibition was occuring inside the cell at the membrane whilst other cellular processes associated with neutrophil spreading, such as $Ca^{2+}$ release and pseudopod extension were able to function.

### 2.4.2. Concluding remarks on diselenides

The diselenide analogues of the α-mercaptoacrylic acid calpain inhibitors were synthesised according to the proposed procedure (Section 2.1.2). Substituted phenyl, indole, and furan derivatives were synthesised to provide a comparison for a range of potential inhibitors.

Analysis of the compounds' redox properties was carried out by UV-Vis spectroscopy to explore the resistance to reduction. The UV-Vis spectra for diselenide **40** rapidly exhibited profound changes upon addition of potent reductants; DTT and TCEP, whilst glutathione and cysteine less so. This suggests that the diselenides are unlikely to remain in their oxidised form for a prolonged period of time in the reducing cellular environment, therefore unlikely to maintain the dihedral bond geometry associated with allosteric inhibition of calpain-1 through PEF(S) binding.

Furthermore, given the MS data, it is likely that thiolate groups can initiate metathesis reactions with diselenides, forming mixed sulfur-selenium species in the process. Covalent bonds formed between reactive cysteine residues and small molecules have been exploited in warhead inhibitor design as discussed in section 1.5.3. They target the active site with a high selectivity, which prevents cross-reactivity with other proteins or indeed any thiol bearing structure. When the diselenides are exposed to surface cysteine residues on proteins it is likely covalent adducts could be formed to a multitude of other proteins, which in the cellular environment could have off target effects.

The diselenides **29 – 41** were tested for calpain-1 inhibitory properties by a FRET-based assay, displaying micromolar $IC_{50}$ values. Attempts at repeating the dramatic increase in potency reported for α-mercaptoacrylic acids, such as PD150606, from a reduced to oxidised state were unsuccessful. This assay was repeated for the diselenides, which also did not display a 100-1000x change in potency. Compound **31** was tested in a neutrophil spreading assay to see if the rapid cell spreading process was affected by diselenides. Within 90 seconds of intracellular $Ca^{2+}$ release, triggered by fMLP, the neutrophil underwent rapid cell spreading, doubling in diameter. When doped with **31** (10 μM), the cell diameter remained unchanged for at least 200 seconds, indicating that these compounds are both cell permeable and effective inhibitors of calpain spreading behavior, likely through calpain-1 inhibition.

# Chapter 3 - Interactions with PEF domains

# 3.1. Small Molecule Interaction with PEF(S)

## 3.1.1. α-Mercaptoacrylic acids

The inhibitory region C of calpastatin binds to a hydrophobic pocket on PEF(S), where α-mercaptoacrylic acid PD150606 (**9**) also binds (Figure 3.1.1).[37] The hydrophobic leucine, phenylalanine and tryptophan residues in the pocket facilitate favourable hydrophobic interactions with the aromatic ring of the ligand, in a similar manner to CAST binding.

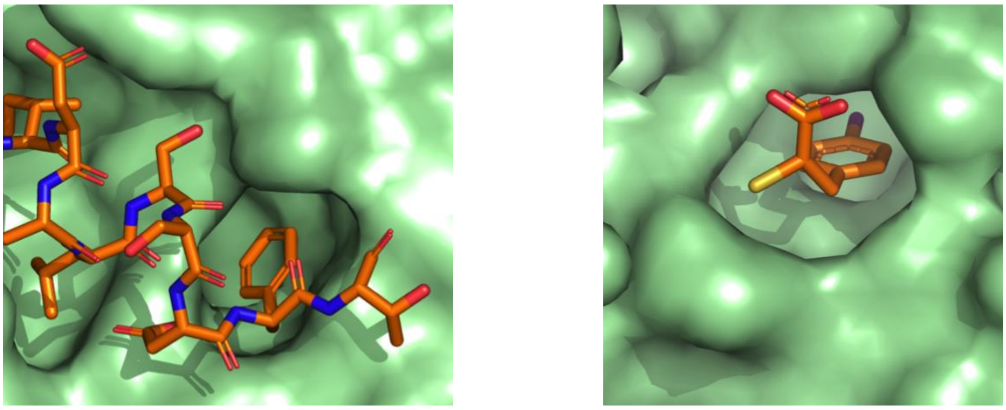


*Figure 3.1.1 - PEF(S) hydrophobic pocket for binding calpastatin (left, PDB:1NX0) and α-mercaptoacrylic acid inhibitor PD150606 (**9**) (right, PDB:1NX3), protein surface represented as green surface and ligands shown as orange sticks.*[37]

The disulfide form of α-mercaptoacrylic acids contain a second aromatic ring that interact with a another hydrophobic region on the protein, suggested to be responsible for the increase in potency of these inhibitors.[128] Multiple conformations of the disulfides have been observed in the electron density map of crystallographic experiments, which implies that the binding geometry isn't optimal.[128] SAR of the inhibitor profile to fit the PEF(S) site may produce even more potent compounds.[63] The dihedral bond geometry of the disulfide group is postulated to be responsible for the increased binding interactions of the disulfides.[128] Prearrangement of a ligand to a receptor reduces the entropic penalty of binding by reducing the conformational changes that need to occur for a binding event.[119]

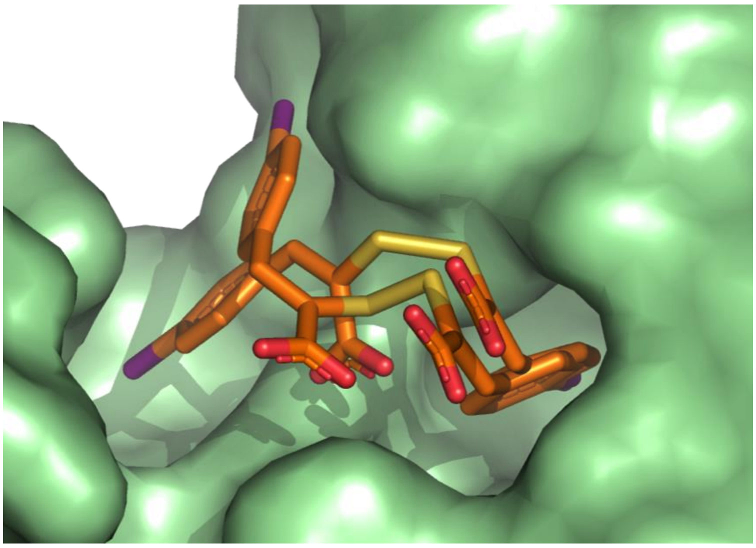


*Figure 3.1.2 - PD150606 (**9**) in disulfide form bound to PEF(S), protein represented as green surface and ligand shown as orange sticks (PDB:5D69).*[128]

In order to investigate whether the diselenide inhibitors displayed analogous allosteric binding interactions to the α-mercaptoacrylic acid derivatives, human PEF(S) was expressed in *Escherichia coli* (*E.* coli) and purified to provide the material for binding and crystallographic experiments.

### 3.1.2. Expression of PEF(S)

In order to produce human PEF(S) in large quantities an *Escherichia coli* expression system was used. Commercially available *E.coli* strains are easy to handle and display fast growth kinetics with a doubling time of only 20 minutes in optimal environmental conditions. Their ability to reach high cell densities makes them a popular choice of host organism for recombinant protein production.[157] These prokaryotes are readily manipulated and culture is inexpensive, in recent years the development of new strains, affinity tags and expression vectors have further increased the power of this protein production technique.[158]

The gene containing human calpain-1 PEF(S) was previously ligated into a pET-21d(+) plasmid and stored in the Allemann plasmid library, entry 1117.[159] Gene transcription in a pET vector is carried out by the highly active and selective T7 RNA polymerase, able to generate vast copies of mRNA.[160] Expression of the gene is under *lac* control where the LacI repressor binds to the operator, preventing T7 polymerase binding to the plasmid for transcription of the gene. The LacI repressor, when bound to a β-galactoside e.g. lactose, cannot bind to the operon, triggering transcription of the gene.[161] As lactose is cleaved by β-galactosidase, a metabolically stable analogue, isopropyl-β-D-thiogalactoside (IPTG) is used for practical purposes.[162]

The pET-21(+) plasmid containing the human PEF(S) gene was transformed into the BL21-CodonPlus(DE3)-RP *E.coli* strain by heat shock at 42 °C using chemically competent cells (Section 6.4.5). These cells are advantageous for protein expression due to the reduction of internal protease production, whilst the DE3 contains the genes for T7 RNA polymerase on the *E. coli* chromosome, additionally the RP CodonPlus strain contains extra copies of the genes that code for the rare tRNAs such as *argU*, *ilyY* and *leuW*, which can limit the heterologous protein production rate.[163] The pET-21(+) plasmid also contains an antiobiotic resitance gene (β-lactamase), allowing the cells containing the plasmid to degrade ampicillin and other β-lactams which is used for selective growth of cells.[164]

The transformed cells were grown in 1 L ampicillin selective LB medium (1 µg / mL) shaking at 37 °C (250 rpm) until an optical density ($OD_{600}$) of 0.6 a.u was reached. Protein expression was induced with IPTG (238 mg) to a final concentration of 1 mM and grown for 4 hours at 37 °C. Cells were harvested by high-speed centrifugation at 4 °C (6080 RCF, 15 min), and re-suspended in PEF(S) buffer A (Tris-base (20 mM), β-ME (5 mM), EDTA (2 mM) at pH 8.0). After sonication and clarification by centrifugation (30310 RCF, 45 min) the lysate was further purified by fast protein liquid chromatography (FPLC).

### 3.1.3. Anion exchange chromatography

The cell lysate mixture in low salt buffer (PEF(S) Buffer A, Section 6.6.1) was loaded onto a Q-Sepharose column to bind the protein to the positively charged resin, an increasing NaCl gradient was passed through the column, displacing proteins off the solid phase through different binding affinities to the resin. The protein was eluted with a linear salt gradient from a concentration of 0 M to 0.5 M NaCl over 5 CV, with 10 mL fractions collected whilst monitoring the UV absorption at 280 nm. (Section 6.6.12)

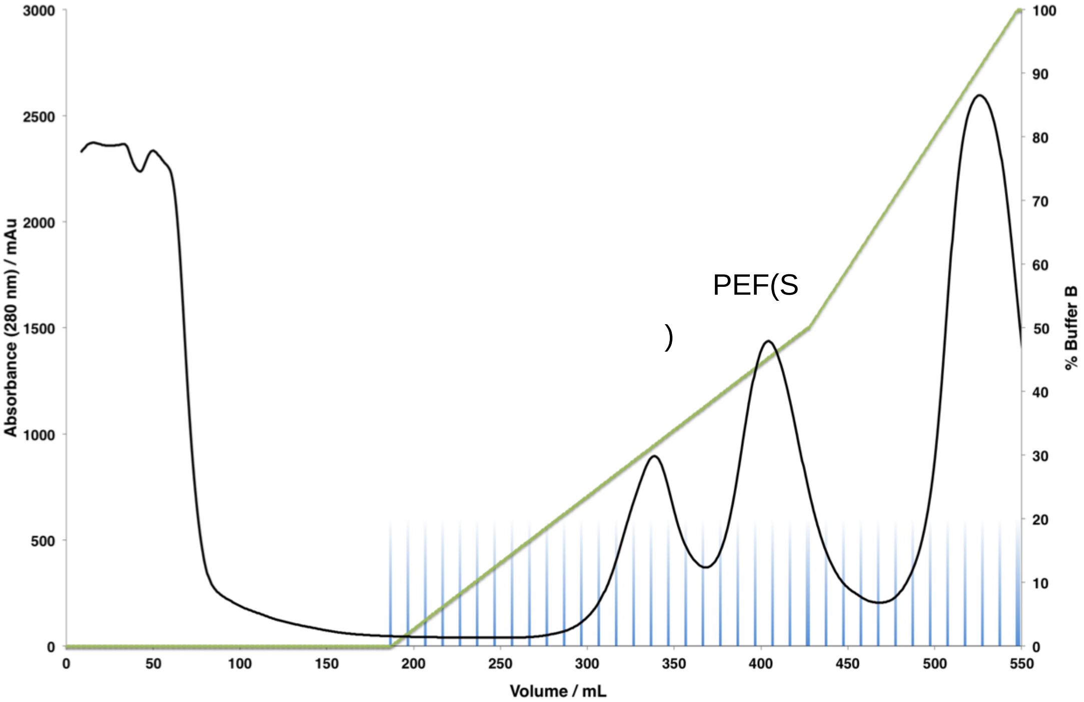


*Figure 3.1.3 - Q-Sepharose purification of human PEF(S) expression from E.coli lysate, the protein was eluted using a gradient from 0 – 5  M NaCl over 5 column volumes collecting 10 mL fractions, monitoring the UV absorbance at 280 nm.*

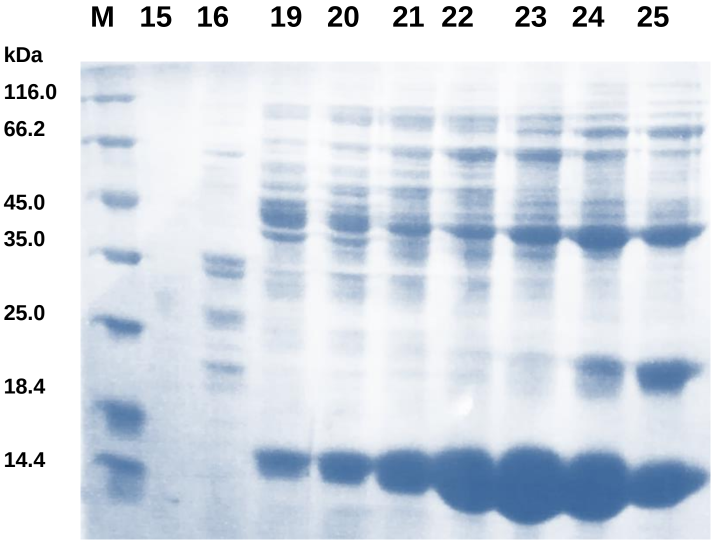


*Figure 3.1.4 - SDS-polyacrylamide gel of anion exchange chromatography (Q-Sepharose) purification of recombinant human PEF(S) where M is the protein marker and the numbered fractions of the eluent.*

Samples (5 µL) were taken from each fraction on a peak of the UV absorbance trace and analysed by SDS-PAGE. The target protein eluted at 0.35 M NaCl in fractions 20-25, these fractions were pooled and concentrated to 5 mL using an Amicon® ultrafiltration membrane with a molecular weight cut off (MWCO) at 30 kDa.

### 3.1.4. Size exclusion chromatography

Size exclusion chromatography, often referred to as gel filtration, separates molecules according to their molecular weight and oligomeric state in solution.[165] Smaller molecules take a longer path through the column whilst larger proteins are excluded from the porous solid phase and elute first.[166] PEF(S) was purified using a SuperDex™ 75 (GE Heathcare Life Sciences, Buckinghamshire, UK) column which readily resolves proteins from 3 – 70 kDa. The column (CV 320 mL) was equilibrated with PEF(S) Buffer C (6.6.3) (Tris-Base (20 mM), EDTA (1 mM), βME (5 mM), NaCl (0.15 M) at pH 8.0). PEF(S) solution (5 mL) was loaded onto the column and eluted with 1.5 CV of PEF(S) buffer C, the UV absorbance at 280 nm was monitored and 10 mL fractions containing protein were analysed by SDS-PAGE.

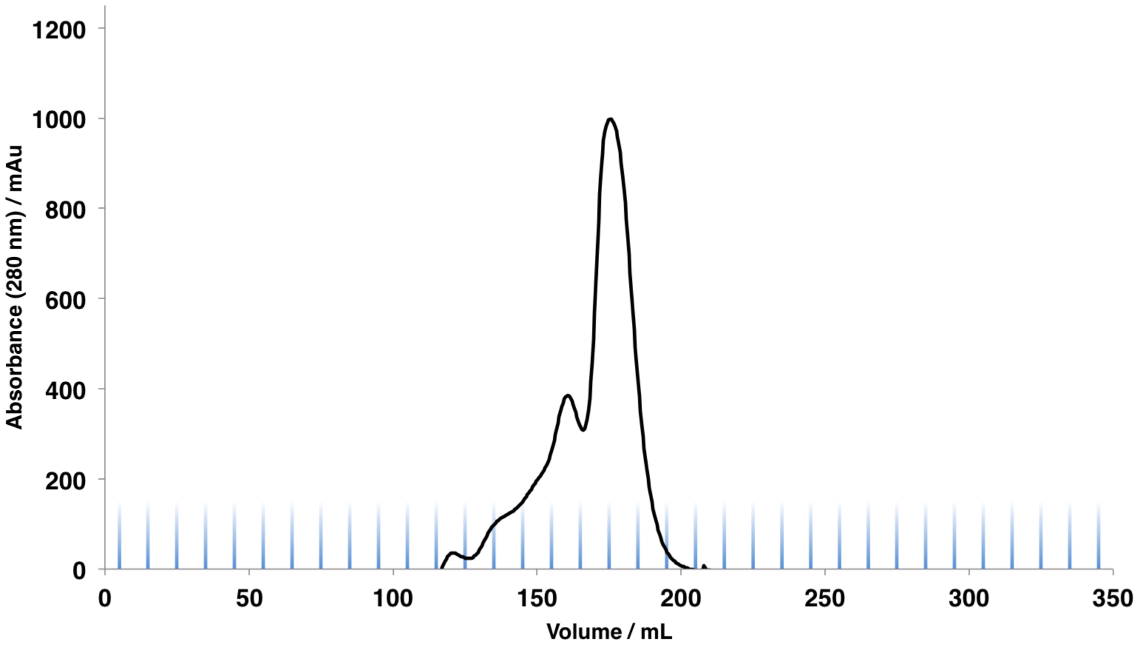


*Figure 3.1.5 – Size exclusion chromatogram with a SuperDex™ 75 (GE Heathcare Life Sciences, Buckinghamshire, UK) column running PEF(S) at a flow rate of 2.5 mL/min,*

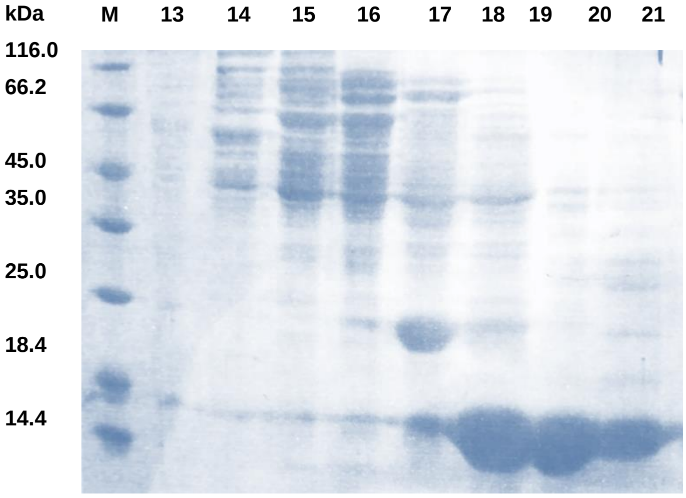


*Figure 3.1.6 - SDS-polyacrylamide gel of Superdex™ 75 purification of combined concentrated Q-Sepharose PEF(S) fractions where M is the protein marker and the numbered fractions of the eluent.*

Fractions containing PEF(S) were pooled and concentrated to 3 mL using an Amicon® ultrafiltration membrane with a molecular weight cut off (MWCO) of 30 kDa.

Protein purity for crystallography is important, at least 95-99% pure is recommended as a 1 p.p.m contaminant at the usual crystallographic concentration range of 10-20 mg/mL protein can equate to $10^9$-$10^{10}$ molecules in a milliliter. This can impact with the nucleation process and poison the crystal growth with defects, reducing data quality.[167]

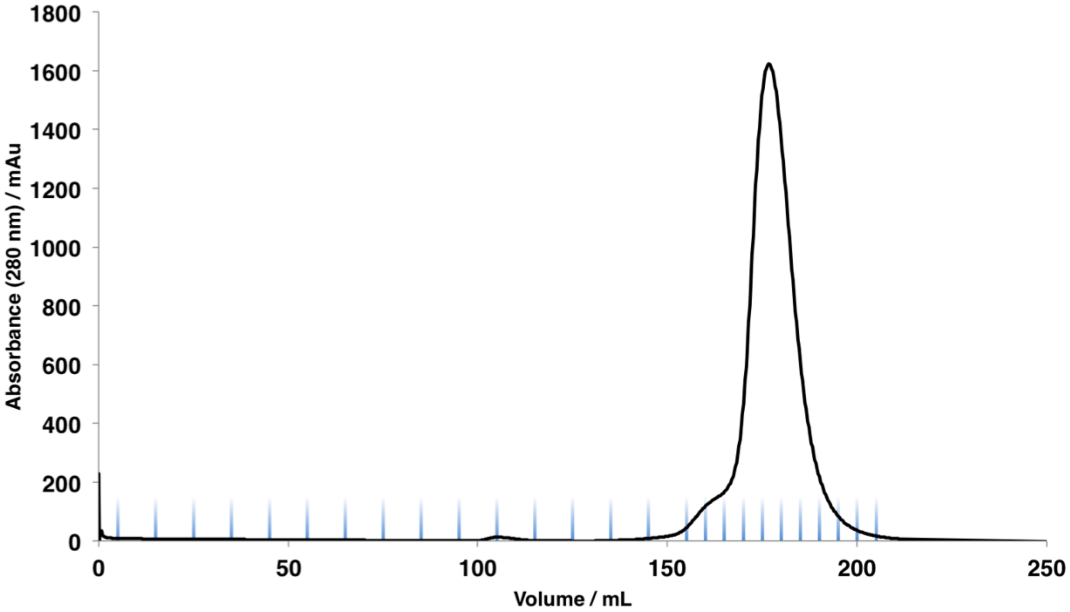


*Figure 3.1.7 - Second elution of size exclusion chromatogram with a SuperDex™ 75 (GE Heathcare Life Sciences, Buckinghamshire, UK) column running PEF(S) at a flow rate of 2.5 mL/min, once significant peaks were started to elute, 5 mL fractions were collected.*

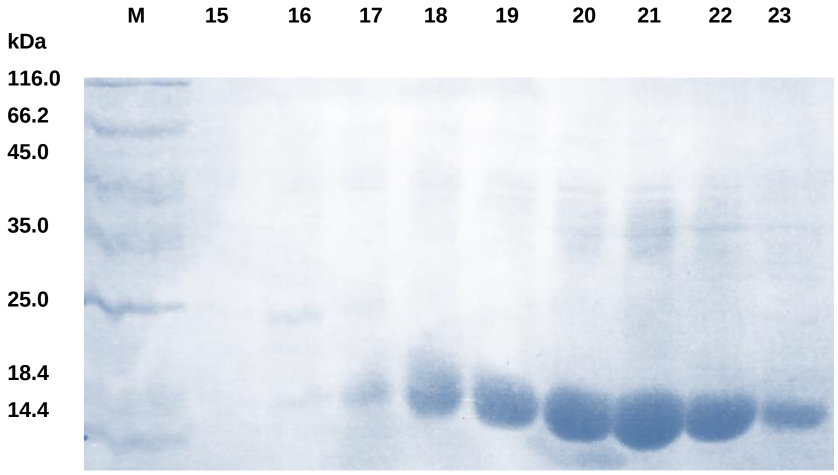


*Figure 3.1.8 - SDS-polyacrylamide gel of second SuperDex™ 75 chromatography step of the PEF(S) purification where M is the protein marker and the numbered fractions of the eluent.*

Sufficiently pure samples of human PEF(S) were produced in three chromatographic steps to proceed with the characterisation and crystallisation experiments. Despite the size of 20 kDa of this construct, it was observed to run on 10% SDS-PAGE gels just below the 18.4 kDa marker on the ladder, which is in accordance with the literature.

### 3.1.5. Mass spectrometry

Mass spectrometry was used to ensure the recombinantly produced PEF(S) was of the correct molecular weight. A sample (0.5 mg / mL) of PEF(S) was analysed by electrospray ionisation mass spectrometry (ES-MS) in positive mode and deconvoluted using the MaxEntropy1 program.

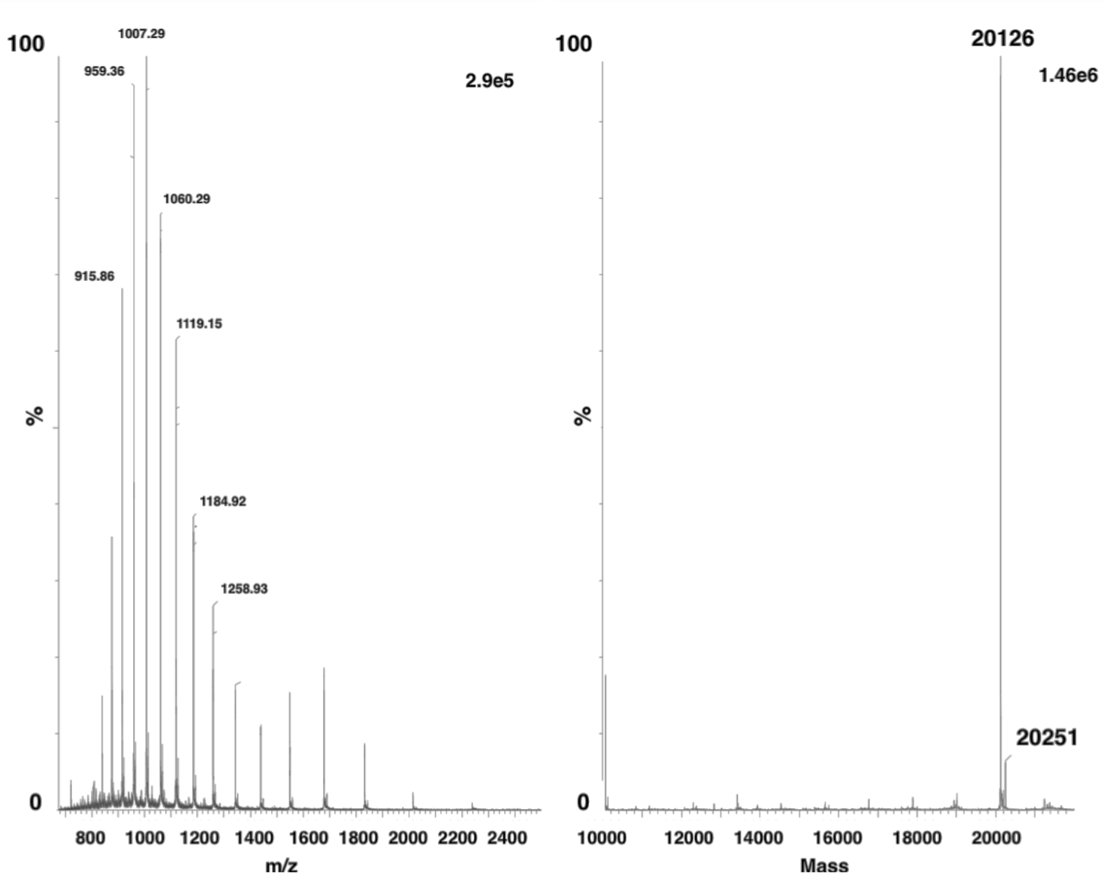


*Figure 3.1.9 – Mass spectrum (ESI$^+$) of purified recombinant human PEF(S) (left), deconvoluted MS using MaxEntropy1 program (right).*

The protein sample mass agreed with the predicted mass of 20126 Da for human PEF(S), indicating that this was indeed the correct protein.

### 3.1.6. Analytical size exclusion chromatography

Size exclusion chromatography using a high resolution analytical column was used to provide insight into the oligomeric state of PEF(S) in solution. The SuperDex™ 200 10/300 GL size exclusion column, composed of cross-linked agarose and dextran, is able to separate proteins between 10-600 kDa at high resolution.[168] By eluting proteins of known sizes and recording the elution volume, an unknown sample can be compared to these values and used to calculate an approximate size of the protein or complex in solution.

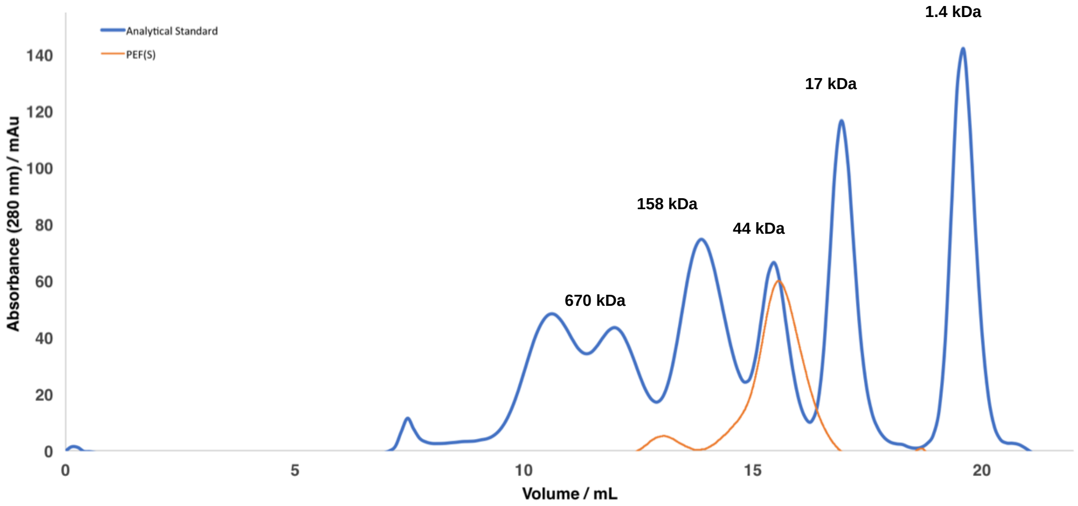


*Figure 3.1.10 – Analytical size exclusion chromatogram using Superdex™ 200 10/300 GL using BioRad™ analytical standards (bovine thyroglobulin - 670 kDa, bovine γ-globulin - 158 kDa, chicken ovalbumin - 44 kDa, horse myoglobin – 17 kDa, cyanocobalamin – 1.4 kDa), 100 µL of 90 µM PEF(S) passed through the column at a flow rate of 0.4 mL / min.*

PEF(S) eluted from the column at (15.9 mL), suggesting it primarily forms a dimer in solution, which agrees with reported observations.[48,51]

### 3.1.7. Circular dichroism

Circular dichroism (CD) was used to ensure the protein was folded into the correct secondary structure in solution. CD is the measurement of far UV wavelengths (180-260 nm) from the absorption of left and right circularly polarised light from chiral environments. [169] It is a widely used tool for the rapid elucidation of protein secondary structure and an important characterisation step in the production of recombinant proteins to ensure correct folding. Unlike NMR or X-ray crystallography It does not give residue specific information however it is advantageous as measurements can be made on samples containing 20 µg or less of protein. [170]

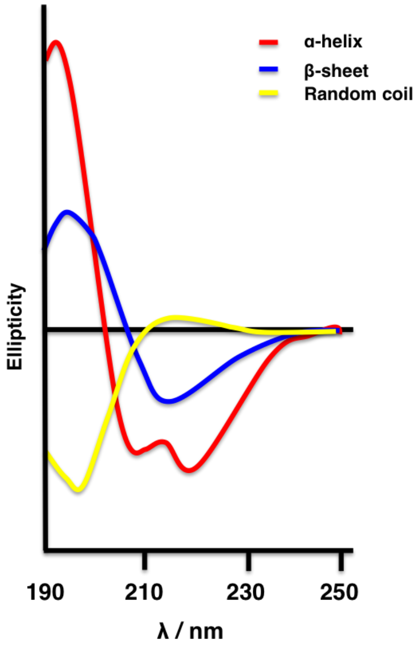


*Figure 3.1.11 – Idealised CD spectra showing the far UV optical rotations for each respective protein secondary structure.*[169]

The secondary structure of a protein can include α-helices, β-sheets and/or random coils/loops (Figure 3.1.11), when circular polarized light hits a chiral molecule the orientation of the light is changed. The change in this orientation at a given wavelength is specific to each secondary structure type, for example the presence of a α-helices in a CD spectra shows a negative peaks at 208 nm and 220 nm and a positive peak at 192 nm.[169] β-Sheets display a negative peak at 216 nm and a positive peak at 200 nm whilst random loop secondary structure presents a negative peak at 195 nm in the CD spectra.[169]

The pure PEF(S) protein sample was diluted to 10 μM in $K_iPO_4$ (10 μM, pH 7.4), the CD spectrum was recorded between 190-300 nm with 1 nm steps at 20 °C.

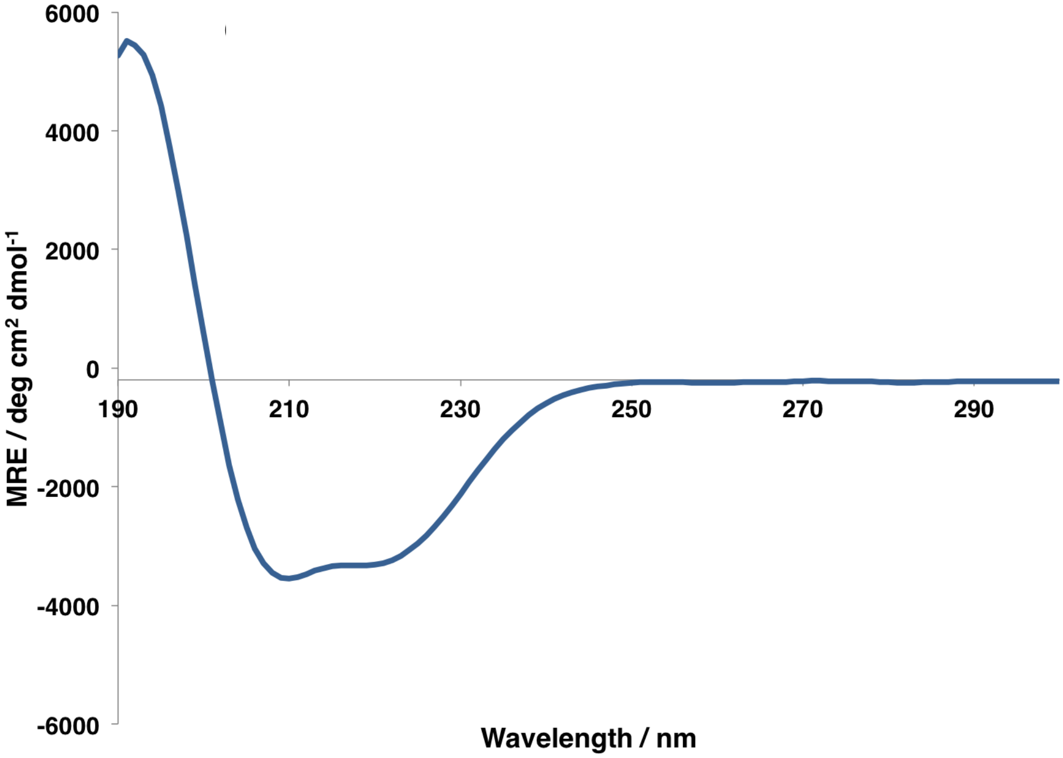


*Figure 3.1.12 - CD spectra of human PEF(S), mean residue ellipticity calculation used Equation 3, section 6.7.8.*

The two negative peaks at 208 nm and 220 nm strongly indicated the secondary structure of PEF(S) was predominantly α-helical.

## 3.2. Fluorescence Binding Assay

A fluorescence assay was used to explore the interaction of small molecules with the purified human PEF(S) domain. Small molecule fluorophores are highly sensitive to environmental variation and thus are an invaluable tool in chemical biology for investigating dynamic biological processes.[171,172] The change in fluorescence of a probe while titrating a small molecule to the protein of interest can provide valuable binding information such as the dissociation constant ($K_d$) of the compound.[106,171,173]

### 3.2.1. TNS displacement assay

6-(*p*-Toluidino)-6-naphthalenesulfonic acid (TNS, Figure 3.2.1) is a fluorescent probe with an excitation wavelength maximum at 318 nm and emission wavelength maximum between 430 and 450 nm when in a hydrophobic environment, however when displaced to an aqueous environment the fluorescence is quenched.[174] TNS binds to a hydrophobic pocket on PEF(S) and is displaced by PD150606 into the solvent, resulting in a decrease in fluorescence, crystallographic data has confirmed the presence of the hydrophobic binding site on PEF(S).[117,128]

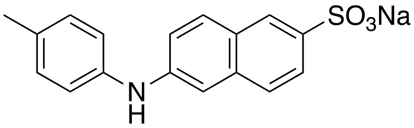


*Figure 3.2.1 - Fluorescent probe 6-(p-Toluidino)-6-naphthalenesulfonic acid. (TNS).*

PEF(S) (10 µM) in TNS binding buffer A, (Tris base (20 mM), $CaCl_2$ (1.1 mM), EDTA (1 mM, pH 7.4 was incubated with TNS (46.7 µM final concentration from 1 mM stock in 40% ethanol) in a Greiner CELLSTAR 96 well black flat bottom plate for 5 minutes at 25 °C. PD150606 (Sigma Aldrich) was stored as a 50 mM DMSO stock solution, which was diluted from 6/400 µM to 500 nM by serial dilution over 10 wells, in triplicate, with an Integra Viaflow 96 multichannel pipette.

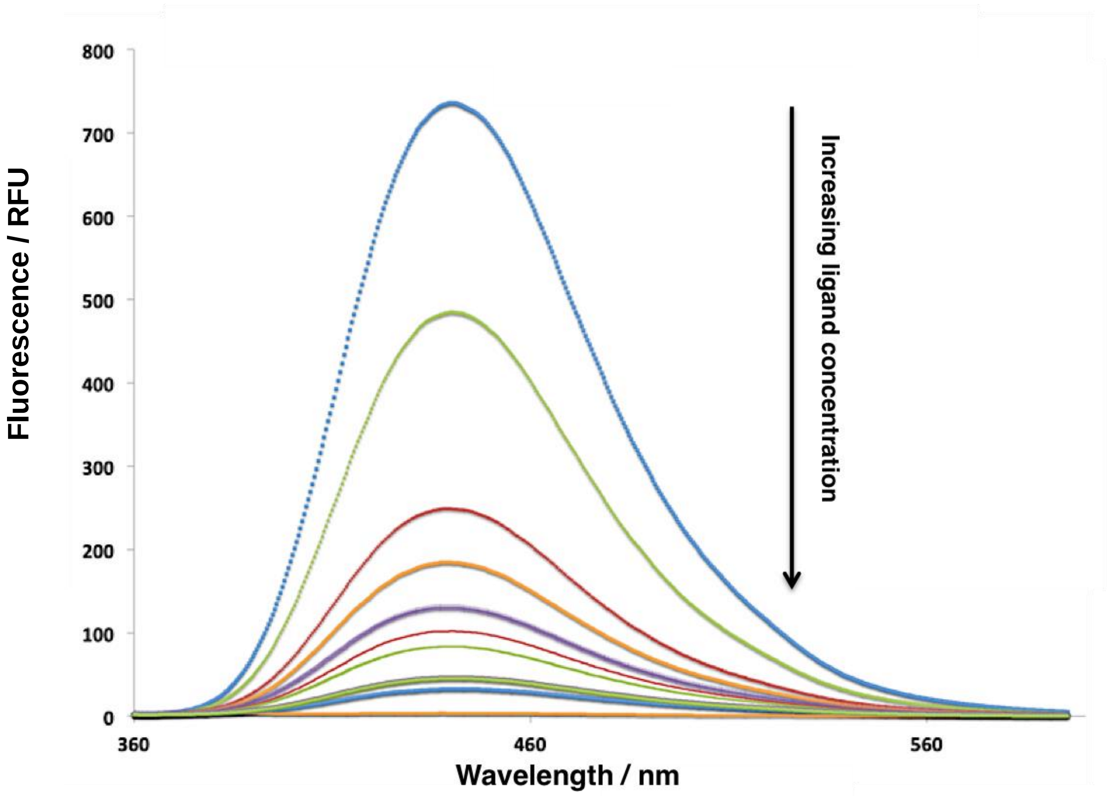


*Figure 3.2.2 - Change in fluorescence between 360-600 nm of TNS fluorescent probe displacement by titrating PD150606 (9) to a 10 μM solution of PEF(S).*

No fluorescence was observed unless TNS and PEF(S) were both present in the reaction mixture. The maximum at 444 nm was recorded for each concentration of PD150606, the data were normalised and inverted by subtracting the fluorescence (B) from the maximum fluorescence, $B_{max}$. These data were then analysed by non-linear regression using the Sigmaplot (Systat Software, San Jose, CA) ligand binding module (Equation 2).

$$y = \frac{B_{max}x}{K_d + x}$$

*Equation 2 – One site saturation equation used to calculate $K_d$ values of inhibitors to PEF(S) where the change in fluorescence associated with the displacement of TNS (y) as a function of ligand concentration (x), $B_{max}$ the maximum at fluorescence at 0 M ligand.*

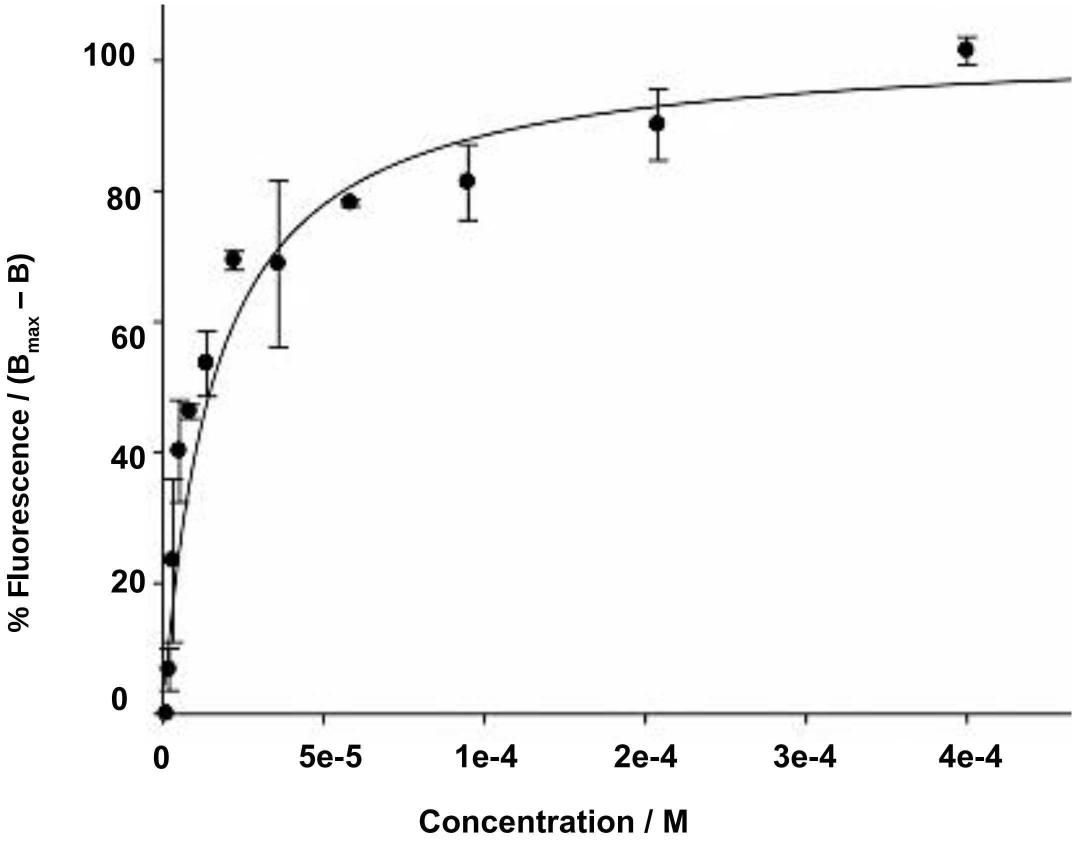


*Figure 3.2.3 - TNS displacement assay on PEF(S) with PD150606 (**9**) measuring the change in fluorescence at different concentrations of ligand, normalised and inverted the peak maxima then analysed by non-linear regression Ligand Binding module on Sigmaplot ((Systat Software, San Jose, CA).*

The calculated $K_d$ for PD150606 was 15.8 ± 2.9 μM, there is no published value however it is in accordance with the low micromolar concentrations required for the observed inhibitory activity.[117] The diselenides substitued with halogens at the 4-position were analysed by the same assay.

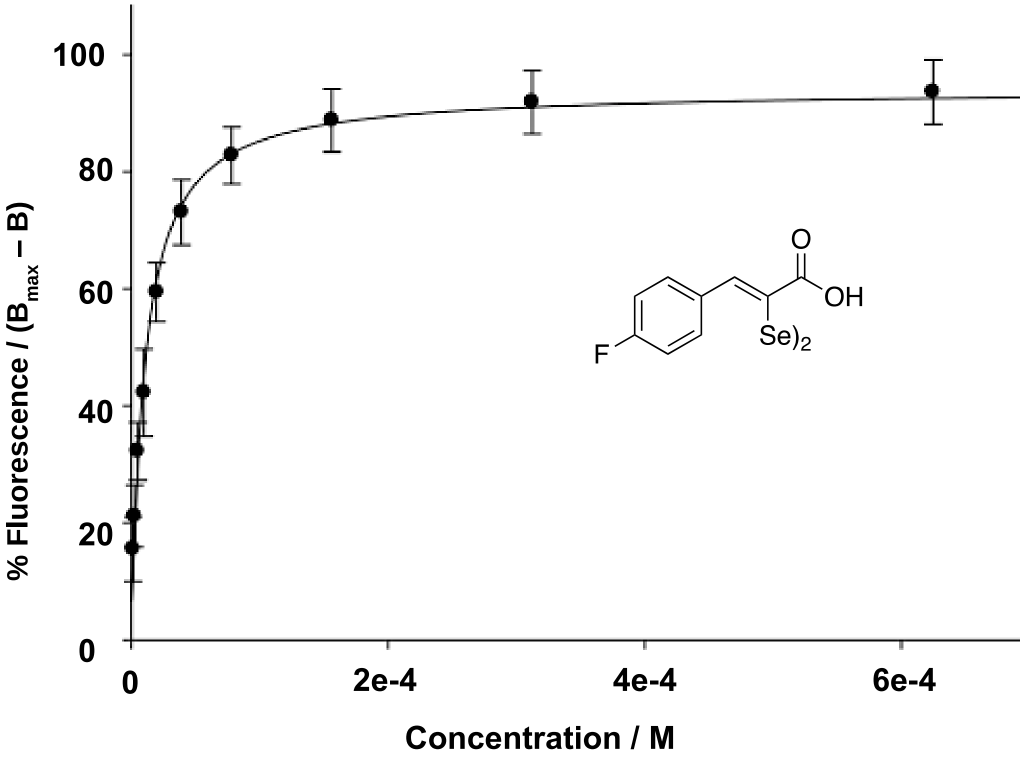


*Figure 3.2.4 - TNS displacement assay on PEF(S) with **35** measuring the change in fluorescence at different concentrations of ligand, normalised and inverted the peak maxima then analysed by non-linear regression Ligand Binding module on Sigmaplot (Systat Software, San Jose, CA).*

| **Compound** | **PEF(S) $K_d$ / μM** |
|---|---|
| **PD150606 - 9** | 15.8 ± 2.9 |
| **35** | 10.4 ± 0.9 |
| **36** | 14.5 ± 1.9 |
| **37** | 20.9 ± 4.3 |
| **38** | 21.0 ± 4.2 |

*Table 3.1 - TNS displacement assay results for $K_d$ values on human calpain-1 calcium binding domain, PEF(S) with compounds **35 – 38** and PD150606.*

In the case of the 4-X phenyl diselenides **35** - **38**, the smaller halogen (F, Cl) produced marginally tighter binding than the larger halogens (Br, I). Compared to **9**, all the phenyl compounds bound with a similar affinity to the α-mercaptoacrylic acid.

# 3.3. Crystallisation of PEF(S)

X-ray crystallography and cryo-electron microscopy (cryo-EM) provide unparalleled structural information of macromolecular drug interactions. Near-atomic resolution electron density maps (<1 Å) can elucidate bond lengths and angles which is valuable data when studying the mode of action of a pharmacophore. This information can be used to further improve binding potency *via* rational changes to the ligand structure.[175] As of 2018, more than 128,000 structures have been published in the Protein Data Bank (PDB), solved by X-ray crystallography, whilst over 12,000 by NMR spectrometry and 2000 by cryo-EM.[176] The technological advances of synchrotron radiation have enabled tunable high-intensity X-ray radiation, multi-wavelength anomalous diffraction and molecular replacement methods allow the phase problem of 3D X-ray crystallography to be readily tackled.[177] Significant developments on the vitrification process of cooling down a crystal quickly enough to inhibit the formation of a crystalline ice lattice, while preserving crystal quality for the duration of cryo-crystallographic data collection have in turn improved data quality.[176] These technological advances have fuelled the development of an incredibly powerful tool that provides unparalleled levels of atomic detail of macromolecular structures.

## 3.3.1. Plate setup

Successful conditions for the crystallisation of PEF(S) have been already reported, although it is common for proteins to form crystals in a variety of crystallisation conditions. The most widely used method of protein crystallisation is vapour diffusion, and of these two variants, sitting drop and hanging drop are popular among crystallographers. The sitting drop method situates the protein next to a larger well containing the precipitant solution, whereas the hanging drop situates the protein directly above the precipitant (Figure 3.3.1). The droplet is commonly a mixture of the protein (approx. 10 mg/mL) and precipitant solution in a 1:1 ratio, the precipitant in the well is more concentrated than in the drop, therefore in order to reach equilibrium, water vapour diffuses into the large well, slowly concentrating the protein and if the conditions are suitable, initiate crystallisation.

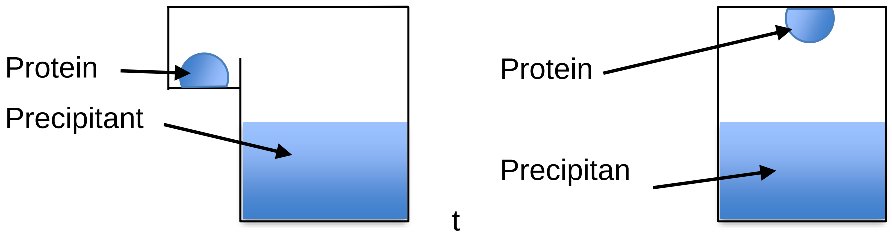


*Figure 3.3.1 - Diagram of sitting drop (left) and hanging drop (right) protein crystallography vapour diffusion techniques.*

The published conditions for crystallisation of PEF(S) are 20 mM $CaCl_2$, 50 mM sodium cacodylate ($(CH_3)_2AsO_2Na$) and 12.5% (w/v) PEG 6000 at pH 7.0, although variation between protein preparations can mean narrow screening around these conditions is advisable.[48] The wells were filled with 1 mL of precipitant and protein droplets set up containing 3 μL precipitant, 1 μL deionized water and 3 μL of 10 mg/mL PEF(S) solution in crystallography buffer (Section 6.7.4) in hanging drop wells. The wells were then sealed and incubated at 20 °C, crystals formed after 2-3 weeks.

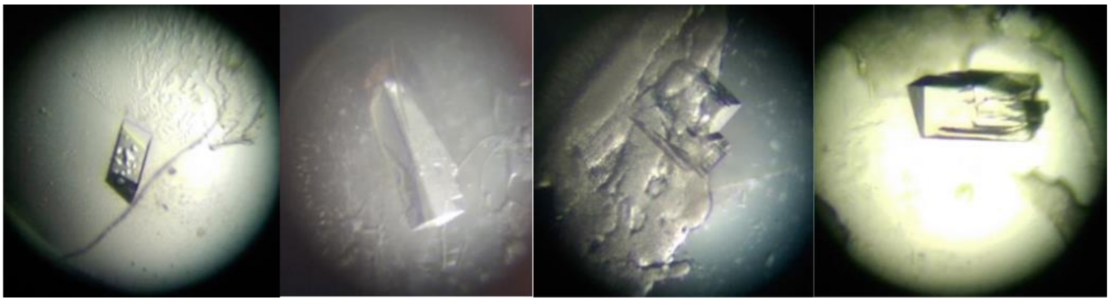


*Figure 3.3.2 - Crystals of PEF(S) in the published conditions (Section 6.7.4) harvested for single crystal X-ray diffraction at Diamond Light Source, Oxford, UK.*[128]

The crystals were cryoprotected with ethylene glycol (20% droplet volume) and harvested using a cryoloop, flash frozen in liquid nitrogen to preserve the crystals and allow data collection at cryogenic temperatures, reducing radiation damage from the high power X-ray source.

Data was collected at Diamond Light Source, the diffraction pattern was recorded up to 1.19 Å, which was reduced and scaled using *xia2.*[178] Molecular replacement was carried out on the mtz file using Phaser-MR from the CCP4 package using the PEF(S) structure (PDB:4PHJ) from the Protein Data Bank.[179]

Molecular replacement with Phaser-MR uses a known structure to phase the data of an unknown crystal dataset, using rigid body refinement to position the model in the density map, avoiding clashes whilst obeying the symmetry operations.[179] A log likelihood score is given based on how well the model fits the data, the most likely outcome from the program produces a PDB file containing the atom coordinates and the resultant phasing information which is used to calculate the observed and difference electron density maps.[180] The difference density maps are used to adjust the position of residues and small molecules within the model manually to further improve the statistics whilst refinement with REFMAC can produce even finer adjustments.[181]

Molecular replacement was carried out on the dataset using Phaser-MR and the 1.6 Å calcium-bound human PEF(S) structure (PDB:4PHJ) as a search model.[48] Iterations of refinement with COOT (Computational Object-Orientated Toolkit) and REFMAC5 were necessary to determine the final structure for the unliganded structure and whether the protein preparation used resulted in the same reported structure.[181,182] The unit cell size and space group was the similar to PDB:4PHJ although a higher resolution dataset was obtained for this crystal (1.19 Å). Several rounds of refinement with COOT and REFMAC5 were carried out until the PEF(S) model and electron density map were in agreement and statistics could not be improved further.[181,182]

| **Data Reduction Statistics** | **Human-PEF(S)** |
|---|---|
| **Unit cell dimensions / Å** | a = 49.31, b = 78.83, c = 56.49 |
| **Unit cell angles / °** | α = 90.00, β = 90.97, γ = 90.00 |
| **Space group** | $P\ 12_11$ |
| **Resolution range / Å** | 37.47 – 1.19 |
| **Total reflections** | 331824 |
| **Unique reflections** | 93245 |
| **Completeness / % (last shell)** | 78.19 (21.32) |
| **I / σ (last shell)** | 12.5 (1.0) |
| **R(merge) / % (last shell)** | 3.8 |
| **B(iso) from Wilson / Å$^2$** | 22.19 |
| **Refinement Statistics** | |
| **Protein atoms excluding H** | 3110 |
| **Calcium atoms** | 8 |
| **R factor / %** | 15.3 |
| **$R_{free}$ / %** | 18.8 |
| **R.m.s. bond length / Å** | 0.028 |
| **R.m.s. bond angle / °** | 2.387 |

*Table 3.2 - Data reduction and refinement statistics for recombinant human PEF(S) unliganded crystal X-ray diffraction dataset.*

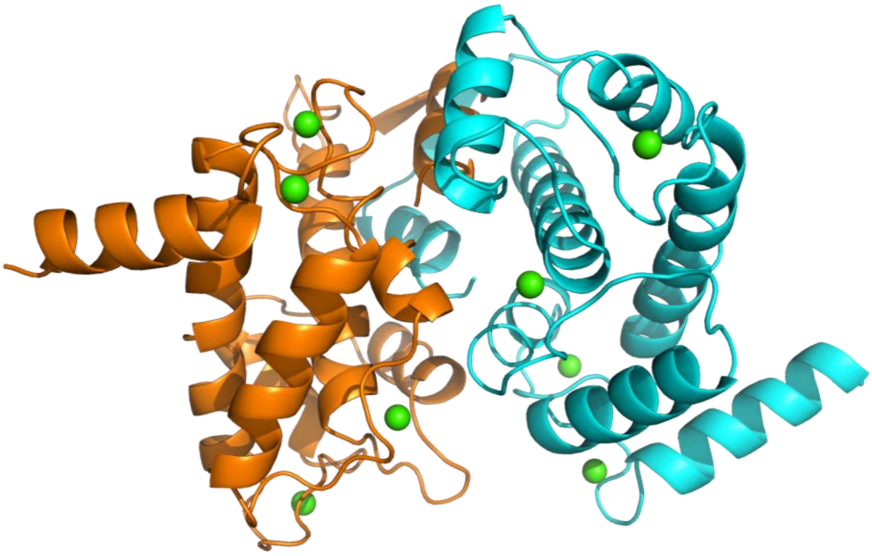


*Figure 3.3.3 - Recombinant human PEF(S) crystal structure chain A and B represented in cartoon form as cyan and orange respectively, 8 bound calcium ions represented as green spheres.*

The PEF(S) protein crystallises in a homodimeric state, hydrophobic residues between the subunits stabilises the formation of the dimer, and the EF-handshake between EFH-5 moieties form the dimer interface for this protein (Figure 3.3.4).

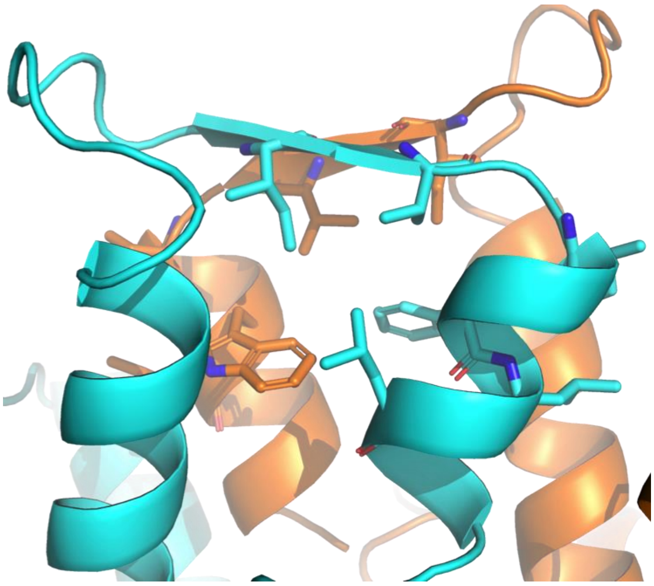
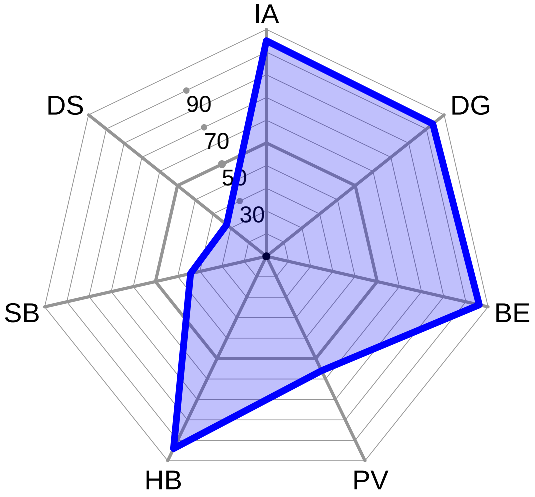


*Figure 3.3.4 - Hydrophobic residues around EFH-5 between chain A (cyan) and chain B (orange) highlighted with stick notations.*



| Parameter | Value |
|---|---|
| IA: Interface Area | 2126.7 Å$^2$ |
| DG: ΔG | -21.6 kcal/mol |
| BE: Binding Energy | -33.7 kcal/mol |
| PV: P-value | 0.1305 |
| HB: Hydrogen bonds | 24 |
| SB: Salt bridges | 4 |
| DS: Disulfide bonds | 0 |

*Figure 3.3.5 - QtPISA analysis of the interface between chain A and B of human PEF(S) crystal structure, the polygon radar (blue) shows the likelihood of each parameter contributing to the formation of the macromolecular assembly (larger area is greater contribution).*

Macromolecular assemblies are complex and abundant in the cellular environment. Functionality is closely related to the 3D structure of protein complexes and thus useful information can be determined by examining these interfaces. PISA (Protein Interactions, Surfaces and Assemblies) is a computational tool as part of the CCP4 toolkit that can calculate parameters associated with the interface of macromolecular assemblies. PISA was used to explore the properties of the unliganded PEF(S) structure (Figure 3.3.5).

The interface in between the two monomers has a total area of 2126.7 Å$^2$ which is relatively large compared to the calculated average interface area (IA) of 1311 Å$^2$ of 6545 protein complexes in the PDB.[183] Hydrogen bonds are the main bonds formed between the PEF(S) subunits compared to only 4 salt bridges and no disulfide bonds. The P-value is a measure of interface specificity indicating the probability of finding an area of the protein surface which has greater hydrophobicity, in this case the low value suggests the interaction is specific and not randomly associated.[184] The binding energy, and negative Gibbs free energy, based upon solvation energy gained upon formation of the complex, suggests that this assembly is an energetically favourable dimer.

The coordination spheres around the $Ca^{2+}$ ions were well resolved in the electron density map (Figure 3.3.6) with calcium ions bearing both one (**A** - Figure 3.3.6) and two water molecules (**B** - Figure 3.3.6) readily visibile.

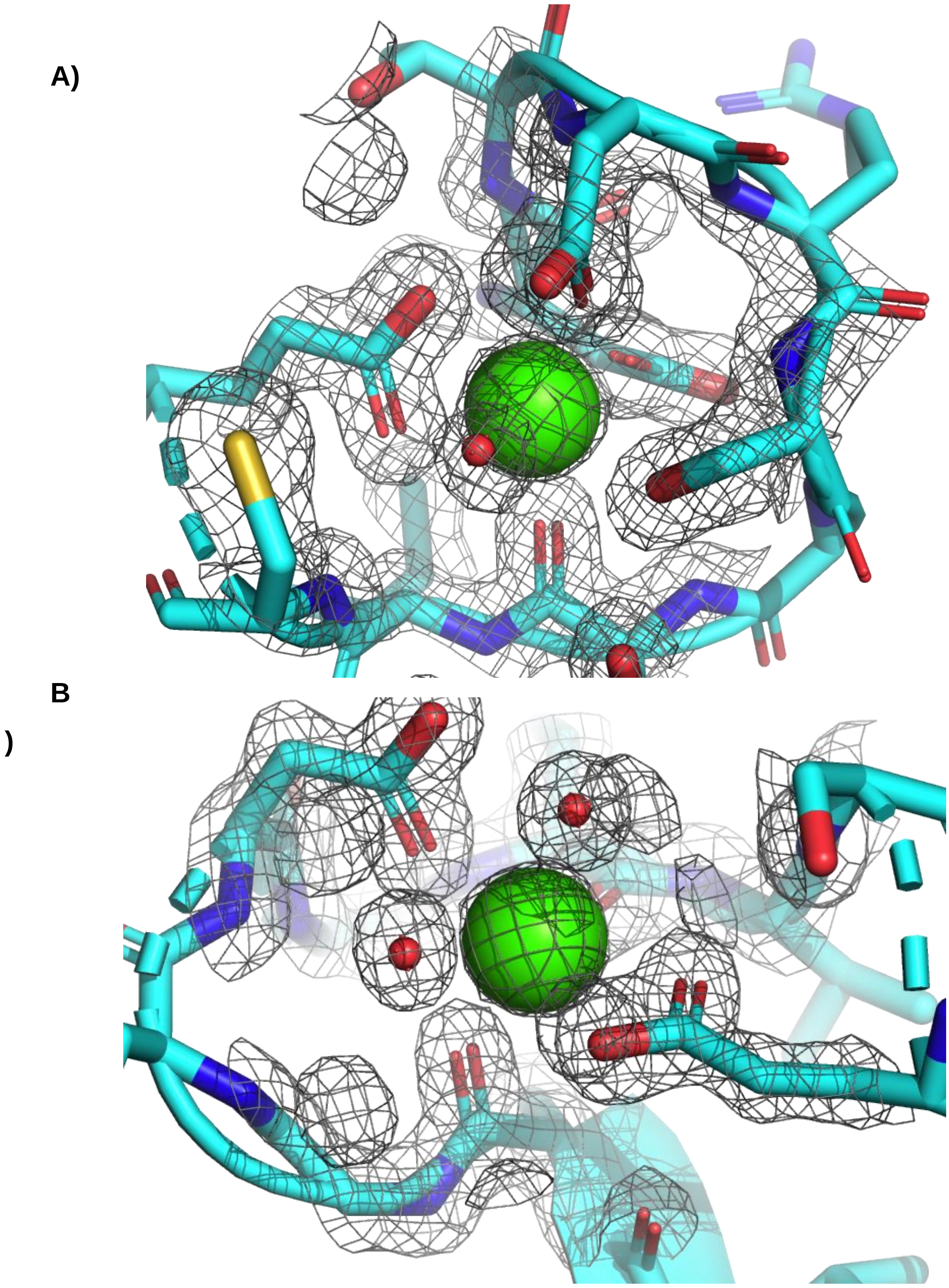


*Figure 3.3.6 - A) Electron density around EFH-3 with calcium and 1 water molecule coordinated bound. B) Calcium bound EFH-1 motif with 2 water molecules coordinated. Density map mesh contoured to 1.5 σ, residues represented as sticks, calcium and water represented as green and red spheres respectively.*

### 3.3.2. Diselenide **35** interaction with PEF(S)

Crystals of PEF(S) were soaked with solutions containing diselenides in order to compare the geometry of the binding with the α-mercaptoacrylic acids. Prior to harvesting the protein crystals, 0.5 µL of diselenide inhibitor (5 mM) diluted from 50 mM DMSO stock with crystallography precipitant (50 mM sodium cacodylate, 12.5% (w/v) PEG 6000, 20 mM $CaCl_2$ at pH 7.4) was added to the droplets containing crystals. After 24 hours these were harvested and cryo-cooled in liquid nitrogen as described (Section 6.7.4).[48]

Crystallographic information files (CIF) for the inhibitors were created using JLigand graphically in 3D for the molecules, using geometric restraints such as bond length and torsion angles.[185] The structures were then fitted into the electron density map to determine if the observed difference density matched the predicted structure of the small molecule for each soaking experiment.

| PEF(S) Structure | Resolution / Å | Space Group |
|---|---|---|
| **Unliganded** | 1.19 – 37.46 | $P\ 12_11$ |
| **35** | 1.58 - 46.36 | $P\ 12_11$ |
| **36** | 1.55 – 79.45 | $P\ 12_11$ |
| **41** | 1.45 – 56.71 | $P\ 12_11$ |

*Table 3.3 - Table of crystallography data from PEF(S) crystals soaked with diselenide inhibitors.*

All the datasets shown for the soaked compounds displayed a lower resolution than the unliganded structure, indicating that the crystals were likely affected by the soaking process, damaging the original crystal, resulting in a lower resolution (Table 3.3). Phaser-MR was used for molecular replacement on the reduced mtz files, using the human PEF(S) structure (PDB:4PHJ), followed by iterations of manual and automated refinement with COOT and REFMAC5 tools respectively.[181,182]

A datatset was recorded to 1.59 Å resolution for PEF(S) soaked with diselenide **35**, During refinement of the protein structure without a ligand, areas of positive density within the hydrophobic pockets on both copies of PEF(S) in the asymmetric unit became apparent. The 3D structure for **35** was drawn in JLigand and geometric constraints applied through regularisation.[185] The density was accountable for that predicted for the ligand.

| **Data Reduction Statistics** | **Human-PEF(S) soaked with 35** |
|---|---|
| **Unit cell dimensions / Å** | a = 49.65 b = 79.81 c = 56.95 |
| **Unit cell angles / °** | $\alpha = 90.00$, $\beta = 90.89$, $\gamma = 90.00$ |
| **Space group** | $P\ 12_11$ |
| **Resolution range / Å** | 46.36 – 1.58 |
| **Total reflections** | 211742 |
| **Unique reflections** | 59135 |
| **Completeness / % (last shell)** | 97.31 (96.67) |
| **I / σ (last shell)** | 11.9 (1.3) |
| **R(merge) / % (last shell)** | 5.1 (78.1) |
| **B(iso) from Wilson / Å$^2$** | 33.80 |
| **Refinement Statistics** | |
| **Protein atoms excluding H** | 3452 |
| **Calcium atoms** | 8 |
| **R factor / %** | 16.91 |
| **$R_{free}$ / %** | 20.76 |
| **R.m.s. bond length / Å** | 0.017 |
| **R.m.s. bond angle / °** | 1.971 |

*Table 3.4 - Data reduction and refinement statistics for recombinant human PEF(S)-**35** co-crystal structure.*

Multiple conformations were observed for the ligand within the pocket which were modelled in to each to satisfy the difference density in each monomer. Multiple conformations have been reported for disulfide PEF(S) soak experiments.[128] The location of the first aromatic ring of the inhibitor remains conserved in the deep pocket whilst the stereo-electronic constraints of the diselenide bond limits the conformation of second the aromatic ring in the hydrophobic groove.[128]

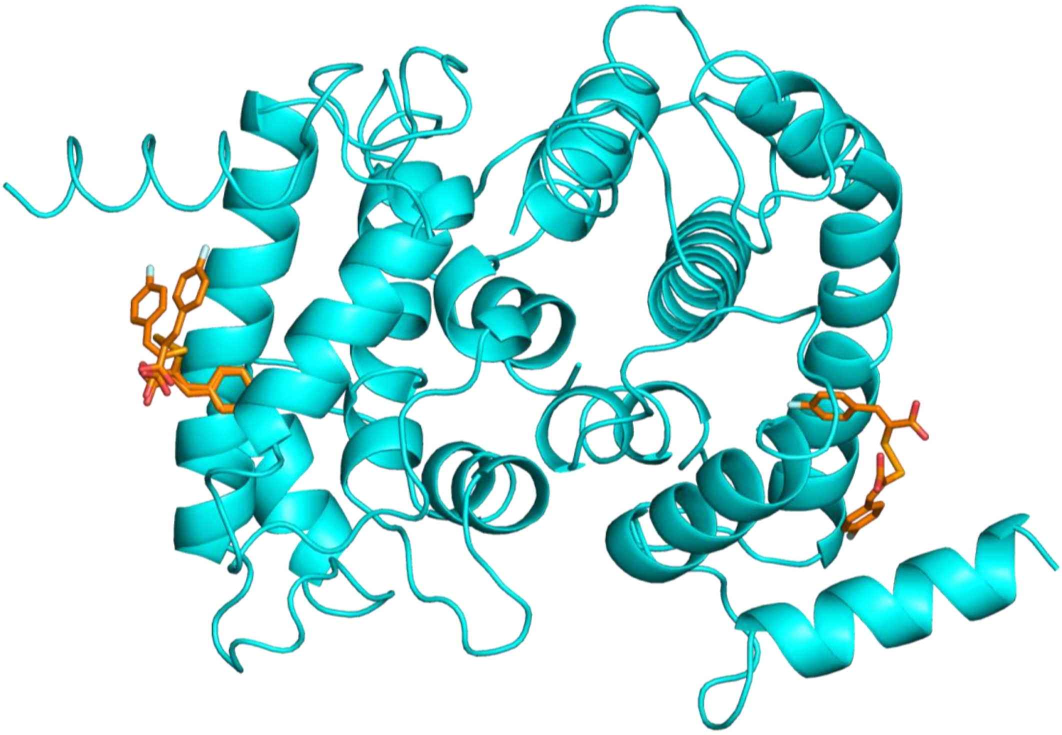


*Figure 3.3.7 - The asymmetric unit of PEF(S)-**35** crystal structure, the protein shown in cartoon representation (cyan) and the diselenide **35** as sticks (orange).*

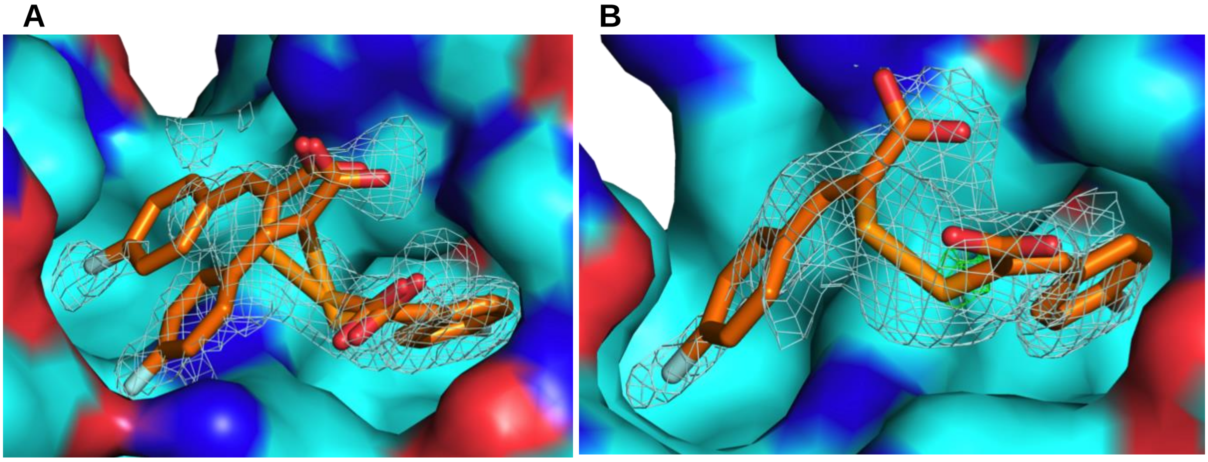


*Figure 3.3.8 - The observed conformations for the ligand **35** (orange sticks), on PEF(S) surface (cyan) chain A (left, **A**) chain B (right, **B**). Map isomesh (pale cyan) contoured to 1.0 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red).*

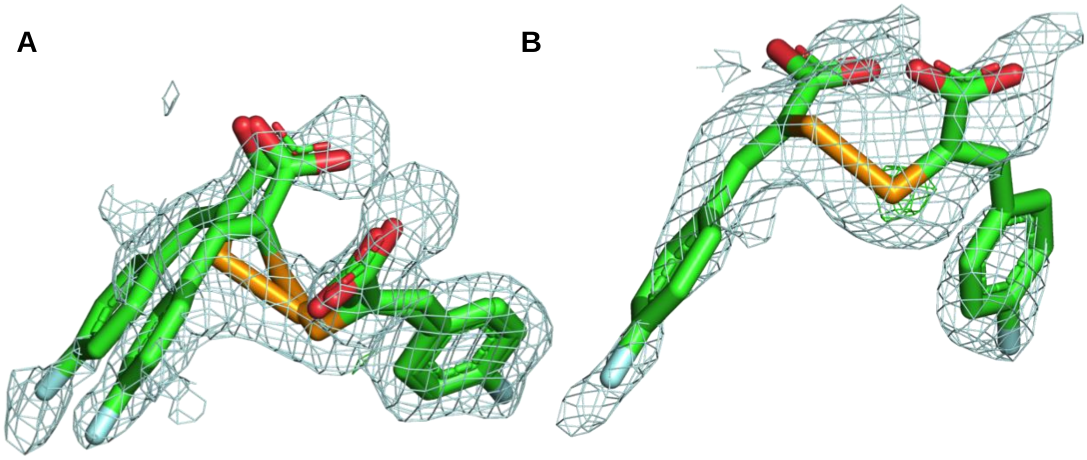


*Figure 3.3.9 - The observed conformations for the ligand* ***35*** *(green sticks), on PEF(S) chain A (left,* ***A****) chain B (right,* ***B****). Map isomesh (pale cyan) contoured to 1.0 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red).*

The bound diselenide **35** was arranged in a similar conformation to the oxidised α-mercaptoacrylic acid PD150606 in the hydrophobic PEF(S) binding site (Figure 3.3.10). When viewed with the protein surface (Figure 3.3.8) it was evident that there is flexibility for the ligand to move around the protein surface. Compared to the crystal structure of disulfide form of PD150606 (**9**) bound to PEF(S), both the diselenide and disulfide geometries link the two aromatic rings effectively, enabling further interaction with the two adjacent hydrophobic grooves. However, restricted flexibility around this dihedral bond allows for multiple conformations to be possible in the binding pocket.

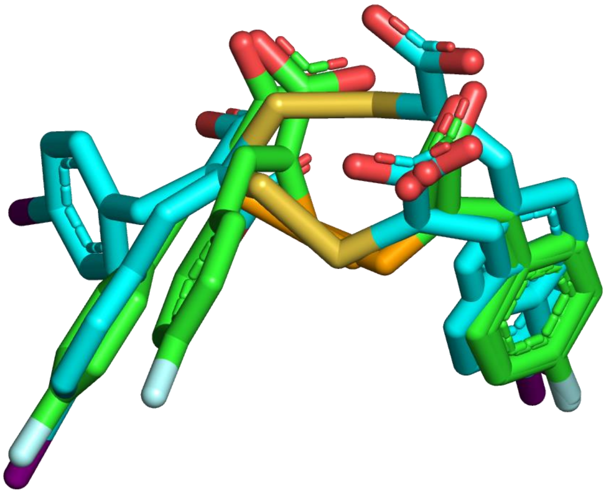


*Figure 3.3.10 - Alignment of PEF(S) chain A in PyMOL with diselenide* ***35*** *dataset represented as green sticks and PDB:5D69 containing PD150606 in disulfide form, shown as cyan sticks.*

Multiple inhibitor conformations of both the disulfide and diselenide ligands suggest that binding of this class of ligand is not optimal in the hydrophobic pocket and relying on several weaker Van der Waals interactions for binding.

### 3.3.3. Diselenide **36** interaction with PEF(S)

The crystal of PEF(S) was soaked with diselenide **36,** containing a chlorine substituent the 4 position, using the same method as described for diselenide **35** (Section 3.3). The determined structural model contained positive density accountable for diselenide **36** in the binding site on both PEF(S) monomers in the asymmetric unit.

| **Data Reduction Statistics** | **Human-PEF(S) soaked with 36** |
|---|---|
| **Unit cell dimensions / Å** | a = 49.64, b = 79.44, c = 56.93 |
| **Unit cell angles / °** | $\alpha$ = 90.00, $\beta$ = 91.48, $\gamma$ = 90,00 |
| **Space group** | $P\ 12_11$ |
| **Resolution range / Å** | 46.36 – 1.55 |
| **Total reflections** | 274845 |
| **Unique reflections** | 74215 |
| **Completeness / % (last shell)** | 97.21 (96.29) |
| **I / σ (last shell)** | 12.4 (1.2) |
| **R(merge) / % (last shell)** | 4.8 (88.4) |
| **B(iso) from Wilson / Å$^2$** | 37.63 |
| **Refinement Statistics** | |
| **Protein atoms excluding H** | 3119 |
| **Calcium atoms** | 8 |
| **R factor / %** | 18.13 |
| **$R_{free}$ / %** | 22.79 |
| **R.m.s. bond length / Å** | 0.0143 |
| **R.m.s. bond angle / °** | 2.189 |

*Table 3.5 - Data reduction and refinement statistics for recombinant human PEF(S)-**35** co-crystal structure.*

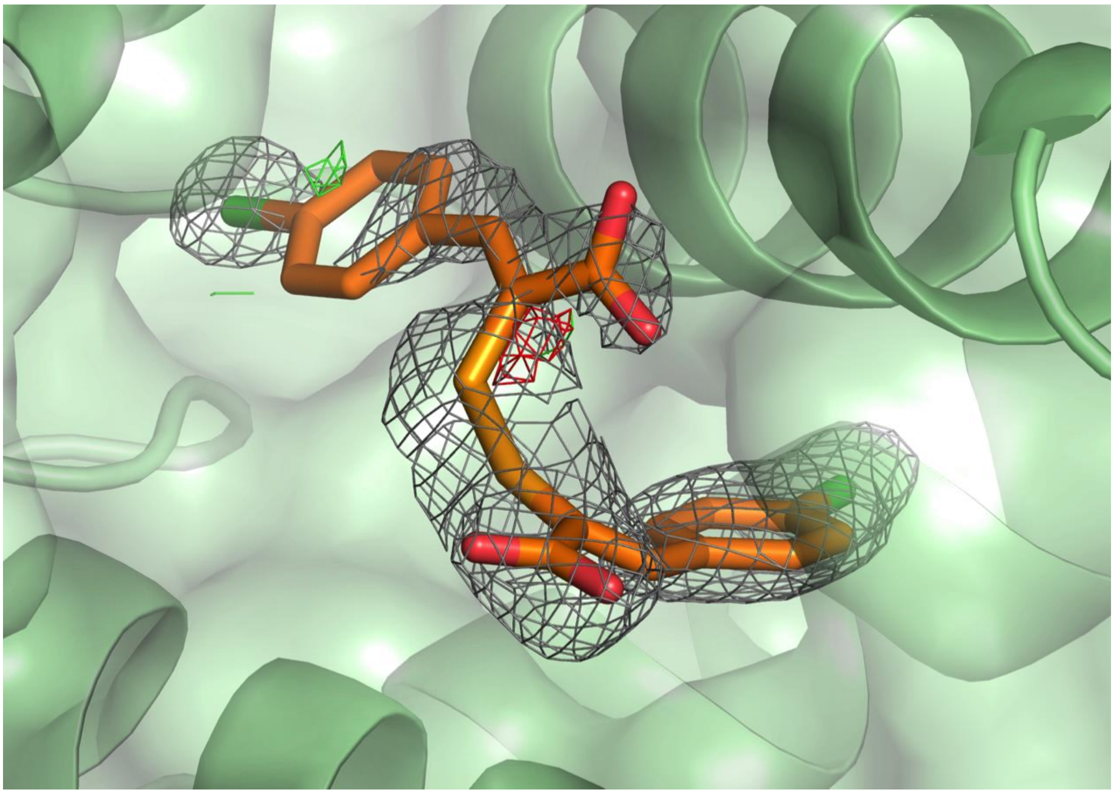


*Figure 3.3.11 - Cartoon represnetation of the electron density around diselenide **36** (orange) in the hydrophobic pocket on chain A PEF(S), protein surface (50% transparent, green, with cartoon secondary structure), Map isomesh (grey) contoured to 1.0 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red).*

The interface between **36** and the PEF(S) hydrophobic pocket were analysed by QtPISA.[186] No hydrogen bonds were formed between the protein and **36**, the primary interactions driving the binding were extensive hydrophobic interactions between the aromatic portion of the diselenide, and the hydrophobic aromatic and alkyl residues.[186]

PEF(S) contains two flexible residues - Arg128 and Gln173 - around the hydrophobic pocket that move substantially upon binding to PEF(S) (Figure 3.3.12), while QtPISA analysis did not highlight these as hydrogen bonded, the flexible nature of a sidechain in solution suggests that it is possible for these residues may play a part in the binding interactions.  Leu104 is a hydrophobic residue on the edge of the PEF(S) binding pocket which can be displaced by the aromatic ring of **36** upon binding.

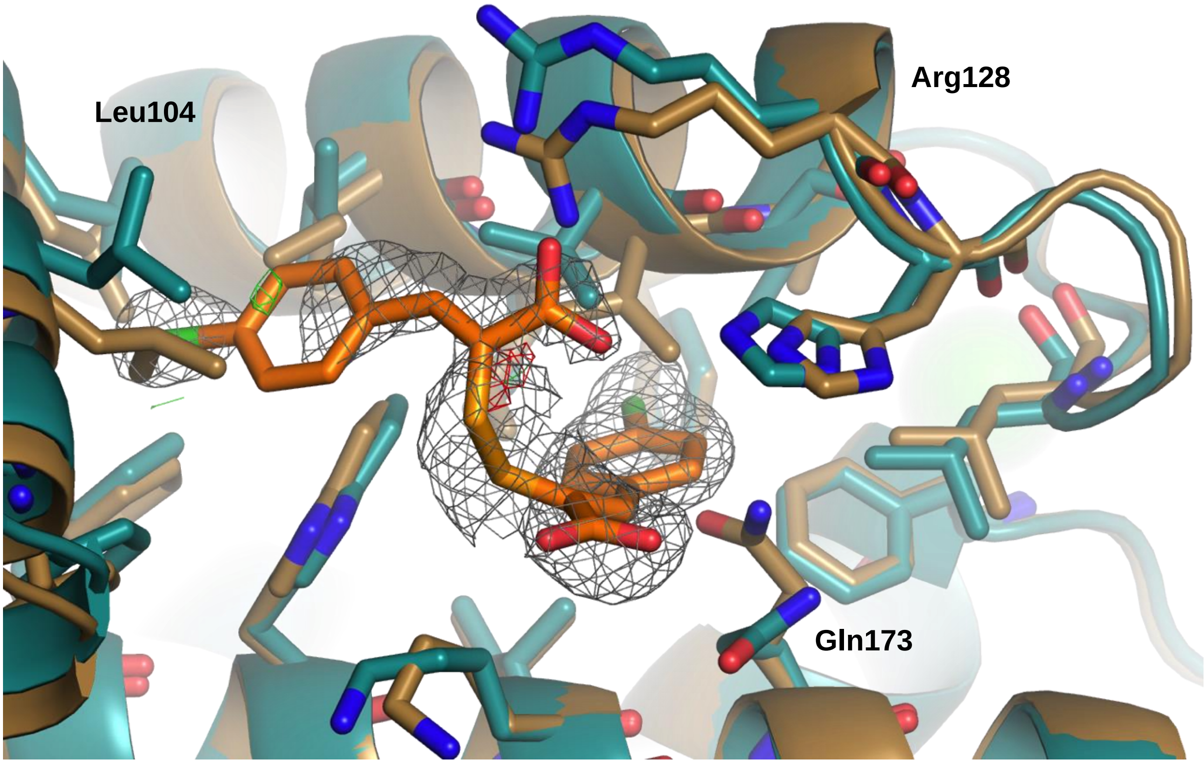


*Figure 3.3.12 - Aligned secondary structure cartoon representations of human PEF(S) (Unliganded, brown) and the complex of* ***36*** *with PEF(S) (****36****, orange, PEF(S) green), Arg128, Gln173 and Leu104 show the greatest deviations upon ligand binding. Map isomesh (grey) contoured to 1.0 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red).*

The phenyl diselenides **35** and **36** displayed similar binding geometries (

Figure 3.3.13), the location of the diselenide in the deep pocket of PEF(S) (Section 1.5.5) is highly conserved between both inhibitors, while the second half of inhibitor displays more flexibility. The diselenide bond is the only group that allows for flexbility in the molecule, however is limited by the stereo-electronic constrains of the dihedral bond.

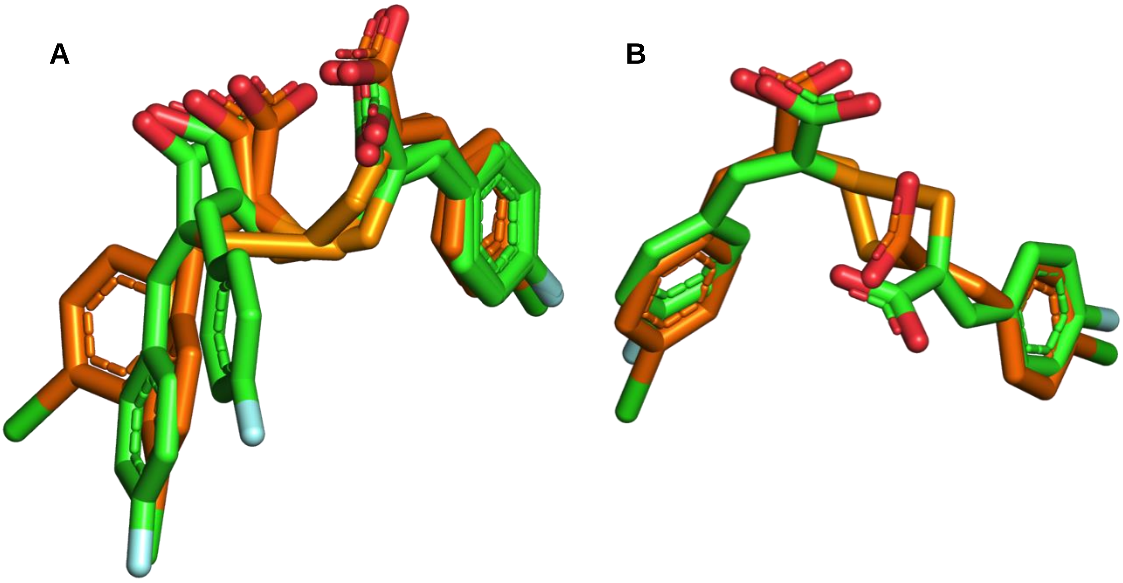


*Figure 3.3.13 – A cartoon representation of PEF(S)-**35** (cyan) and PEF(S)-**36** (green) ligands on chain A (**A**) and chain B (**B**) aligned in PyMol.*

The dihedral angle of the diselenide **36** was compared to the disulfide form of PD150606 **9** bound in PEF(S) (Figure 3.3.14) (PDB:5D69). The diselenide bond was marginally longer (2.3 Å) than the disulfide bond (2.1 Å). The C-Se-Se bond angle was larger than the C-S-S bond, in diphenyl diselenide the reported C-Se-Se and C-Se-Se-C bond lengths are ~ 110° and ~ 82° respectively which is similar to the diselenide bond angles observed in PEF(S)-**36**.

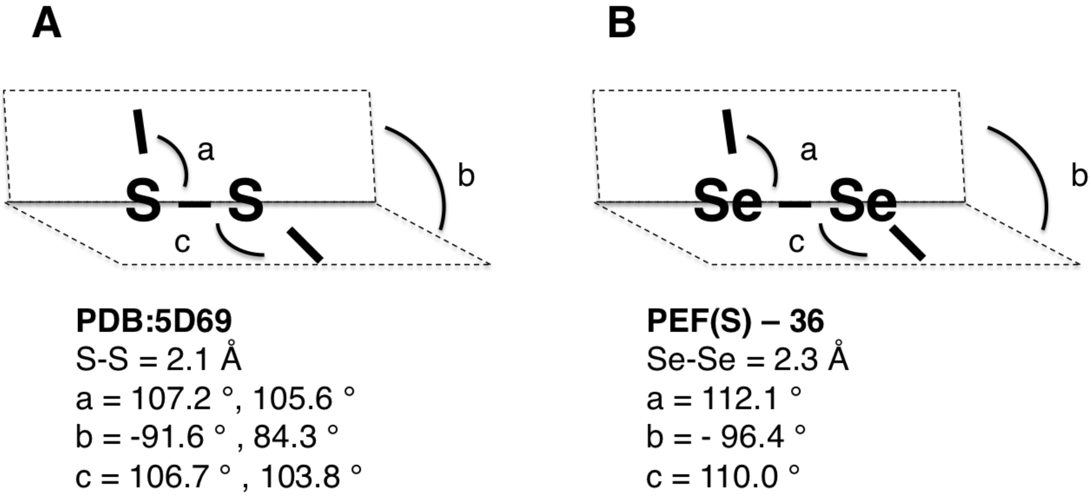


*Figure 3.3.14 - Comparison of the observed geometries of the disulfide and diselenide dihedral angles, taken from chain B of PDB:5D69 and the PEF(S)-**36** crystal structures as example disulfide and diselenide bonds, analysed in PyMol using the dihedral angle tool. Two angles are listed for both reported conformations of PD150606 disulfide in chain B of PDB:5D69.*

### 3.3.4. Diselenide **41** interaction with PEF(S)

The indole based diselenide (**41)** has a larger aromatic ring than the phenyl based ligands. It was soaked into a crystal of human PEF(S) which were harvested and data collected as described as for PEF(S)-**35** (Section 3.3).

| **Data Reduction Statistics** | **Human-PEF(S) soaked with 41** |
|---|---|
| **Unit cell dimensions / Å** | a = 49.26, b = 78.74, c = 56.72 |
| **Unit cell angles / °** | α = 90.00, β = 91.29, γ = 90.00 |
| **Space group** | $P12_11$ |
| **Resolution range / Å** | 56.71 – 1.45 |
| **Total reflections** | 226302 |
| **Unique reflections** | 61617 |
| **Completeness / % (last shell)** | 98.60 (99.97) |
| **I / σ (last shell)** | 15.1 (1.5) |
| **R(merge) / % (last shell)** | 4.3 |
| **B(iso) from Wilson / Å²** | 38.12 |
| **Refinement Statistics** | |
| **Protein atoms excluding H** | 3476 |
| **Calcium atoms** | 8 |
| **R factor / %** | 20.90 |
| **$R_{free}$ / %** | 26.49 |
| **R.m.s. bond length / Å** | 0.010 |
| **R.m.s. bond angle / °** | 1.830 |

*Table 3.6 - Data reduction and refinement statistics for recombinant human PEF(S)-**41** co-crystal structure.*

The protein structure was determined through molecular replacement using PDB entry 4PHJ, difference density was observed in the hydrophobic pocket which was modelled in as **41** using JLigand to create a 3D model of the molecule using geometric restraints.[185] This dataset displayed less ligand density than the other diselenide co-crystal structures, however refinement without a ligand in the pocket produced positive density inside the binding site, implying partial occupancy of a molecule, which fit the geometry of the **41**. The lack of difference density implies this is a correct interpretation of of the observed density, improvement on this structure may provide further information on the binding nature. Further refinement of the soaking conditions by increasing the ligand concentration, duration of the soak or alterations to the cryoprotection process may increase the occupancy of the ligand.

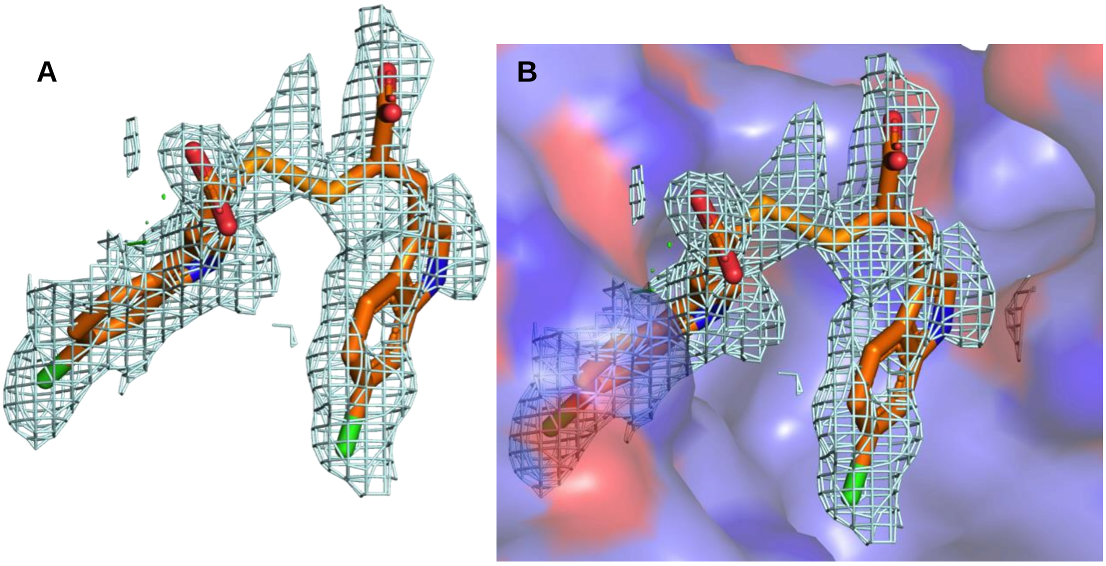
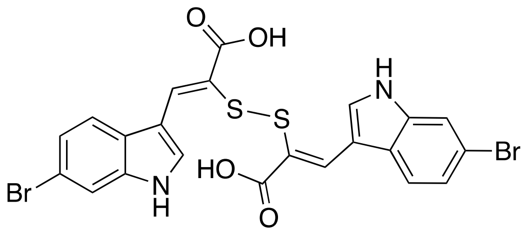
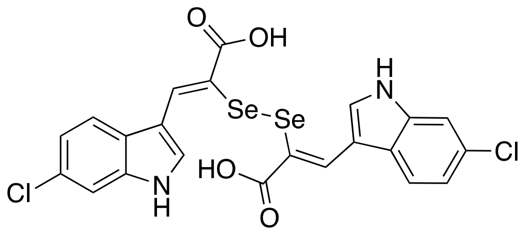


*Figure 3.3.15 – **A)** - Electron density around **41** in the PEF(S) hydrophobic pocket, isomesh (grey) contoured to 0.5 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red). **B**) Electron density of inhibitor as **A**, with surface of PEF(S) shown at 40% transparency (purple).*

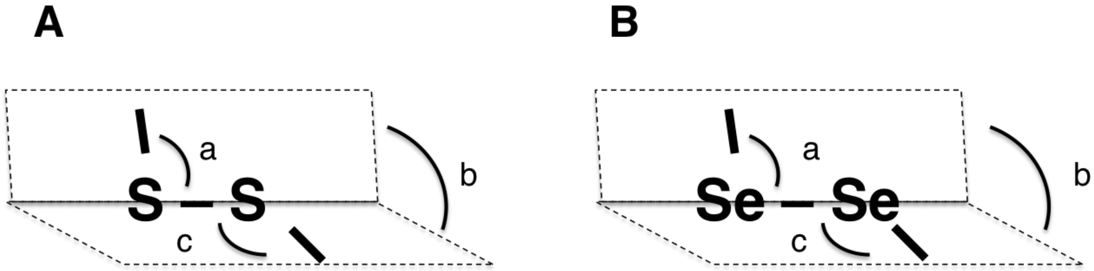
The dihedral angle of the diselenide **41** was compared to the indolde based disulfide inhibitor, **44,** bound in PEF(S) (Figure 3.3.14) (PDB:4WQ3). The diselenide bond is marginally longer (2.3 Å) than the disulfide bond (2.1 Å). The dihedral angles observed for these indole based inhibitors display larger dihedral angles than the phenyl based disulfide and diselenides. This could be due to the steric bulk of the indole rings enforcing a different conformation by engaging with the protein surface residues, resulting in a larger dihedral angle.



**44 –PDB:4WQ3**



**41**



**PDB:4WQ3**
S-S = 2.1 Å
a = 109.8 °, 110.1 °
b = 128.4 ° , 169.4 °
c = 111.5 ° , 103.3 °

**PEF(S) – 41**
Se-Se = 2.3 Å
a = 107.8 °
b = 154.5 °
c = 96.7 °

*Figure 3.3.16 - Comparison between the observed geometries of the indole based disulfide **44** and diselenide dihedral angle, taken from chain A of PDB:4WQ3 and the PEF(S)-**41** crystal structures as example disulfide and diselenide bonds, analysed in PyMol using the dihedral angle tool. Two angles are listed for both reported conformations of disulfide **44** in chain A of PDB:4WQ3.*

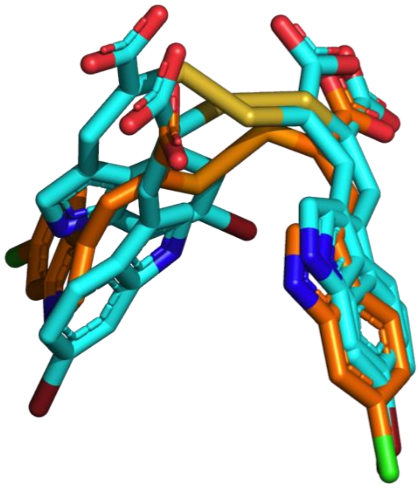


*Figure 3.3.17 - Cartoon representation of diselenide **41** (orange sticks) and disulfide **44** (cyan sticks, PDB:4WQ3) bound to PEF(S), aligned in PyMol.*

The aligned diselenide **41** and disulfide **44** shows that, as with the phenyl based ligands, the first aromatic ring in the deep pocket position is conserved whilst multiple conformations are possible for the second ring through different orientations of the dihedral bond.

### 3.3.5. Concluding remarks on PEF(S) interactions

The diselenides synthesised in this project bind to the allosteric site on the calpain regulatory subunit calcium binding domain, PEF(S), with a similar affinity to the original PD150606 inhibitor **9**. The TNS displacement assay was developed beyond the previous scope of qualitatively examining PEF(S) binding, by measuring fluorescence over an extended range of ligand concentration with enough datapoints to quanitfy $K_d$ values by non-linear regression calculations.

Single crystal X-ray diffraction data aided in the visualisation of the binding interaction between the novel diselenide inhibitors and the PEF(S) site, exhibiting binding primarily *via* hydrophobic interactions between the aromatic ring and hydrophobic residues surrounding the pocket. The multiple observed ligand conformations suggests that the diselenide bond geometry is analagous to the disulfide bond. While the dihedral bond is an important aspect of the binding of disulfides and diselenides to PEF(S), the observed flexibility and multiple conformations suggests the binding is sub optimal.

## 3.4. Small Molecule Interaction with PEF(L)

A decrease in TNS fluorescence upon the addition of PD150606 (**9**) to PEF(L) suggested that binding occurs, however more detailed experiments were not carried out to determine the $K_d$ values.[117] Small molecule inhibition *via* this site could be appealing for drug development programmes seeking selective calpain-1 allosteric inhibition. α-Mercaptoacrylic acids are moderately selective between the isoforms despite the fact PEF(S) is identical between calpain-1 and -2, therefore an interaction with PEF(L) has been postulated to explain the difference in potency between the isoforms.[187] Ultimately, binding to a site that is entirely unique to the large subunit of calpain-1 would be preferable for isoform specific inhibition.[117,118] Inhibiting *via* an allosteric isoform specific domain such as PEF(L) would also be an attractive option over reactive warhead inhibitors that target the active site, allosteric binding could be tailored to modulate the activity of the enzyme *via* conformational changes without permanently altering the chemistry of the active site residues.[56] The expression and purification of PEF(L) was carried out to determine if the diselenides bind to this domain.

### 3.4.1. Expression of PEF(L)

An expression construct for PEF(L) was previously prepared in the Allemann group by S. Adams containing PEF(L) ligated in a pET-21d(+) vector, deposited in the plasmid libarary, entry 1121. The gene codes for a protein that is a total of 197 residues, containing an *N*-terminal polyhistidine tag, a short linker and PEF(L), excised from the full length calpain-1 sequence from residues 539 - 714.

PEF(L) was grown from a frozen glycerol stock of BL21 (DE3) CodonPlus RP cells containing the plasmid, in ampicillin selective LB media overnight at 37 °C. 10 mL of this culture was then grown in 1 L of TB ‘Enhanced’ media until $OD_{600}$ = 0.8-1.0, IPTG (238 mg) was added to reach a final concentration of 1 mM and the culture was grown for 20 hours at 20 °C, the cells were then harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C.

The cells were re-suspended in 20 mM HEPES, 100 mM NaCl, 0.5 mM TCEP and 5% glycerol pH 7.6 (buffer A) and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 RCF in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 µm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 CV buffer A, 15 CV buffer A containing 20 mM imidazole and eluted with 10 CV buffer A containing 250 mM imidazole. The fractions were analysed by SDS-PAGE.

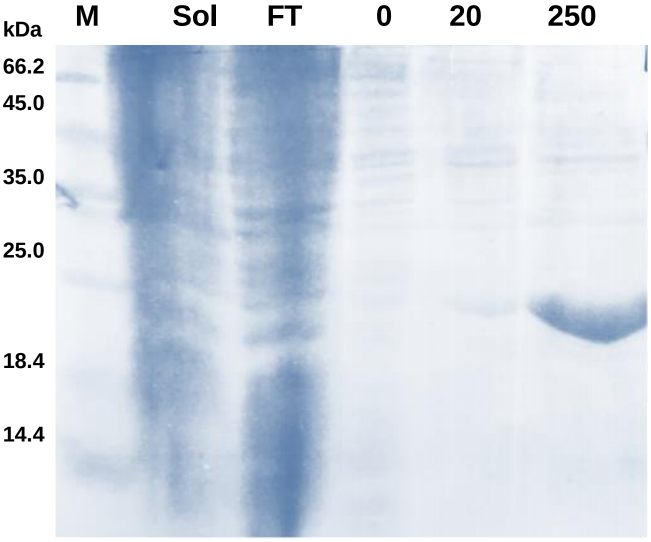


*Figure 3.4.1 - SDS-Polyacrylamide gel of Ni-NTA purification of recombinant human PEF(L). Where M is the protein ladder, Sol is the soluble fraction, FT is the flow-through and 0, 20 and 250 mM imidazole fractions.*

The protein was initially soluble however appeared to be unstable in solution, forming aggregates hours after purification at 4 °C. Attempts at improving the solubility of PEF(L) by the addition glycerol (5%), which can stabilise protein in solution by providing an amphiphilic interface between exposed hydrophobic regions of a protein surface and the polar aqueous solvent, was found to slow the aggregation of this PEF(L) construct.[188]

## 3.4.2. Anion exchange chromatography

A Resource™Q column (7 mL, GE Healthcare Life Sciences, Buckinghamshire, UK) was equilibrated with PEF(L) Buffer C (Section 6.6.6) at a flow rate of 3 mL / min using a Bio-Rad NGC FLPC instrument (California, USA). PEF(L) elution from Ni-NTA column was diluted to 50 mM NaCl using PEF(L) Buffer C, Section 6.6.6) and loaded to the column. A gradient of 0 to 0.5 M NaCl was run over 10 CV with 2 mL fractions collected whilst monitoring the UV-absorbance at 280 nm. The peak fractions containing protein were analysed by SDS-PAGE.

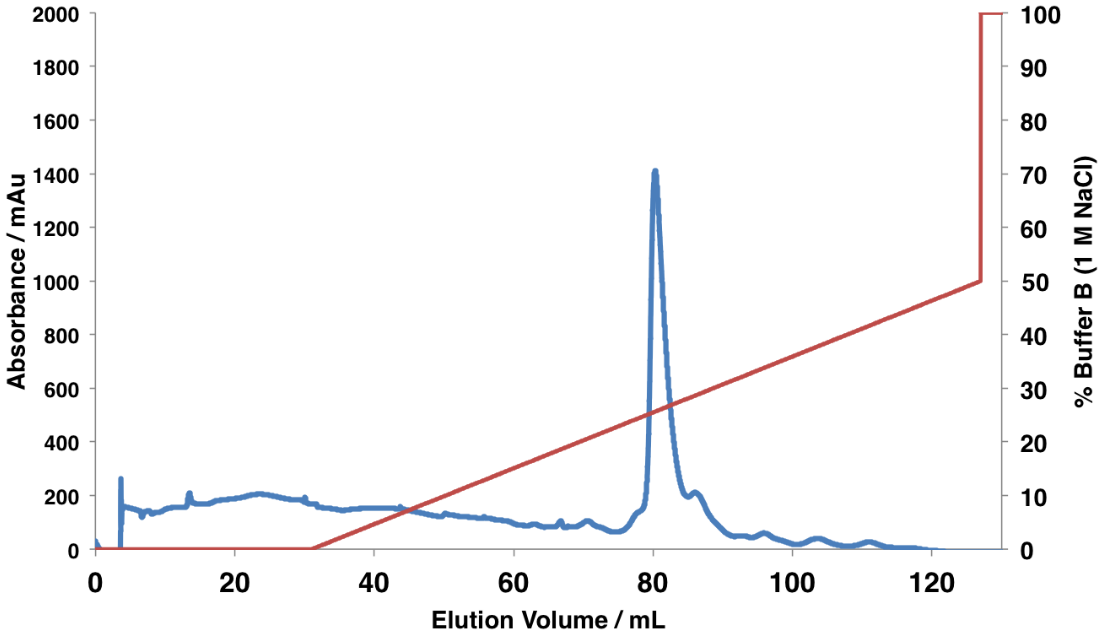


*Figure 3.4.2 – Anion exchange (Resource™Q) purification of PEF(L) homodimer, monitoring the absorbance at 280 nm over a gradient of 0-0.5 M NaCl, collecting 2 mL fractions.*

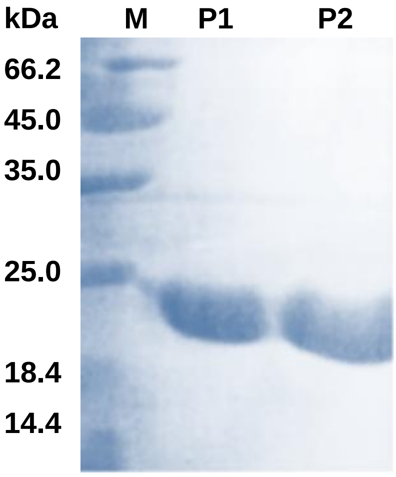


*Figure 3.4.3 - SDS-polyacrylamide gel of the peak fractions (P1, P2) of anion exchange purification of PEF(L), M is the protein ladder.*

### 3.4.3. Analytical size exclusion chromatography

Analytical size exclusion chromatography was used to determine the oligomeric state of PEF(L) in solution, using the method described for PEF(S) (Section 3.1.6).

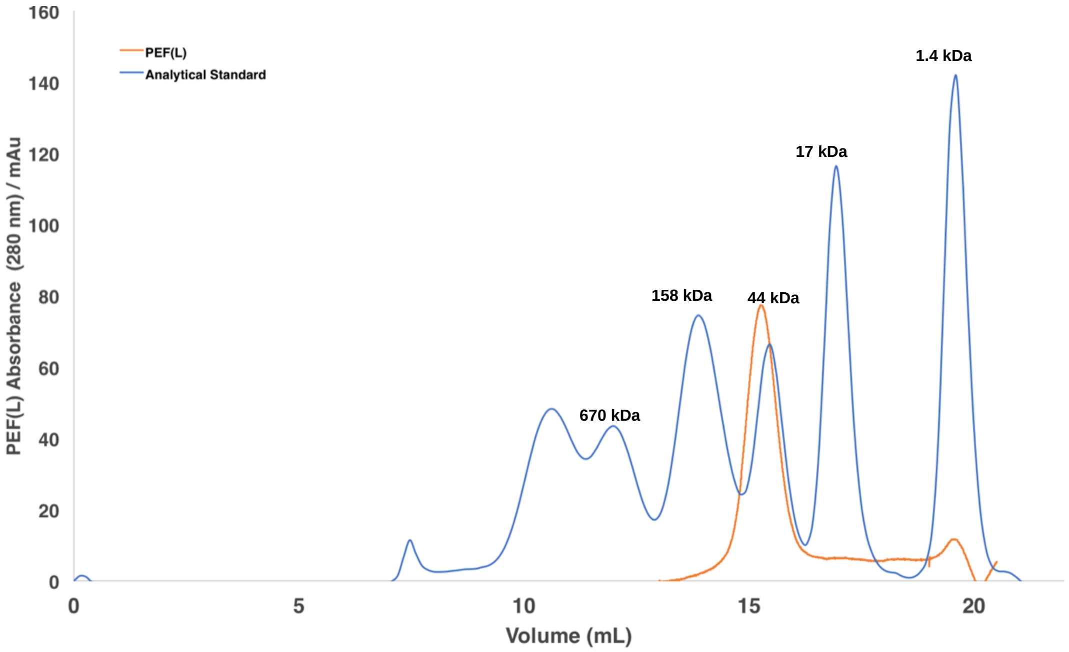


*Figure 3.4.4 - Analytical size exclusion chromatogram using Superdex™ 200 10/300 GL using BioRad™ analytical standards (bovine thyroglobulin - 670 kDa, bovine γ-globulin - 158 kDa, chicken ovalbumin - 44 kDa, horse myoglobin – 17 kDa, cyanocobalamin – 1.4 kDa), 100 µL of 90 µM PEF(L) passed through the column at a flow rate of 0.4 mL / min.*

PEF(L) eluted from size exclusion column just before the 44 kDa peak in the standard, indicating that the protein readily formed a dimer in solution. In the full length heterodimeric calpain-1 or -2 complex, PEF(S) and PEF(L) form the interface between the dimer however in calpain-3, a 180 kDa homodimer between the PEF(L) domains forms instead, believed to be through the lack of a small subunit.[189] Therefore it is reasonable that PEF(L) is capable of forming a homodimer in solution when PEF(S) is not present.

### 3.4.4. Mass spectrometry

A sample (0.5 mg / mL) of PEF(L) was analysed by electrospray ionisation mass spectrometry (ES-MS) in positive mode and deconvoluted using the MaxEntropy1 program (Section 6.7.6).

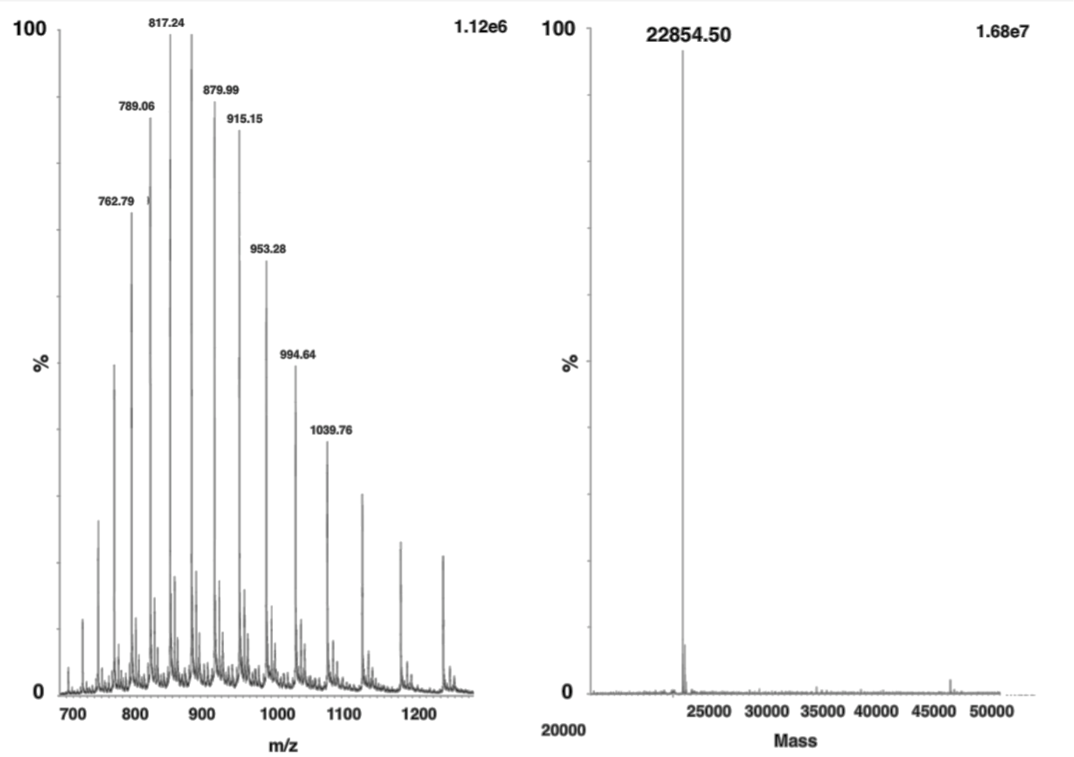


*Figure 3.4.5 - Mass spectrum (ESI$^{2+}$) of purified recombinant human PEF(L) (left), deconvoluted using the MaxEntropy1 program (right).*

The mass of 22854 corresponded to the mass of the $His_6$-PEF(L) construct, confirming the successful expression and purification of this protein.

### 3.4.5. Circular dichroism

To ensure the protein displayed the correct secondary structure, a CD spectrum was obtained for the sample. The pure PEF(L) protein sample was diluted to 10 µM in $K_iPO_4$ (10 µM, pH 7.4), the CD spectrum was recorded between 190-300 nm with 1 nm steps at 20 °C (Figure 3.4.6).

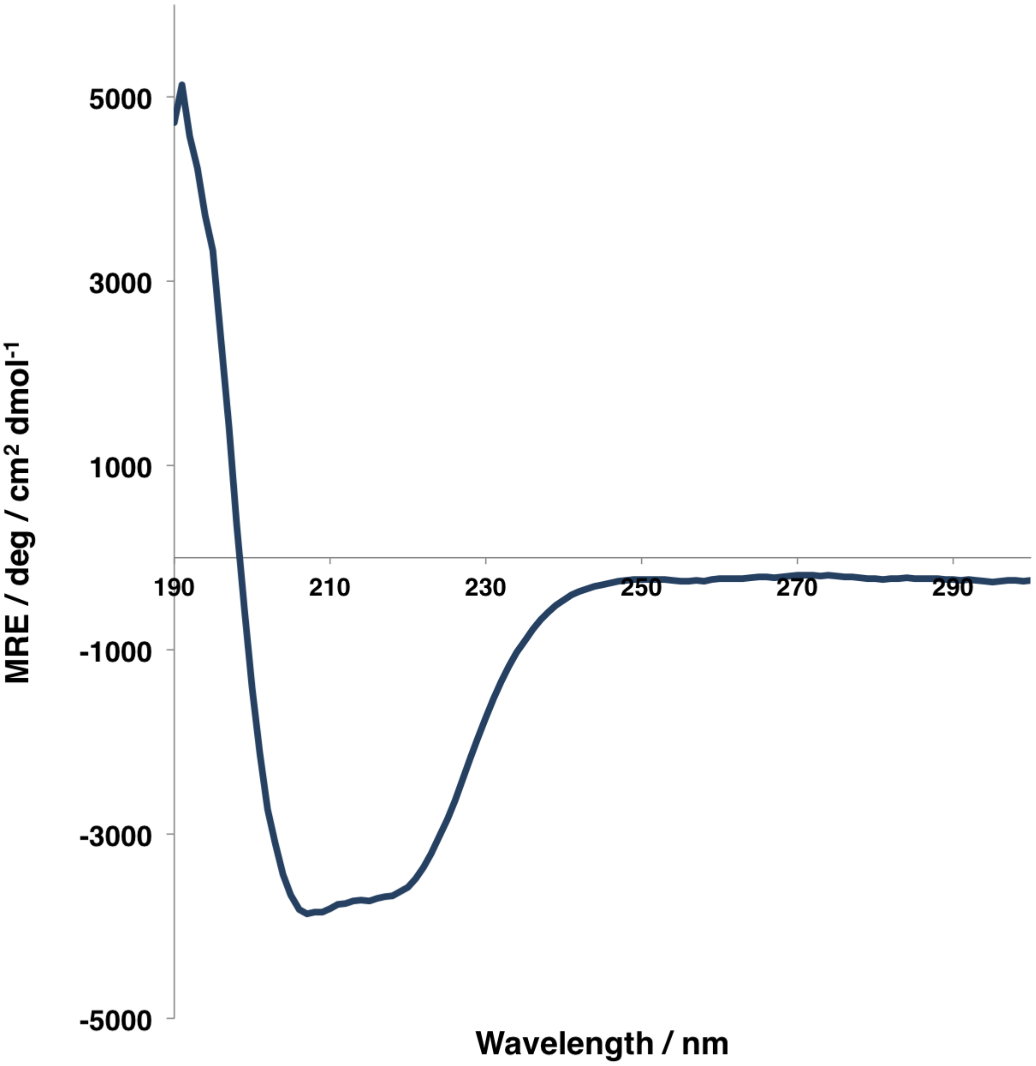


*Figure 3.4.6 - CD spectra of human PEF(L) showing the secondary structure to be mostly α-helical. Mean residue ellipticity calculation used (Equation 3, Section 6.7.8).*

The CD spectrum showed that PEF(L) was primarily α-helical and thus featured the predicted secondary structure. With confirmation that the protein was the correct size and secondary structure, binding studies of the diselenide inhibitors were attempted on human calpain-1 PEF(L).

### 3.4.6. TNS displacement assay

The EF hand domains on the large and small subunit are believed to be similar in structure on both calpain-1 and -2, the capastatin-bound calpain-2 model shows the α-helix of calpastatin subdomain A binds to PEF(L) *via* key hydrophobic residues L579 and L583 ( Figure 3.4.7).[37]

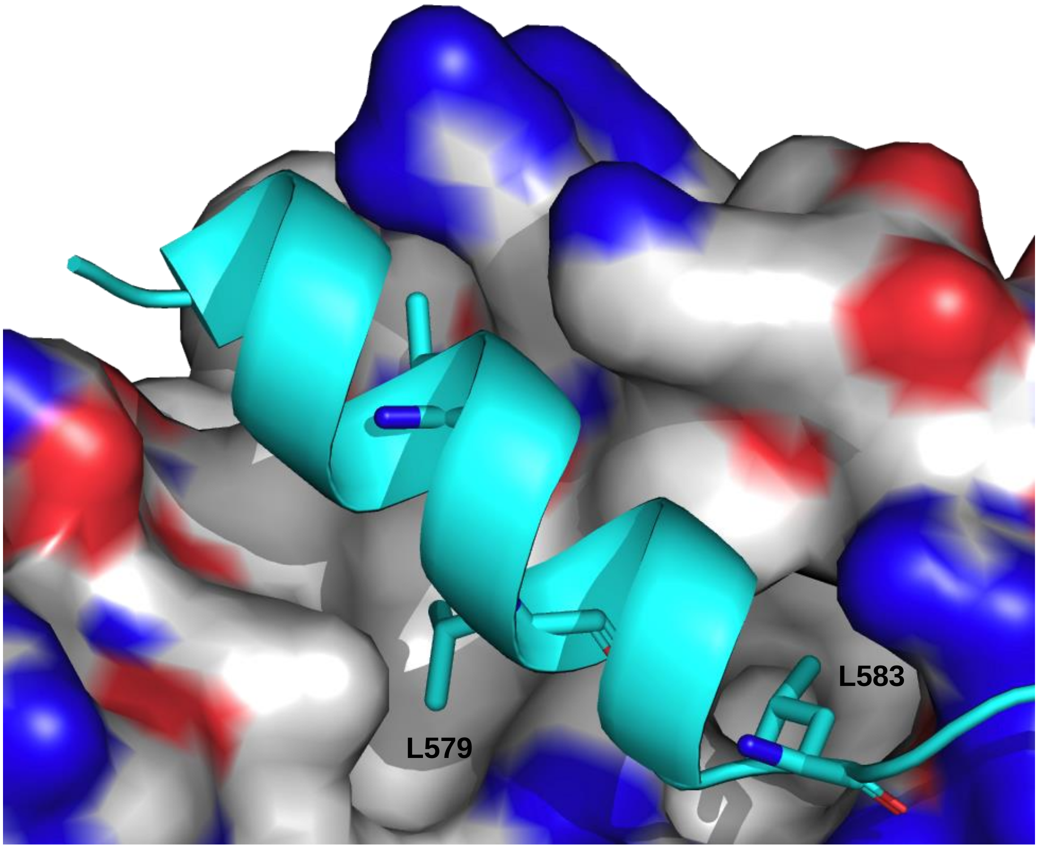


*Figure 3.4.7 - Cartoon representation of calpastatin (cyan) binding to PEF(L) of calpain-2 (surface, grey) with the key hydrophobic residues L579 and L583 represented as sticks (PDB:3BOW).*[37]

PEF(L) and PEF(S) crystal structure domains from calpain-2 (PDB:3BOW), were aligned in PyMol which revealed these domains share significant tertiary structure around the respective calpastatin binding grooves, suggesting that the mode of small molecule binding could be analogous to PEF(S) binding.[37]

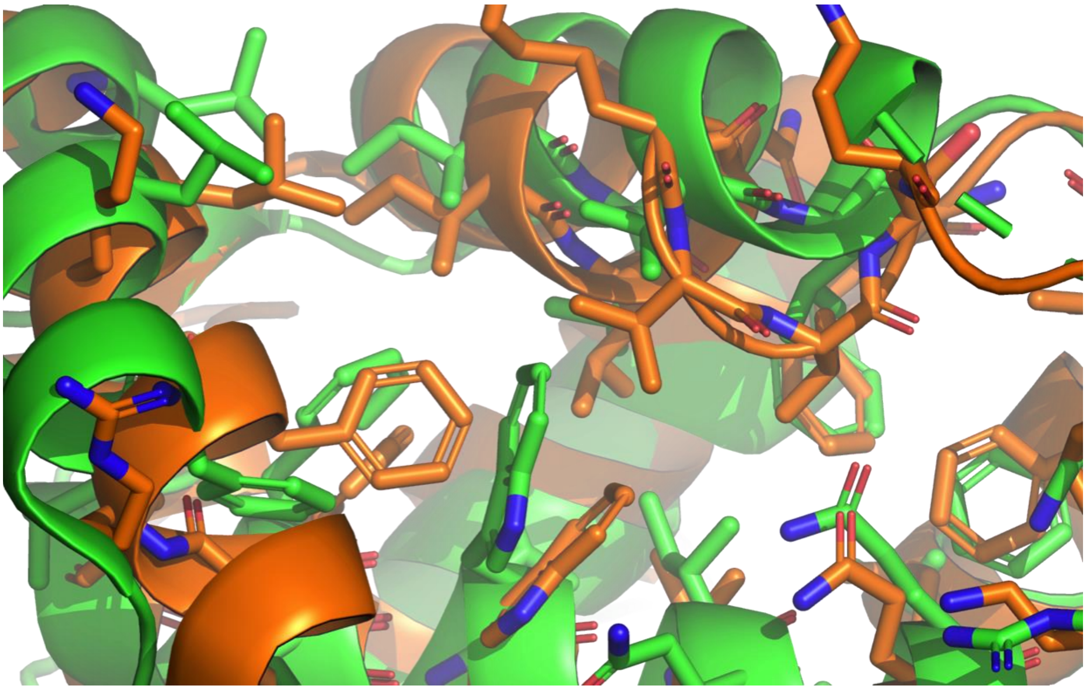


*Figure 3.4.8 – Cartoon representation of the residues surrounding the hydrophobic binding site on PEF(S) (green) and PEF(L) (orange) of calpain-2 aligned in PyMol (PDB:3BOW).[24,47]*

The TNS displacement assay was carried out on the 4-halogenated phenyl diselenides **35**-**38** and data analysed as described (Section 6.8.2), substituting PEF(S) for PEF(L) to the same concentration (10 μM). Glycerol and imidazole were removed by washing the protein sample thoroughly using a Vivaspin protein concentrator (10 kDa MWCO), into fluorescence assay buffer (Section 6.8.2) Despite the structural similarities between PEF(S) and PEF(L), PEF(L) expressed as a recombinant construct was significantly less stable, tending to form aggregates over time, which limited the material available for structural and binding studies.

The maxima at 444 nm was taken for each concentration of inhibitor as described (Section 3.2.1), the data was normalised and inverted by subtracting the fluorescence (B) from the maximum fluorescence, $B_{max}$. The data was then analysed by non-linear regression using Sigmaplot (Systat Software, San Jose, CA) Ligand Binding module.

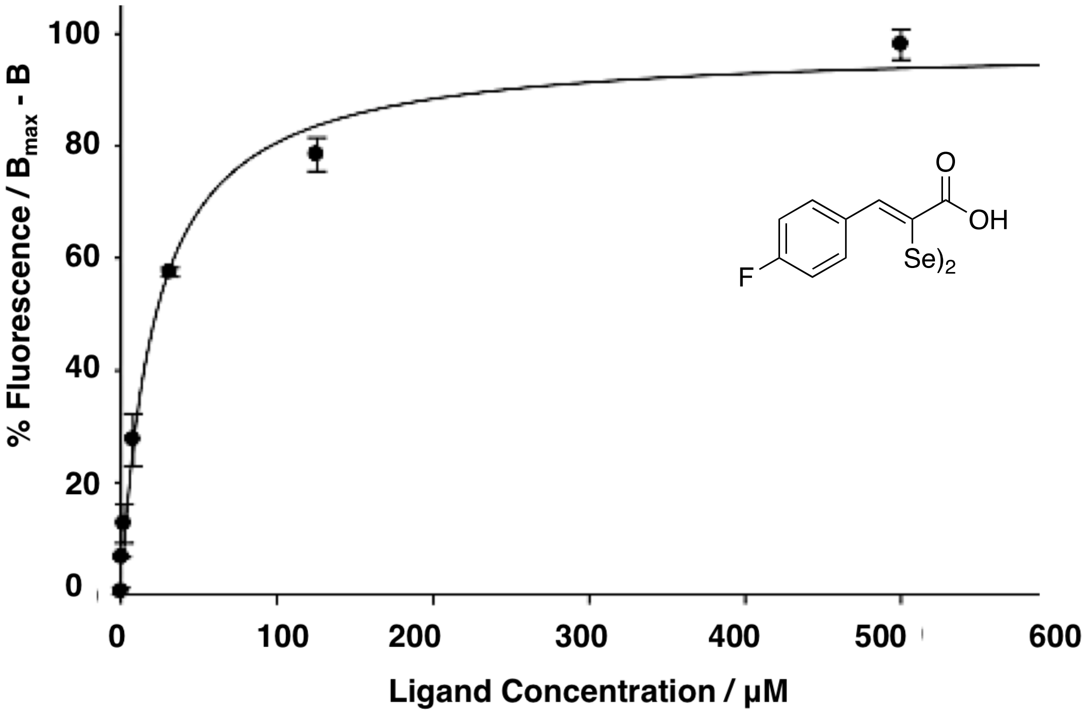


*Figure 3.4.9 - TNS displacement assay on PEF(L) with* ***35*** *measuring the change in fluorescence at different concentrations of ligand, normalised and inverted the peak maxima then analysed by non-linear regression Ligand Binding module on Sigmaplot (Systat Software, San Jose, CA) and Equation 2.*

| **Compound** | **$K_d$ / µM** |
|---|---|
| **9** | 14.0 ± 3.1 |
| **35** | 21.4 ± 2.7 |
| **36** | 15.4 ± 3.3 |
| **37** | 12.1 ± 4.8 |
| **38** | 9.6 ± 1.9 |

*Table 3.7 – A table displaying the $K_d$ values obtained by the TNS displacement assay for calpain-1 PEF(L) by measuring the change in fluorescence at different concentrations of α-mercaptoacrylic acid PD150606 and diselenide compounds* ***35-38****.*

The data obtained from the TNS displacement assay demonstrates that the diselenides halogenated at the 4 position (**35 – 38)** and PD150606 bind to PEF(L) with a comparable apparent affinity as to the calcium binding domain on the small subunit, PEF(S). This is likely due to the similarity between the hydrophobic binding sites on both the PEF(S) and PEF(L) domains, resulting in a similar affinity for the protein. A greater affinity for PEF(L) was observed with increasing size of the halogen. These observations should be backed up by further structural confirmation that the ligands bind to the hydrophobic pocket on PEF(L), currently the only crystal structures of an inhibitor bound to PEF(L) is the calpastatin calpain-2 complex as discussed (Section 1.5.1).

Attempts at concentrating PEF(L) by Amicon Ultraflltration (10 kDa MWCO) or Vivaspin centrifugal concentrator to concentrations required for crystallography (>10 mg mL$^{-1}$) resulted in significant aggregation. The expressed PEF(L) construct contains regions that are not usually solvent exposed in the full length protein which may explain the inherant instability of the recombinant construct, a more stable construct would perhaps be a heterodimeric form of the two PEF domains, PEF(S) / PEF(L).

# 3.5. PEF(L) / PEF(S) heterodimer

The solubility and stability of PEF(S) as a standalone protein has enabled a crystallographic data to be collected. As PEF(S) and PEF(L) form a dimer in the full length protein, a more stable dimer than PEF(L) homodimer was considered to be the PEF(L)/PEF(S) heterodimer. The heterodimer was synthesised by mixing harvested cell pellets of both PEF(S) and PEF(L) together then lysing both simultaneously by sonication. The mixture was then passed through a Ni-NTA column, following the protocol for PEF(L) (Section 3.4.1), as both the heterodimer and PEF(L) homodimer contain polyhisitidine tags, it was likely the elution fraction would contain a mixture of these which was observed by an larger band for PEF(L) over PEF(S) on the SDS-PAGE gel (250, Figure 3.5.1).

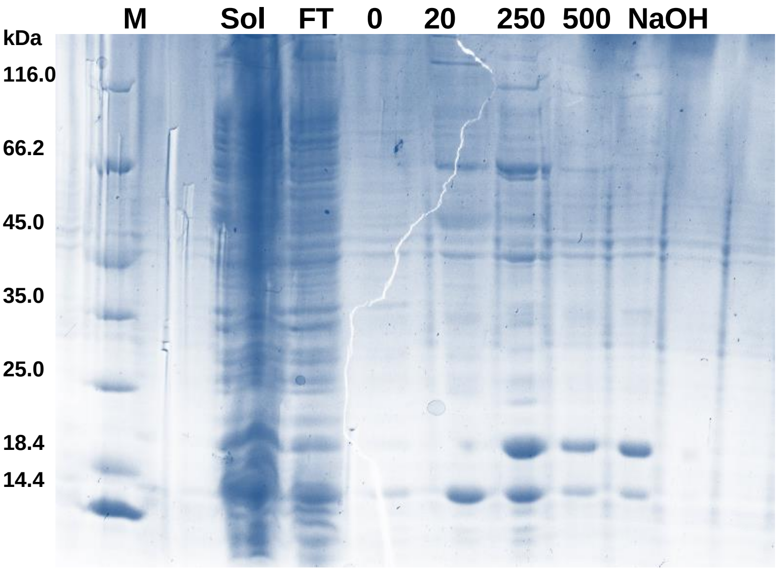


*Figure 3.5.1 - SDS-polyacrylamide gel of PEF(L) / PEF(S) heterodimer Ni-NTA purification, M is protein ladder, Sol is the soluble fraction, FT is the unbound flow-through, 0, 20, 250 and 500 mM imidazole fractions and the sodium hydroxide wash.*

### 3.5.1. Anion exchange chromatography

The elution fraction (PEF(L) Buffer A containing 250 mM imidazole) NaCl concentration was reduced to 50 mM NaCl using (PEF(L) Buffer C, Section 6.6.6) as a dilutant. The sample was then applied to a Resource™Q high resolution anion exchange column pre-equilibrated with PEF(L) Buffer C.

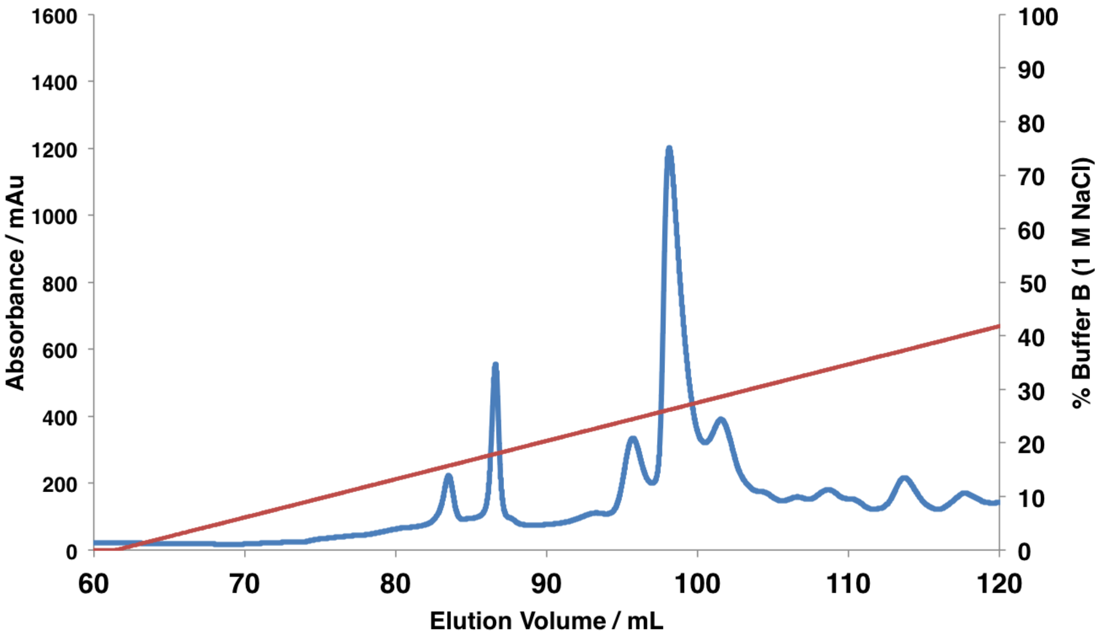


*Figure 3.5.2 - Anion exchange (Resource™Q) purification of PEF(L)/PEF(S) heterodimer, monitoring the absorbance at 280 nm over a gradient of 0-0.5 M NaCl, collecting 2 mL fractions.*

The fractions containing both PEF(L) and PEF(S) together were combined and concentrated to 300 µL for a final purification step using a Superdex™ 200 10/300 GL gel filtration column to ensure the protein was indeed both pure and the correct size in solution.

### 3.5.2. Analytical size exclusion chromatography

Analytical size exclusion chromatography was used to check if the size of the PEF(L)/PEF(S) heterodimer in solution was correct, using the method described for PEF(S) and PEF(L) (Section 3.1.6).

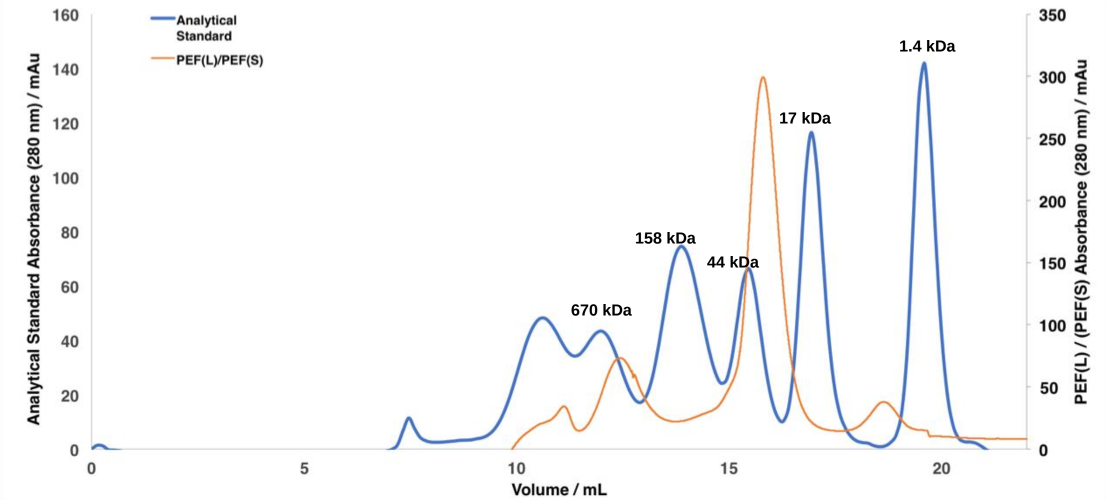


*Figure 3.5.3 - Superdex™ 200 10/300 Gl size exclusion chromatography of the PEF(L) / PEF(S) heterodimer complex (yellow) overlaid on the analytical standard (Bio-Rad, blue).*

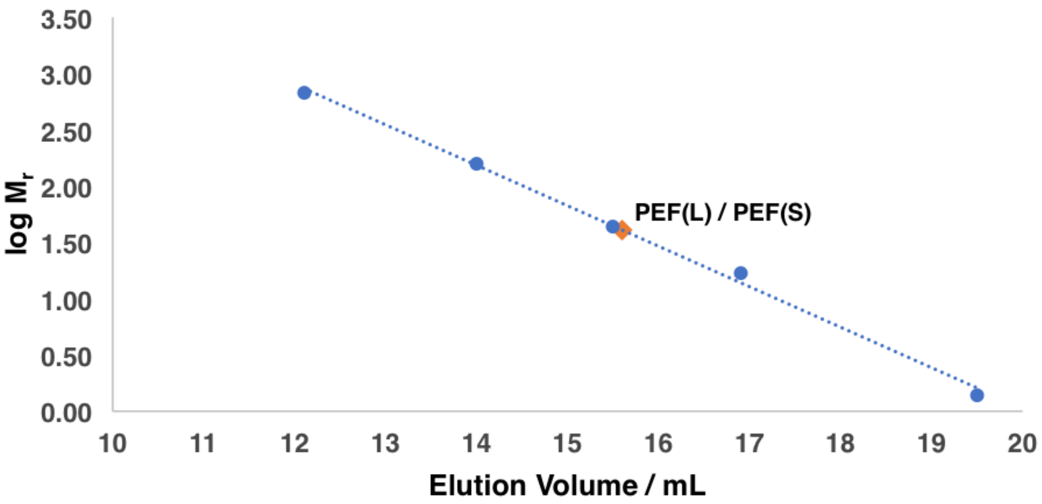


*Figure 3.5.4 – Calibration log$M_r$ graph of the elution volumes of the analytical standards (Bio-Rad) from the Superdex™ 200 10/300 Gl size exclusion column compared to the PEF(L) / PEF(S) heterodimer complex (yellow) overlaid on the analytical standard (Bio-Rad, blue).*

Calibration of the size exclusion column was carried out using elution volumes of the analytical standard peaks to create a linear log$M_r$ graph (Figure 3.5.4). The elution volume of the of the heterodimer complex suggests a molecular weight of 40.6 kDa which is close to the size of the dimer. The SDS-polyacrylamide gel of the purified product (Figure 3.5.5) also shows both PEF(S) and PEF(L) in an equal ratio.

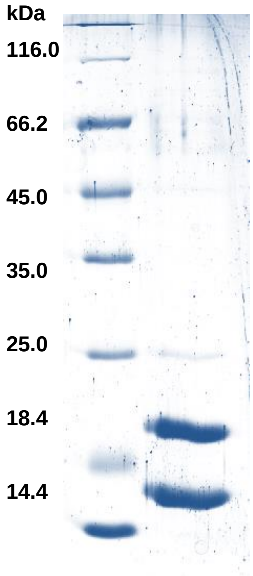


*Figure 3.5.5 – SDS-polyacrylamide gel of the purified heterodimer complex, M is the protein ladder and L/S corresponds to the sample from the size exclusion purification run.*

### 3.5.3. Circular dichroism

The pure human calpain-1 PEF(L)/PEF(S) heterodimer complex sample was diluted to 10 $\mu$M in $K_iPO_4$ (10 $\mu$M, pH 7.4), the CD spectrum was recorded between 190-300 nm with 1 nm steps at 20 °C. The two negative peaks at 208 nm and 220 nm (Figure 3.5.6).

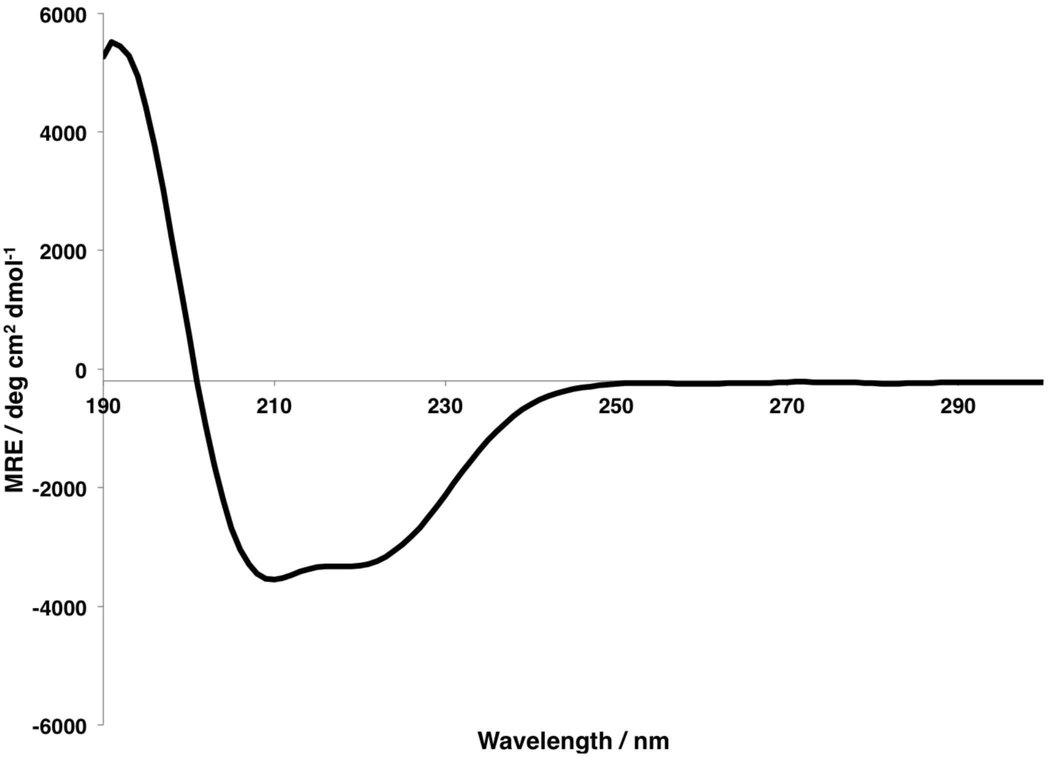


*Figure 3.5.6 - CD spectra of human calpain-1 PEF(L)/PEF(SL) heterodimer complex, showing the secondary structure to be mostly α-helical. Mean residue ellipticity calculation used (Equation 3, Section 6.7.8)*

The production of the correctly folded heterodimeric construct could serve as a tool for development of isoform selective allosteric compounds. Structural data of this complex may provide a model for designing compounds with a preference for PEF(L) binding over PEF(S), which may ultimately yield more selective allosteric inhibitors.

### 3.5.4. Concluding remarks for PEF domains

The calcium-binding domain from human calpain regulatory subunit, PEF(S), was expressed, purified and characterised. Quantification of the binding of α-mercaptoacrylic acid calpain inhibitor PD150606 (**9**) and several novel diselenide compounds to this protein was explored using a TNS displacement fluorescence assay. X-ray co-crystal structures of inhibitor-bound PEF(S) enabled the comparison of disulfide and diselenide based compounds, showing that the geometry of the dihedral bond is important in the potent binding observed.

The calcium-binding domain from human calpain-1 large subunit, PEF(L), was also successfully expressed, purified and characterised. The TNS displacement assay was also used to show that both PD150606 (**9**) and the diselenides **35 – 38** bind to this domain with a similar affinity to PEF(S). The instability of the PEF(L) homodimer was overcome by producing a heterodimeric PEF(L)/PEF(S) complex. Although not further explored used in this project, this complex could be useful in the further development of selective allosteric calpain-1 inhibitors and compounds that selectively bind to PEF(L) over PEF(S).

# Chapter 4 - CysPC Interactions

# 4.1. Interactions with CysPC

## 4.1.1. PD150606 inhibition of CysPC

The CysPC domain is an active, stable autolysis fragment of calpain-1 approximately 35 kDa in size with a lower $Ca^{2+}$ requirement than the full length enzyme.[39] Recombinant CysPC was produced for use in activity assays at a concentration of 100 nM to re-evaluate the mode of inhibition of PD150606.[39] PD150606 was shown to reduce the activity of CysPC by approximately 50% at 50 μM concentration, using a ((EDANS)-EPLFAERK-(DABCYL)) FRET substrate.[39] It was concluded that PD150606 cannot inhibit calpain-1 by binding to the PEF domains.[39] This observation confirmed that the α-mercaptoacrylic acid compounds are not allosteric in their mode of action, despite displaying allosteric binding properties. If a compound inhibits the full length calpain-1 enzyme but not the isolated active site alone, the observed activity must arise from allosteric inhibition. The methodology used to evaluate the allosteric inhibitory properties of the α-mercaptoacrylic acids was then used on the diselenides.

## 4.1.2. Cloning of CysPC

Residues 29-325 of calpain-1 were excised from the pET24b(+) CAPN1 plasmid and amplified by PCR using primers which added a BsaI restriction enzyme cut sequence flanking the CysPC gene, for use in a golden gate cloning protocol. Golden gate cloning is a one-pot restriction-ligation protocol that joins a template and an insert, using the type-II restriction enzyme BsaI which recognises and cuts at 5’-GGTCTC(N1)/(N5)-3’, the cut produces sticky ends between the two fragments, and a high activity T4 ligase is then used to ligate the DNA into a new target vector.[190]

A pET-28(a) plasmid was chosen for ligation of the CysPC gene fragment which contained an *N*-terminal (6x) polyhistidine tag, B1 domain of *Streptococcal* protein G (GB1) solubility enhancement tag (SET) and a short linker to CysPC containing a tobacco etch virus (TEV) protease cleavage site.

SETs such as GB1, maltose binding protein (MBP) and thioredoxin serve as highly soluble fusion proteins which can aid the folding and solubility of the target recombinant protein, and thus an increase in yield.[191] These are often then cleaved to yield the native protein, free of any tags.

In the pET-28(a) construct, the GB1 tag can be cleaved by the sequence specific TEV protease. TEV is a 27 kDa cysteine protease that is well characterised and cleaves peptides with the ENLYFQ/S sequence.[192] The product of this reaction leaves a serine residue prior to the first residue of the CysPC gene, unlikely to have a major effect on the properties of the protein.[42]

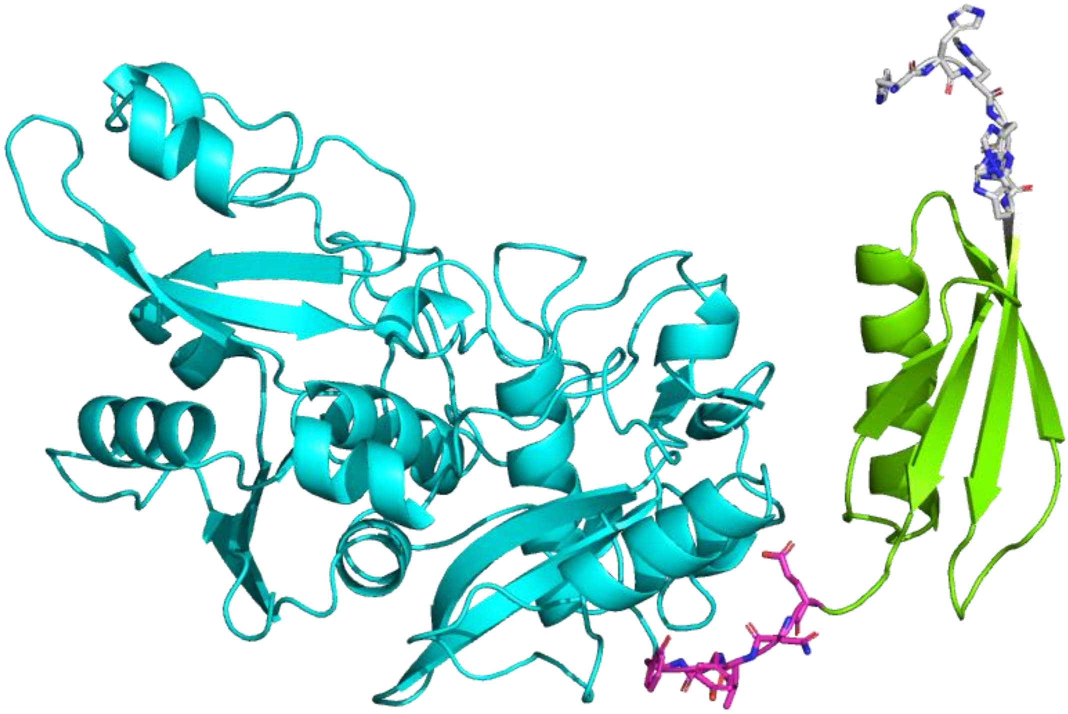


*Figure 4.1.1 – Graphical representation of CysPC-TEV-GB1-6xHis construct modelled using Protein Homology/analog Recognition Engine (PHYRE2) based on calpain-1 CysPC domain (cyan) (PDB:2ARY) B1 domain from Streptococcal protein G (green) [PDB:1EM7], TEV cleavage site (magenta) and 6x His tag (grey).*[193]

The golden gate ligation reaction contained BsaI, T4 ligase, the insert (CysPC) and the pET-28(a) acceptor plasmid, cycles of cutting and ligating enabled the accumulation of the ligated gene in the final plasmid which contained no BsaI sites (Section 6.4.11). The reaction was then transformed into chemically competent XL1-Blue cells and grown overnight on kanamycin selective agar plates. A single colony of the bacteria was grown overnight in 10 mL kanamycin selective media, then plasmid DNA isolated from the culture by the QIAprep spin miniprep kit manufacturers instructions (Section 6.4.6). The isolated DNA was sequenced to reveal that CysPC had been ligated into the pET28-(a) vector (Section 6.4.7).

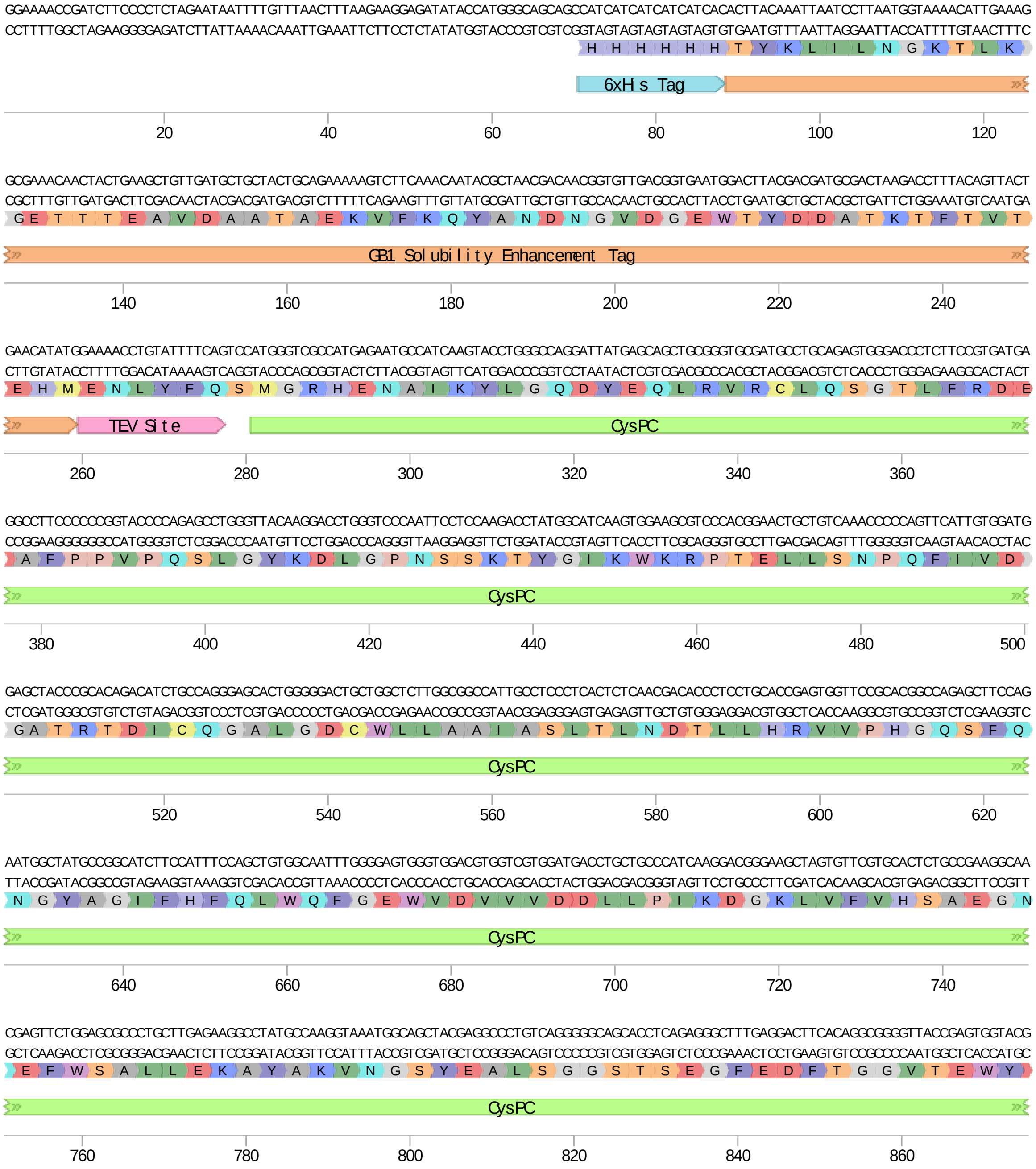


*Figure 4.1.2 – DNA Sequencing of the pET28-(a)-6xHis-GB1-TEV-CysPC golden gate construct.*

The sequencing of the golden gate cloning protocol showed the gene had been successfully incorporated into the target plasmid.

### 4.1.3. Expression and purification of TEV protease

TEV protease was required in the purification of CysPC free of the purification and solubility enhancement tag. TEV was produced using an *E.coli* expression system. The BL21 (DE3) cells containing the TEV protease gene were grown from a frozen glycerol stock at 37 °C in ampicillin selective LB media until $OD_{600}$ = 0.6 then induced with IPTG to reach a final concentration of 1 mM. The protein was expressed overnight at 20 °C and cells harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C. The cells were re-suspended in 20 mM HEPES, 100 mM NaCl, 0.5 mM TCEP pH 7.6 (buffer A) and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 RCF in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 μm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 CV buffer A and eluted with 10 CV buffer A containing 250 mM imidazole. The elution fraction was analysed by SDS-PAGE, then mixed with 20% v/v glycerol (20 mL final volume), and stored at -80 °C in 1 mL aliquots.

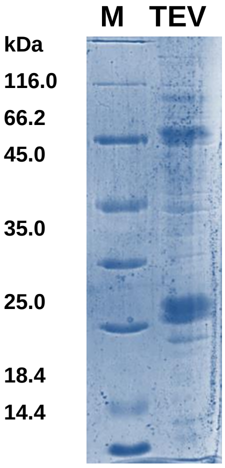


*Figure 4.1.3 - SDS-polyacrylamide gel of the isolated TEV protease expressed in E.coli bacteria and purified by Ni-NTA chromatography.*

The SDS-polyacrylamide gel showed a significant band at 27 kDa which was approximately the size of TEV.

### 4.1.4. Expression and purification of CysPC

BL21-CodonPlus (DE3) RP cells containing the human calpain-1 CysPC-GB1 plasmid were grown at 37 °C in kanamycin selective TB ‘Enhanced’ media until $OD_{600}$ = 0.6-0.8 then induced with 1 mM IPTG. The protein was expressed overnight at 20 °C and cells harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C. The cells were re-suspended in 20 mM HEPES, 100 mM NaCl, 0.5 mM TCEP pH 7.6 (CysPC Buffer A - Section 6.6.8) and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 RCF in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 μm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 CV buffer A and eluted with 10 CV buffer A containing 250 mM imidazole, which was further dialyzed in buffer A overnight in a 10 kDa MWCO membrane containing a 1 mL aliquot of TEV protease for ‘reverse’ Ni-NTA purification. The cleavage product was then passed back through a Ni-NTA column to remove the cleaved 6xHis-GB1 solubility tag and TEV protease, whilst the flow through contained CysPC as confirmed by SDS-PAGE.

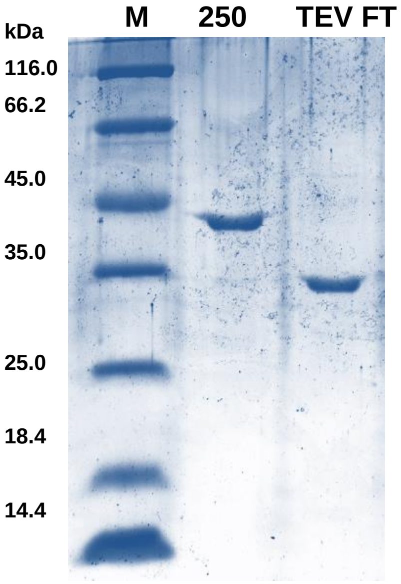


*Figure 4.1.4 – SDS-PAGE of the TEV cleavage reaction of eluted CysPC (250 – buffer A 250 mM imidazole), subsequently removing the 6x His Tag + GB1 fusion protein and the potein being present in the flowthrough (FT).*

### 4.1.5. Analytical size exclusion chromatography

In order to detemine the oligomeric state of CysPC in solution, size exclusion chromatography was carried out as described for the PEF domains (Section 3.1.4).

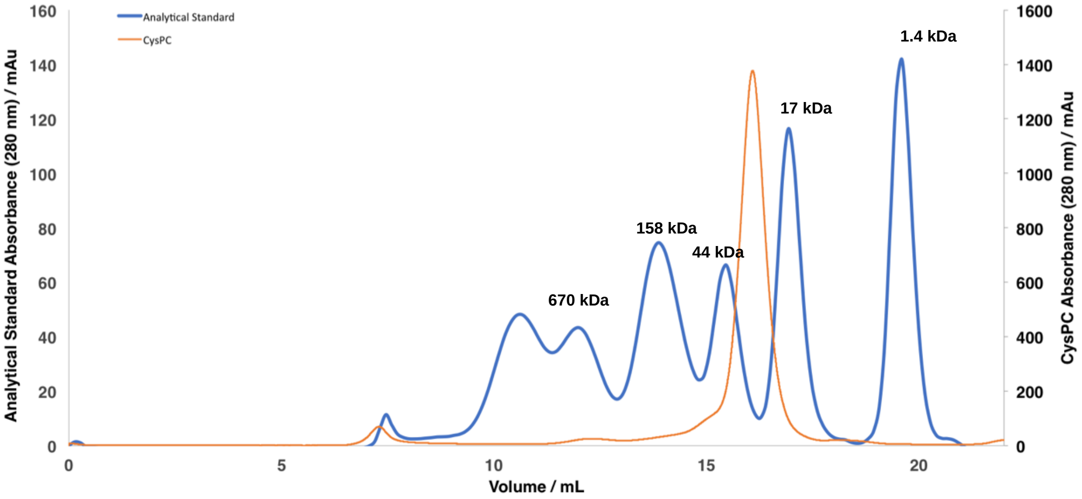


*Figure 4.1.5 - Size exclusion chromatogram using the Superdex™ 200 10/300 GL using BioRad™ analytical standard (bovine thyroglobulin - 670 kDa, bovine γ-globulin - 158 kDa, chicken ovalbumin - 44 kDa, horse myoglobin – 17 kDa, cyanocobalamin – 1.4 kDa), 300 μL of 60 μM CysPC was loaded onto the column and eluted at a flow rate of 0.4 mL / min*

The absorbance at 280 nm of purified CysPC unveiled the major peak to be between 17 and 44 kDa from the analytical standard, indicating that CysPC was a monomer in solution (Figure 4.1.5).

### 4.1.6. Circular dichroism

A CD spectrum of CysPC was obtained to ensure the protein was folded correctly after TEV cleavage.

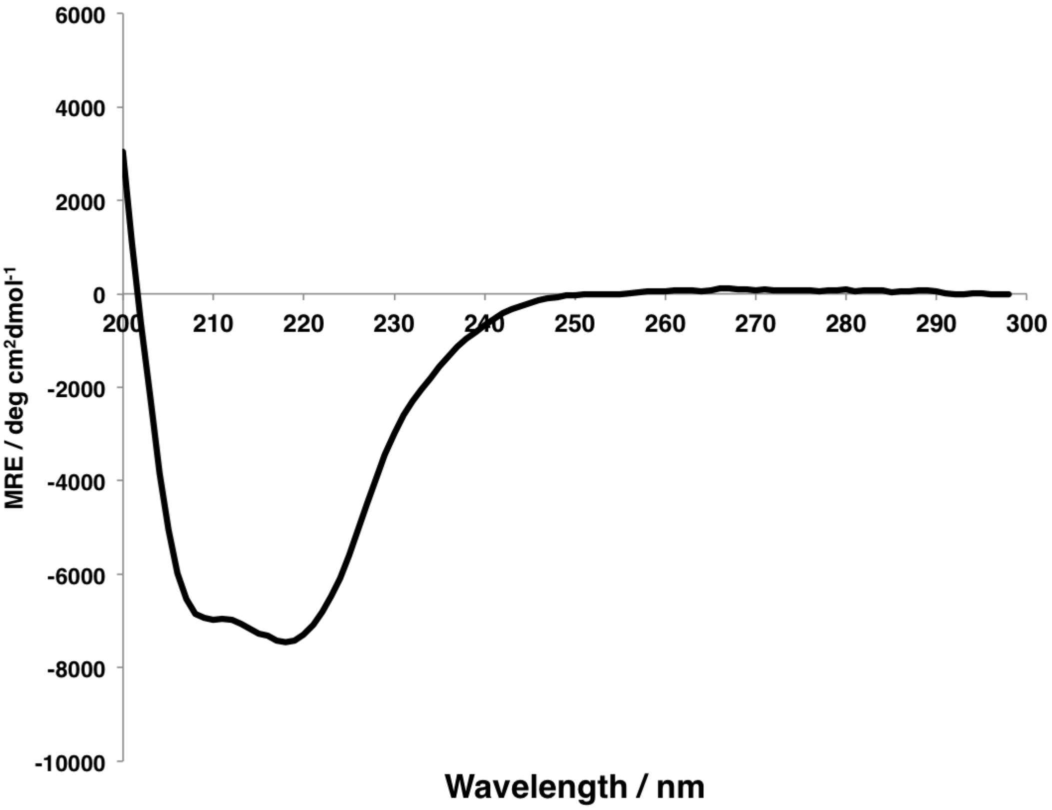


*Figure 4.1.6 – CD spectrum of purified TEV cleaved CysPC.*

The two negative peaks at 208 nm and 220 nm strongly indicated the secondary structure of CysPC was primarily α-helical, which is in accordance with the available structural data.[56]

### 4.1.7. FRET assay

CysPC has been reported to be stable and active as a standalone protein without other calpain domains present, it is also an active autolysis fragment after *in vivo* calpain activation.[36,104] Therefore this domain was used as a way of testing whether the synthesised diselenides displayed allosteric inhibitory characteristics. The FRET assay used to test for full length calpain-1 inhibition as described (Section 2.3) was adapted for use with the purified CysPC domain, to calculate apparent $IC_{50}$ values for the diselenide inhibitors.

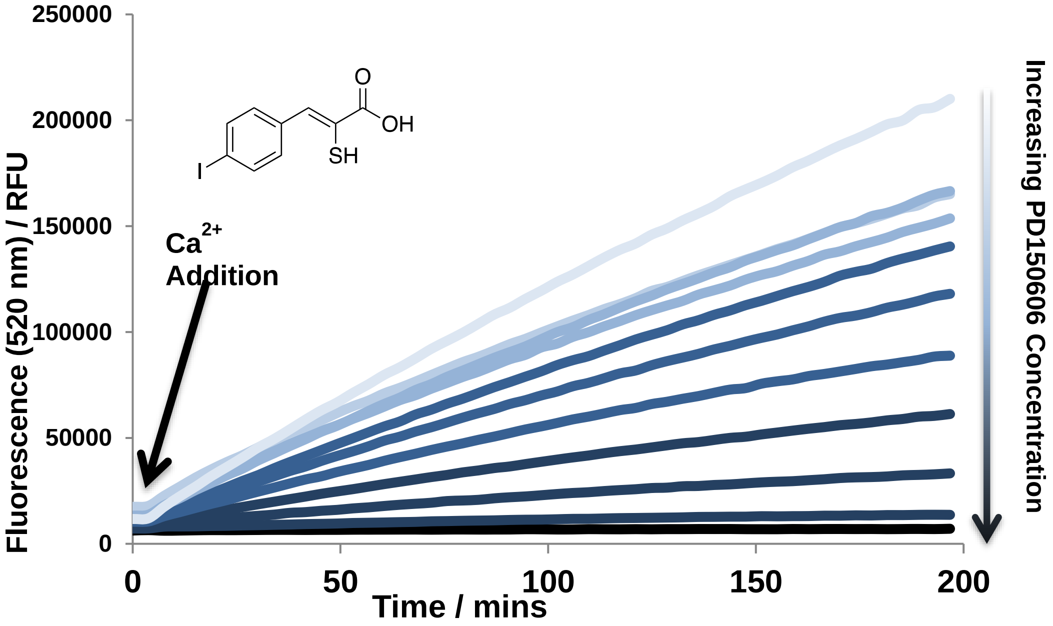


*Figure 4.1.7 - CysPC inhibition assay monitoring the change in fluorescence at 520 nm at different concentrations of PD150606 from 50 mM DMSO stock. Assays were performed with the calpain-1 FRET FAM substrate (1 µM), CysPC (100 nM), HEPES (10 µM), glutathione 9:1 (reduced : oxidised), 10 (µM), EDTA (0.5 mM), BSA (0.1%) at pH 7.0. Addition of $CaCl_2$ (5 mM) initiated the reaction*

The CysPC domain was capable of cleaving the calpain-1 FAM substrate in a similar manner to the full length enzyme, with the exception of CysPC displaying linear kinetics over a much longer duration. Upon calpain-1 activation, many factors are able to affect the kinetics such as subunit dissociation and autolysis, as CysPC is a single domain and is reported to be resistant to autolysis it is more predictable in the substrate cleavage assay.[34,50]

Increasing additions of PD150606 decreased the rate of the fluorescent FAM product formation, indicating that this compound inhibited calpain-1 CysPC activity (Figure 4.1.7). The gradients of the inhibition assay were analysed by non-linear regression in Sigmaplot (Systat Software, San Jose, CA) to determine the apparent $IC_{50}$ of PD150606 (Figure 4.1.8) as described for full length calpain-1 (Section 2.3).

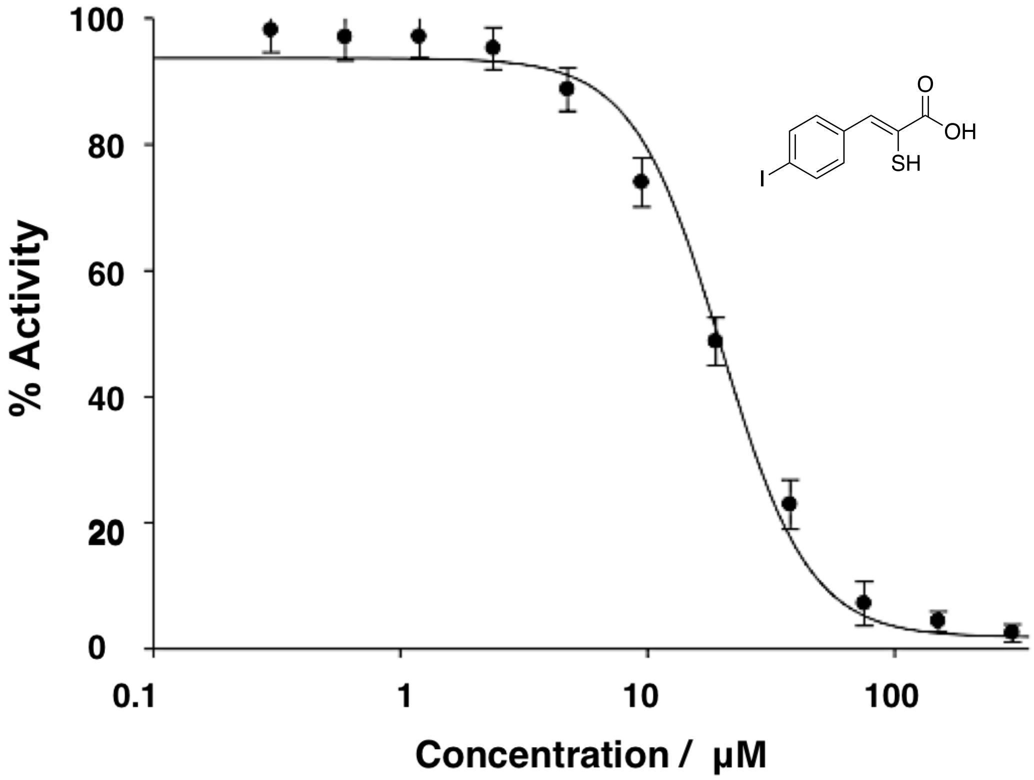


*Figure 4.1.8 - Normalised fluorescence gradients for increasing PD150606 (**9**) concentration in human calpain-1 CysPC assay, analysed by non-linear regression in Sigmaplot (Systat Software, San Jose, CA).*

The calculated apparent $IC_{50}$ for PD150606 on human calpain-1 CysPC was 17.8 ± 2.8 µM, confirming that this compound is not an allosteric inhibitor of calpain-1. The apparent $IC_{50}$ for PD150606 on calpain-1 was 14.1 ± 1.6 µM, which taking into account the error range for both values, could be the same value. The source of inhibition of this compound is therefore likely located on the active site alone and not dictated solely by binding to the PEF domains. Allosteric inhibition is important for calpains due to active site homology with other cysteine proteases. These proteolytic cores are often conserved domains sharing similar topology and chemistry, therefore may produce off target effects when used *in vivo*.[40]

The inhibition assay was repeated for several diselenides in the series to explore whether any difference in $IC_{50}$ was observed for these compound for full length calpain-1 or CysPC, which may be indicative of allosteric inhibition.

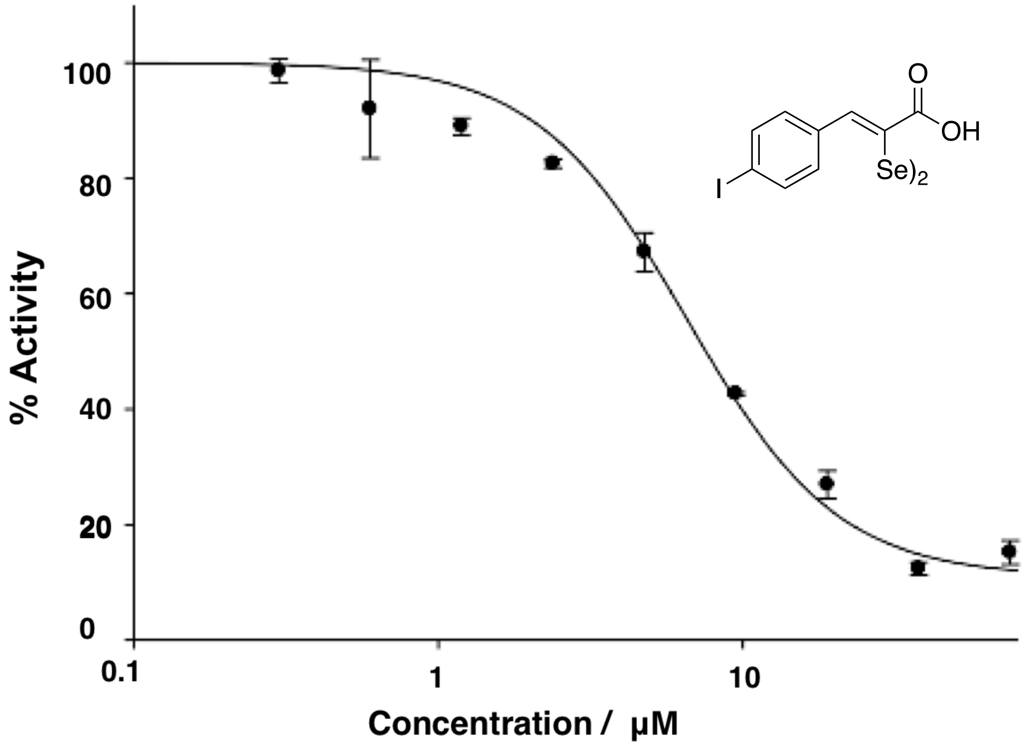


*Figure 4.1.9 - Normalised fluorescence gradients for increasing* ***38*** *concentration in human calpain-1 CysPC assay analysed by non-linear regression in Sigmaplot (Systat Software, San Jose, CA).*

The calculated apparent $IC_{50}$ for the 4-iodo diselenide analogue of PD150606, **38**, on CysPC was 7.0 ± 1.5 µM. The apparent $IC_{50}$ for **38** on the full length enzyme was 17.0 ± 3.6 µM, thus displaying a more potent inhibitory effect on the active site over the full length enzyme. Whilst the immediate reason for this is not apparent, the dynamic nature of the diselenide bond towards thiol containing molecules such as cysteine could potentially be a factor. Fewer binding sites are present on CysPC whereas calpain-1 contains both the PEF(S) and PEF(L) domains as well as the CysPC domain, therefore it is possible that the active site domain experiences a lower effective concentration of the diselenide inhibitor due to PEF binding are present.

Apparent $IC_{50}$ values were calculated for several other diselenide inhbitors, using the method described for the FAM substrate cleavage assay (Table 4.1).

| **Compound** | **Apparent $IC_{50}$ (μM)** |
|---|---|
| PD150606 - **9** | 17.8 ± 2.4 |
| **29** | 14.5 ± 1.1 |
| **30** | 9.9 ± 0.7 |
| **32** | 7.8 ± 0.6 |
| **33** | 10.3 ± 1.9 |
| **34** | 4.8 ± 2.2 |
| **35** | 3.2 ± 0.3 |
| **36** | 11.5 ± 1.0 |
| **38** | 7.0 ± 1.5 |
| **39** | 2.3 ± 0.4 |

*Table 4.1 - Apparent $IC_{50}$ values calculated for PD150606 (**9**) and various diselenide inhibitors using calpain-1 active site CysPC (100 nM) and the FAM substrate in the buffer conditions as described (Section 2.3) for the full length calpain-1 FRET assay.*

The apparent $IC_{50}$ values of the diselenides tested were compared for the active site and full length calpain-1, it was apparent that they are more potent inhibitors of the active site domain.

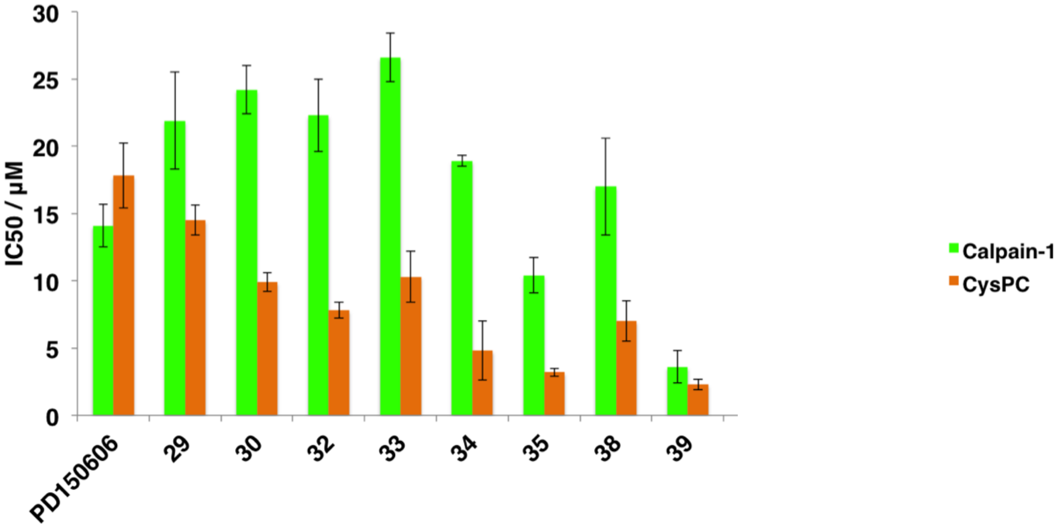
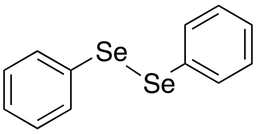


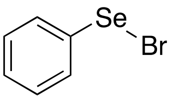
*Figure 4.1.10 – Apparent $IC_{50}$ values of diselenide inhibitors from tripclicate FRET assays on full length calpain-1 and CysPC using the FAM substrate.*

The crystallographic and fluorescence data (Section 3.4.6) supports the notion that the diselenides bind non-covalently to the calcium binding domains on the large and small subunits, PEF(L) and PEF(S). Despite displaying allosteric binding, a more potent inhibitory effect is observed on the calpain-1 active site domain, without the other calpain domains present. All diselenide inhibitors tested displayed a decrease in $IC_{50}$ for CysPC to calpain-1, compounds **32**, **34** and **35** exhibited a 3 – 4-fold decrease in $IC_{50}$ for the active site compared to the full-length enzyme. The UV-Vis spectra of the diselenides with cysteine (Section 2.2.4) and associated MS data suggests that these compounds are not redox stable and are readily able to react with cysteine residues, possibly forming selenyl-sulfide bonds in the process. The redox sensitivity of the compounds, combined with the proposed covalent and non-covalent interactions, creates a highly complex mode of action of the diselenide inhibitors on calpain-1.

Diphenyl diselenide (**45**) and phenylselenyl bromide (**46**) were tested for CysPC inhibition to further probe the involvement of selenium containing compounds as inhibitors of calpain-1.



**45**



**46**

*Figure 4.1.11 - Commercially available organoselenium compounds diphenyldiselenide (**45**) and phenylselenylbromide (**46**).*

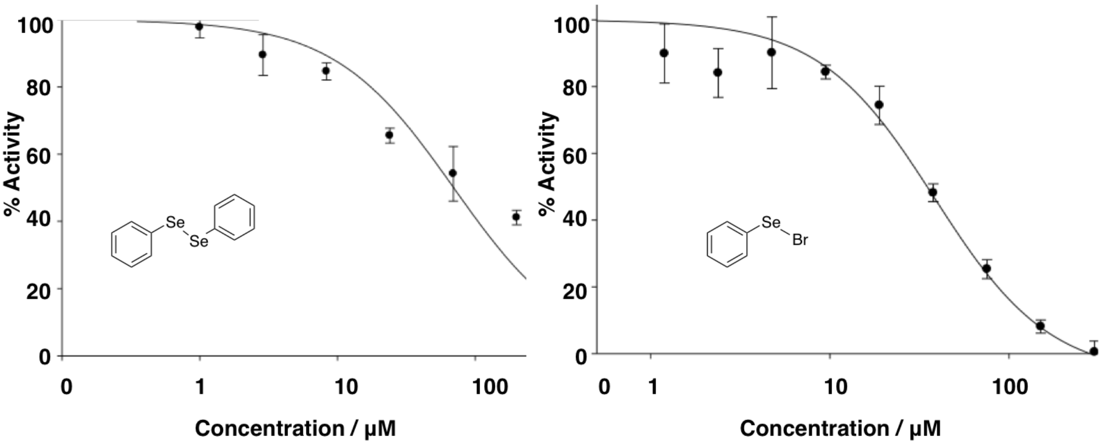


*Figure 4.1.12 – Normalised fluorescence gradients for different concentrations of diphenyldiselenide (**45**) and phenyl selenylbromide (**46**) concentration in human calpain-1 CysPC assay analysed by non-linear regression in Sigmaplot (Systat Software, San Jose, CA).*

Diphenyldiselenide displayed weaker inhibition of CysPC than the synthesised diselenide compounds with an apparent $IC_{50}$ of approximately 100 µM.

Phenylselenylbromide displayed much more potent inhibition with an apparent $IC_{50}$ of 38.4 ± 9.3 μM. Both these compounds have the propensity undergo nucleophilic attack from thiolate groups, which is therefore likely to be involved in the inhibition observed for the calpain-1 active site domain (Figure 4.1.13).

.

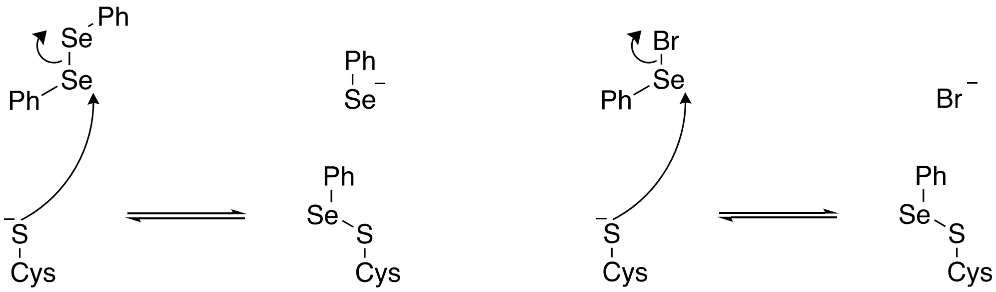


*Figure 4.1.13 - Proposed mechanisms for the formation of the phenylselenylsulfide adduct via nucleophilic attack by the cysteine thiolate anion on diphenely diselenide (**45**) and phenylselenylbromide (**46**).*

In order to further probe the interaction of diselenide inhibitors with the active site, crystallisation trials were carried out with the purified protein.

### 4.1.8. Crystallisation of CysPC

Purified CysPC solution was concentrated to 10 mg/mL *via* Amicon® ultrafiltration membrane (30 kDa MWCO). Crystallisation plates were set up using an Oryx4 dispensing robot (Douglas Instruments, Hungerford, UK) to dispense 200 nL of protein solution and 200 nL of precipitant from the $JCSG^{++}$ screen (Section 6.7). Crystals formed within 1 week of setting up the plates in 200 mM Bis-Tris, 200 mM NaCl, 25% w/v PEG 3,350 at pH 5.5. These were soaked with **35** 24 hours before harvesting as described (Section 6.7.2), and frozen in liquid nitrogen until diffraction at the synchrotron.

Data was collected at Diamond Light Source, the diffraction pattern was recorded up to a high resolution limit 2.29 Å, it was reduced and scaled using *xia2.*[178] Molecular replacement was carried out on the mtz file using Phaser-MR from the CCP4 program using the PEF(S) structure (PDB:2ARY) from the Protein Data Bank.[179] Several rounds of refinement with COOT (Computational Object-Orientated

Toolkit) and REFMAC5 were carried out until the model and electron density map were in agreement and statistics could not easily be improved further.[181,182]

| **Data Reduction Statistics** | **Human-CysPC soaked with 35** |
|---|---|
| **Unit cell dimensions / Å** | a = 102.92 b = 102.92 c = 63.73 |
| **Unit cell angles / °** | α = 90.00, β = 90.00, γ = 90.00 |
| **Space group** | I4 |
| **Resolution range / Å** | 72.78 – 2.29 |
| **Total reflections** | 104348 |
| **Unique reflections** | 14823 |
| **Completeness / % (last shell)** | 98.9 (100.0) |
| **I / σ (last shell)** | 9.8 (1.2) |
| **R(merge) / % (last shell)** | 12.3 (147.7) |
| **B(iso) from Wilson / Å$^2$** | 53.00 |
| **Refinement Statistics** | |
| **Protein atoms excluding H** | 2723 |
| **Calcium atoms** | 2 |
| **R factor / %** | 20.13 |
| **$R_{free}$ / %** | 26.99 |
| **R.m.s. bond length / Å** | 0.0109 |
| **R.m.s. bond angle / °** | 1.8618 |

*Table 4.2 - Data reduction and refinement statistics for recombinant CysPC soaked with diselenide **35** crystal X-ray diffraction dataset.*

Upon refinement of the CysPC crystal structure, cysteine residues C49 and C108 displayed large volumes of positive difference density on the sidechain thiol group in the electron density map. The predicted selenylsulfide adduct was created in JLigand and 3D geometric restraints applied through regularisation (4FP-CYS, Figure 4.1.14).[185] The cysteine residues that displayed large difference density were then removed and replaced with the 4FP-CYS ligand in COOT. Several rounds of refinement of the 4FP-CYS model were carried out with COOT and REFMAC5 until the CysPC-**35** model and electron density map were in agreement and statistics could not be improved further.[181,182]

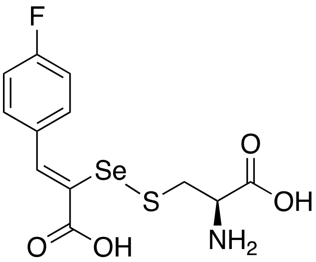


*Figure 4.1.14 - 4FP-CYS adduct drawn and regularised in JLigand then inserted into the CysPC protein backbone using COOT, replacing C49 and C108 residues.*[182,185]

Both predicted selenyl-sulfide adducts protruded out into the solvent region between CysPC monomers in the asymmetric unit (**A, C** - Figure 4.1.15). Flexible side chains in the solvent region can result in an apparent loss of visible electron density due to many available conformations from rotatable bonds.[194] The result of this is blurring of the electron density map resulting in reduced visible density for the mobile atoms.

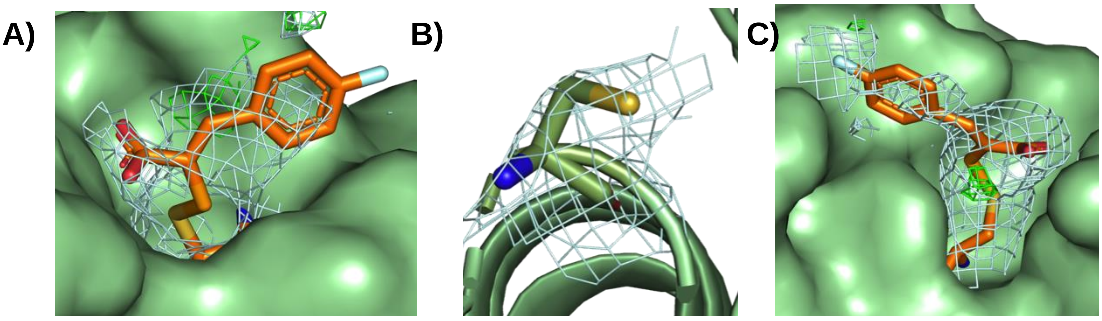


*Figure 4.1.15 - Cartoon representation of the electron density around CYS-C4F (orange, **A** – C49, **B** – C115, **C** – C108) on CysPC protein (green, with cartoon secondary structure). Map isomesh (grey) contoured to 1.0 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red).*

A flat bulk-solvent model is used in refinement to calculate constant electron density for any given region of the unit cell that is not occupied by the atomic model to account for the solvent electron density.[195] If the density for solvent exposed ligand atoms in the model is weak then the contribution of the bulk-solvent calculation to the density map for that region may further decrease the visible density for the ligand.[195]

The bulk solvent calculation for the ligand was switched off to see if any increase in the selenyl-sulfide density map could confirm the presence of the ligand, and provide possible geometric information indicating spatial orientation of the adduct.

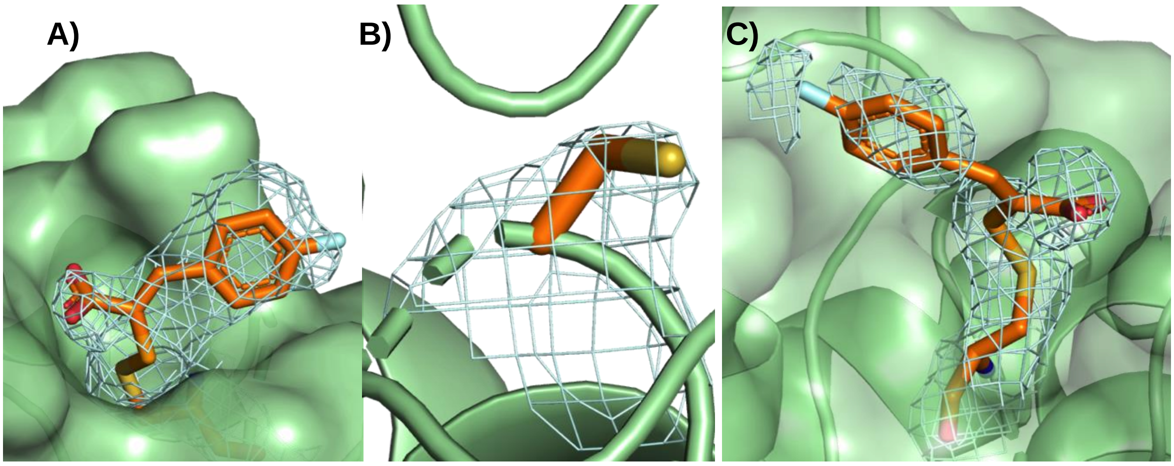


*Figure 4.1.16 - Cartoon representation of the electron density around CYS-C4F, without bulk solvent contribution during refinement (orange,* ***A*** *– C49,* ***B*** *– C115,* ***C*** *– C108) on CysPC protein (green, with cartoon secondary structure). Map isomesh (grey) contoured to 1.0 σ, carve = 2.0, difference map contoured to 3.0 σ, carve = 2.0 (positive = green, negative = red).*

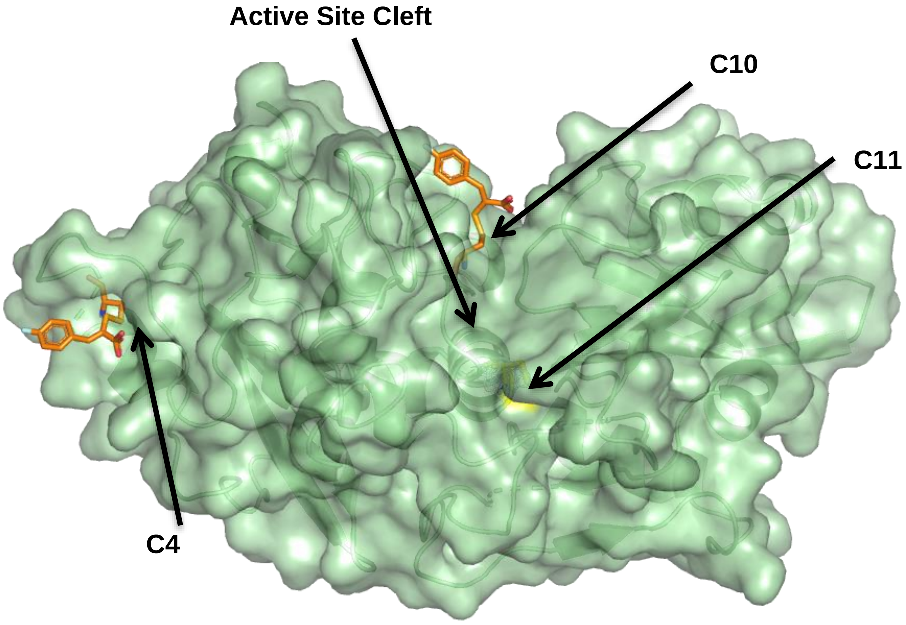
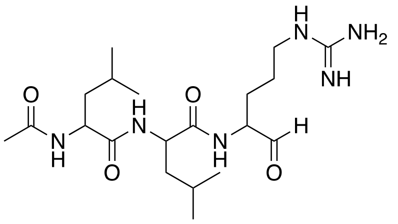


*Figure 4.1.17 - Cartoon representation of the CysPC-****35*** *model (green, surface), highlighting the surface cysteine residues C49, C108 and C115. C49 and C108 have selenyl-sulfide adducts shown (orange sticks, 4FP-CYS).*

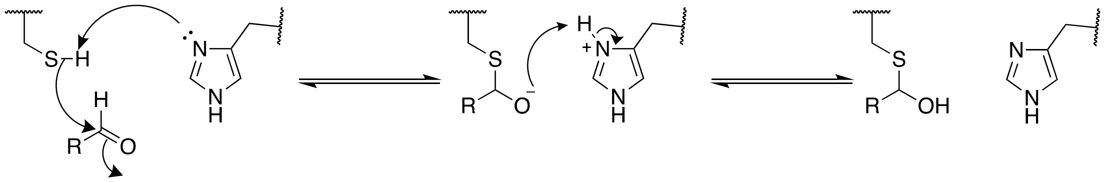
The difference electron density observed on C49 and C108 residues, attributed to the suspected selenyl-sulfide adduct, was not apparent on the active site cysteine, C115 ( Figure 4.1.16). C115 is significantly more buried in the active site cleft of the protein compared to C49 or C108 (Figure 4.1.17).

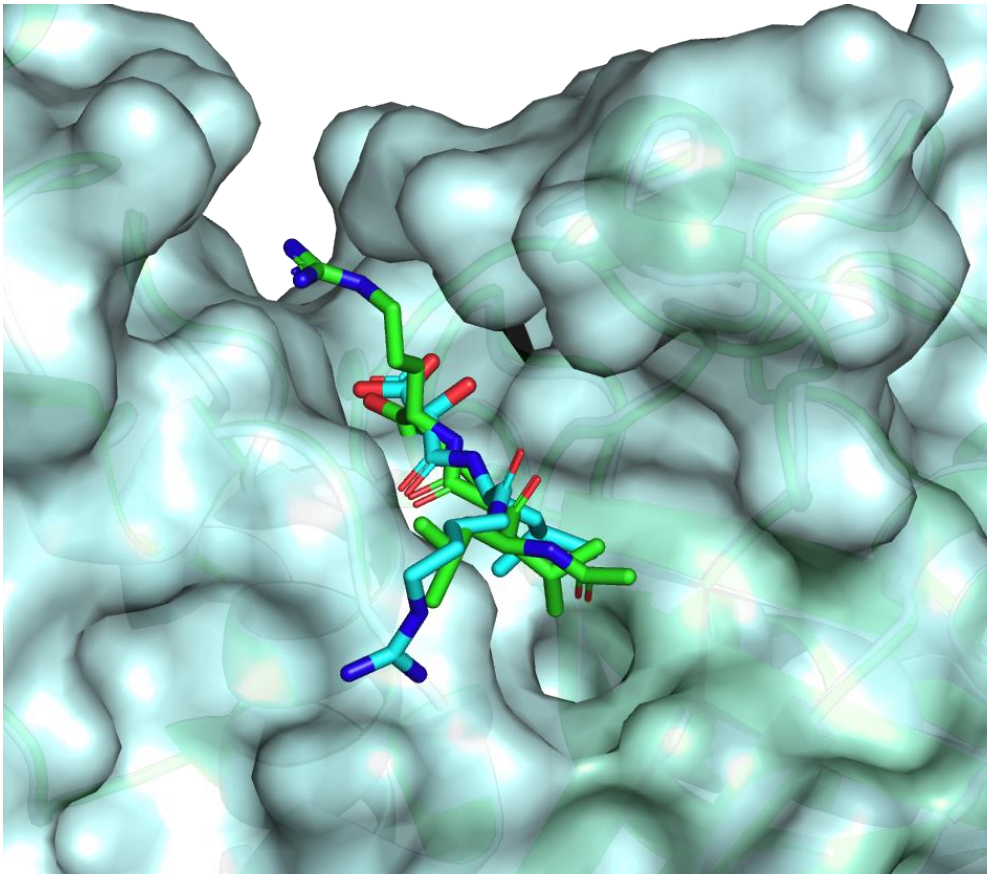
The warhead peptide based inhibitors E64 (**1**, Figure 1.5.5) and leupeptin (**3**, Figure 1.5.6)  bind to the active site cleft of calpain to C115 (Figure 4.1.18). The conformation of the bound inhibitors extends across the S1, S2 and S3 sites flanked by the deep sides of the canyon-like protease core.[104]

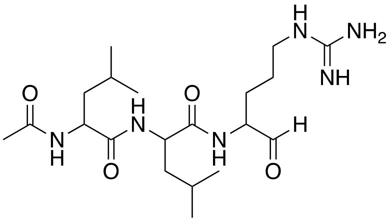


**Leupeptin - 3**

**Mechanism**







**Leupeptin - 3**

**Mechanism**

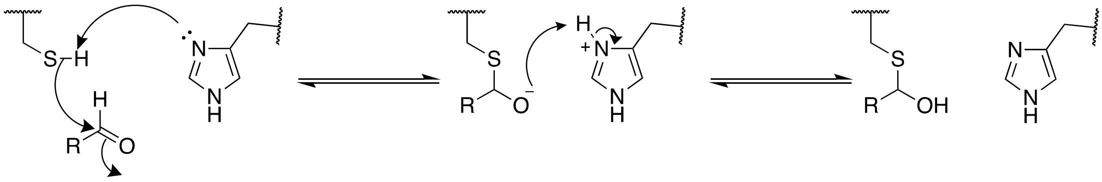


*Figure 4.1.18 - Cartoon representation of the CysPC (grey surface) inhibitor complexes Leupeptin (green sticks ,PDB:1TL9) and E64 (cyan, sticks, PDB:1TLO).*[104]

Extensive salt bridges, hydrogen bonding and hydrophobic van der Waals interactions are responsible for the binding of these inhibitors.[104] The active site plasticity observed upon binding of E64 and leupeptin is postulated to be analogous to the binding of substrates.[104] The toplogy of the CysPC active site celft, bound to leupeptin and E64 (Figure 4.1.18,) was compared to unbound CysPC and CysPC-**35**. The topology of the CysPC-**35** bound active site cleft was not the same as the warhead bound structures. As the diselenide compounds do not feature extended amide moeities, they would not be expected to bind into the active site cleft in the same way as the peptides. Whilst the active site does show preference for hydrophobic residues in the P1 and P2 positions (Table 1.2), the salt bridges and extensive hydrogen bonding interactions reported to be important in binding of the peptide warhead inhibitors are not present in the diselenide compounds. This would not alter the occluded CysPC conformation to expose the C115 residue sufficiently to react and form the selenyl-sulfide adduct.9

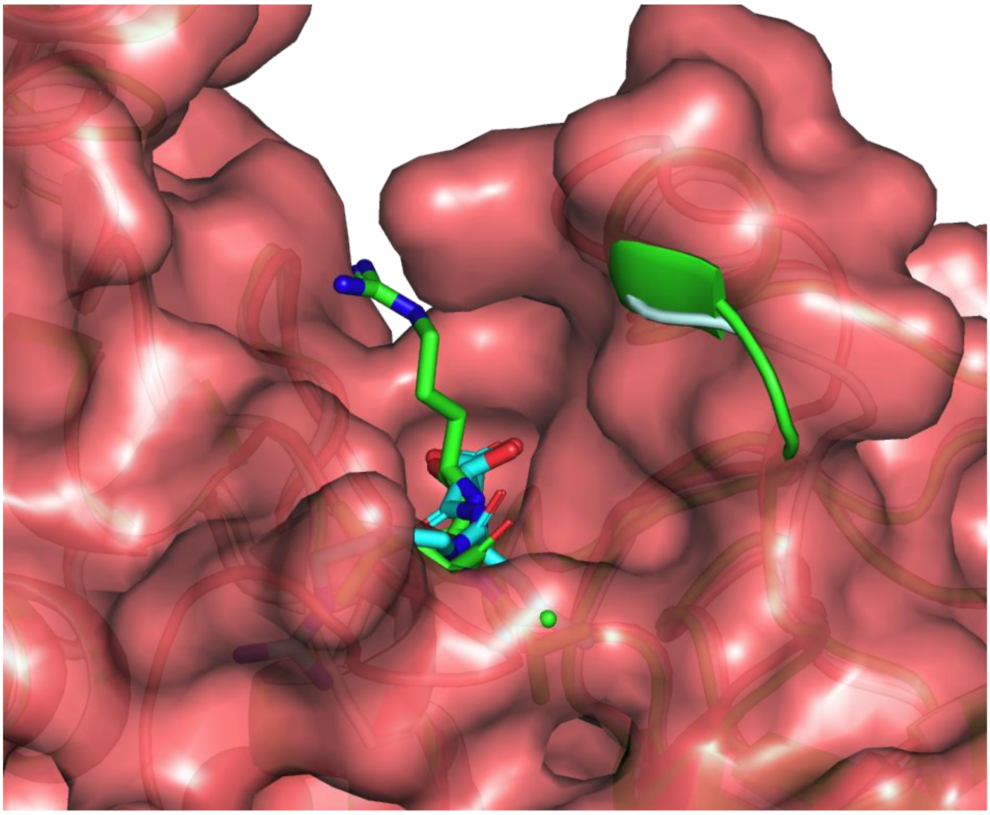


*Figure 4.1.19 - Cartoon representation of the CysPC inhibitor complexes Leupeptin (green sticks, PDB:1TL9) and E64 (cyan, sticks, PDB:1TLO), aligned with CysPC-**35** (red surface).*[104]

*The α-helix between residues 202-218 in the E64 and leupeptin bound structures*

Figure 4.1.20) is a flexbile loop in the unbound active site and CysPC-**35** structures. It is possible that upon substrate binding, these occluding, flexbile gating loops undergo conformational change upon substrate binding to expose the reactive catalytic cysteine residue.

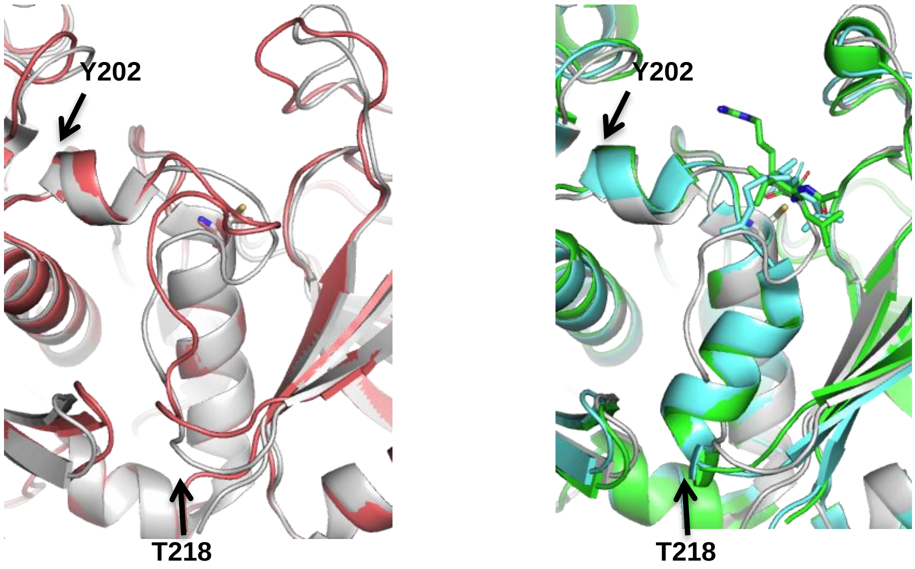


*Figure 4.1.20 - Cartoon representation of the CysPC unbound (grey, PDB: 2ARY), and CysPC-**35** (red) structures (left), and inhibitor complexes leupeptin (green, PDB:1TL9) and E64 (cyan, PDB:1TLO) (right).*[104]

The CysPC conformational changes observed upon binding of the warhead inhibitors provide insight into the possible mechanism of the inhibition of the diselenides. The exposure of the catalytic cysteine residue is therefore limited to substrates that exhibit efficient binding to the active site cleft. While reaction of the diselenides with the catalytic cysteine residue was not observed in the electron density map, it is likely that were this reactive thiolate more exposed, a selenyl-sulfide adduct would form. As calpain is a modulatory protease, it is imperative to prevent the reactive catalytic cysteine residue from participating in nonspecific proteolysis by maintaining substrate specificity.[104]

## 4.2. Summary of CysPC Interactions

The human calpain-1 active site, CysPC, was produced using an *E. coli* expression system and purified by Ni-NTA chromatograpy, followed by TEV cleavage of the poly-histidine-GB1 tag and ‘reverse’ Ni-NTA purification. CysPC presented calcium activated activity in a similar manner to the full length protein and thus was used in the FAM FRET substrate cleavage assay to evaluate the diselenide inhibitors and PD150606 for calpain-1 allosteric inhibitory properties. PD150606 has been reported to inhibit the active site of calpain-1, using the FAM substrate an apparent $IC_{50}$ of 17.8 ± 2.8 μM compared to 14.1 ± 1.6 μM for the full length enzyme, therefore suggesting that this compound is not allosteric in its mode of action. The diselenide inhibitors tested displayed more potent apparent $IC_{50}$ values for the active site, suggesting that they are also not allosteric in their mode of action, despite also binding non-covalently to the calcium binding PEF domains.

The pure protein was crystallised by vapour diffusion and soaked with **35** in order to gain insight into how the diselenide inhibitors interact with CysPC. Significant difference density was apparent on the C49 and C108 residues, which was tentatively assigned to be the selenyl-sulfide adduct formed upon reaction of the inhibitor and solvent exposed cysteine residues. The catalytic residue C115 displayed no difference density associated with the electron density map, suggesting that the inhibitor did not bind to this residue. Comparison of the secondary structure of peptide-based active site inhibitors E64 and leupeptin with the unbound CysPC and CysPC-**35** structure suggest that upon binding the peptide warhead inhibitors, conformational changes expose the active site residue C115 in the cleft to the substrate to partake in catalytic function. The structural features of this process were not observed in the CysPC-**35** structure which suggests that **35** was not able to bind in an analgous manner to peptide based inhibitors, therefore not readily able to react with C115. It is likely that binding to C108, in the active site cleft is responsible for the inhibitory properties of the diselenides, whereby partial obstruction of the cleft results in reduced capacity for substrate binding.

# Chapter 5 - Structure Based Design and Evaluation of Calpain Inhibitors

# 5.1. Structure Based Design and Evaluation of Calpain Inhibitors

*Computational docking experiments carried out by L. Kalash and Dr A. Bender, University of Cambridge, UK.*[103]

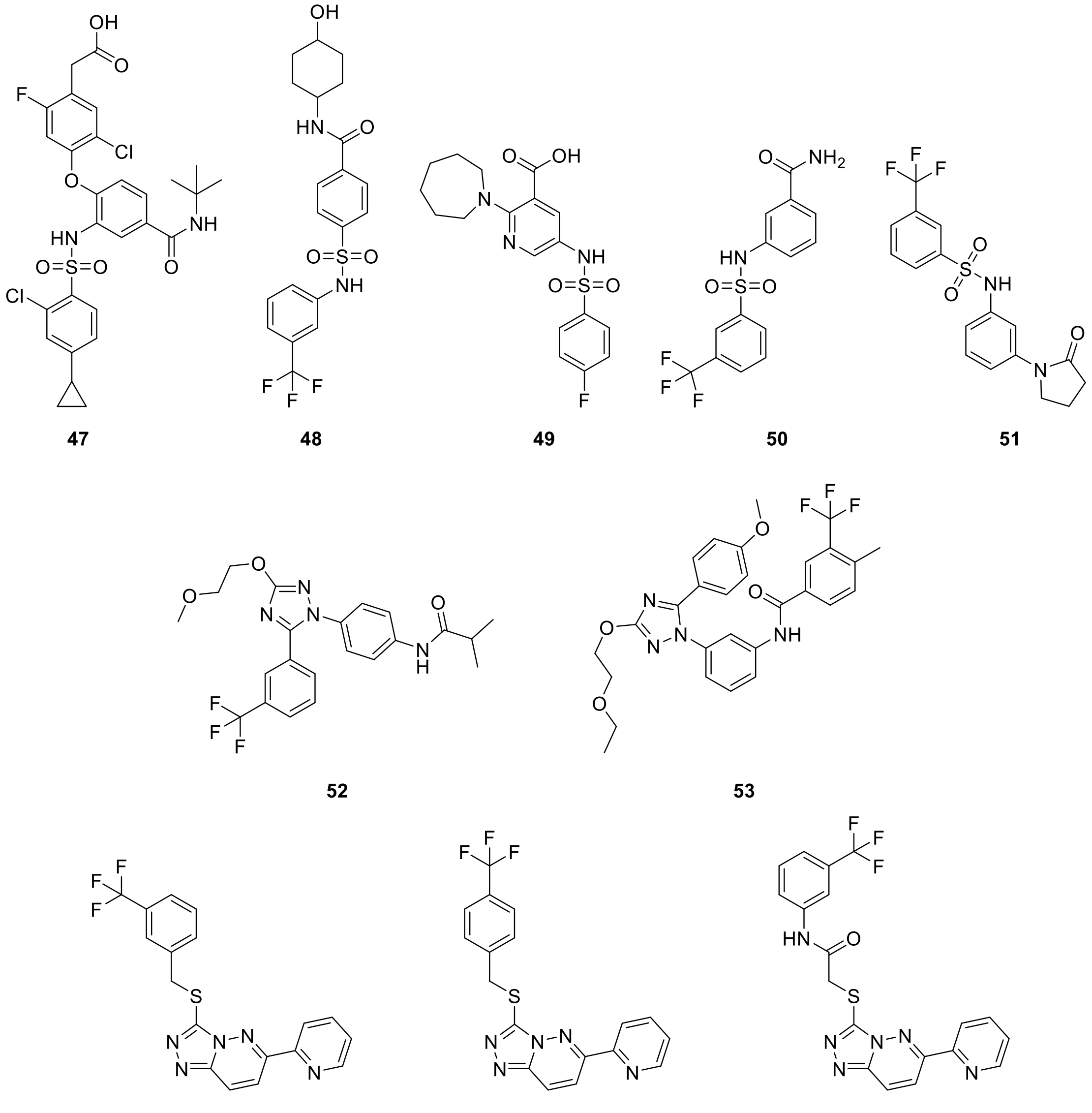
Computational design and docking tools can produce tight binding, potent compounds for a biological target through the optimisation of ligand geometry to the topology of the binding site. The high resolution crystal structure of human PEF(S) provided an accurate 3D model of the hydrophobic pocket in the search for novel calpain-1 PEF(S) binding molecular scaffolds.

## 5.1.1. Docking of computational compounds to PEF(S)

The search for novel inhibitors of calpain-1 *via* binding to the PEF(S) domain was carried out using Glide which is a versatile computational ligand binding tool. The v7irtual screening of 36,503 purchasable ligands were docked using Glide into the crystal structure of human PEF(S) (PDB:4WQ2).[196] A distribution of docking scores for the active and inactive compounds was obtained, and a shortlist of candidates for PEF(S) binding were further screened against PAINs, to filter out Pan-Assay Interference Compounds – common false positives.[197] These were then screened for absorption, distribution, metabolism, excretion and toxicity (ADMET) properties using FAF-Drugs2 tool.[198] This method was used to evaluate compounds for drug-like properties.

Early attempts at carrying out accurate predictions based on structures derived the Rule of 5 (Ro5) by Dr Lipinski. Lipinski analysed the physiochemical properties of 2000 drug candidates beyond phase II clinical trials and predicted that poor absorption is more likely when a molecule contains more than 5 H-bond donors, 10 H-bond acceptors, has a molecular weight above 500 or the calculated logP (partition coeffiction, the ratio of a compound between a hydrophobic and hydrophillic phase) is greater than 5.[199] Interactions with drug transportation systems are not considered in the Ro5, only accounting for absorption and diffusion throughout the body. The Ro5 is only applicible to orally adminstered drugs, which is the preferred method of drug delivery allowing patients to noninvasevely self-administer over potentially traumatic intravenous treatments.[200] Natural products are known to ignore the physiochemical restraints of the Ro5 as they occupy a diverse and complex chemical space and these molecules are often able to interact directly with biomolecules in a highly specific fashion such as through the use of native molecular recognition systems.[201]

The compounds that passed the screening process were five sulfonamides (**47**-**51**), two substituted 1,2,4-triazoles (**52**,**53**), and three substituted 6-(pyridin-2-yl)-[1,2,4]triazolo[4,3-*b*]pyridazine scaffolds (**54-56**).



**47** **48** **49** **50** **51**

**52** **53**

**54** **55** **56**

*Figure 5.1.1 – Potential PEF(S) binding compounds **47**-**56** selected by passing through computational docking screening filters by the Bender research group (Cambridge University, UK).*[202]

The docking studies predicted favourable molecular interactions of the compounds with several residues in the vicinity of the binding site on the PEF(S) domain. The 2-chloro-4-cyclopropylsulfonamido phenyl ring of sulfonamide **47** displayed π-stacking interactions with His131, whilst the carbonyl of its carboxylic acid was predicted to form H-bonds with the same residue (Figure 5.1.2, **A**). The carbonyl of its amide moiety was expected to form H-bonds with Trp168, and the phenyl ring attached to the tert-butylcarbamoyl moiety π-stacked with the same residue.

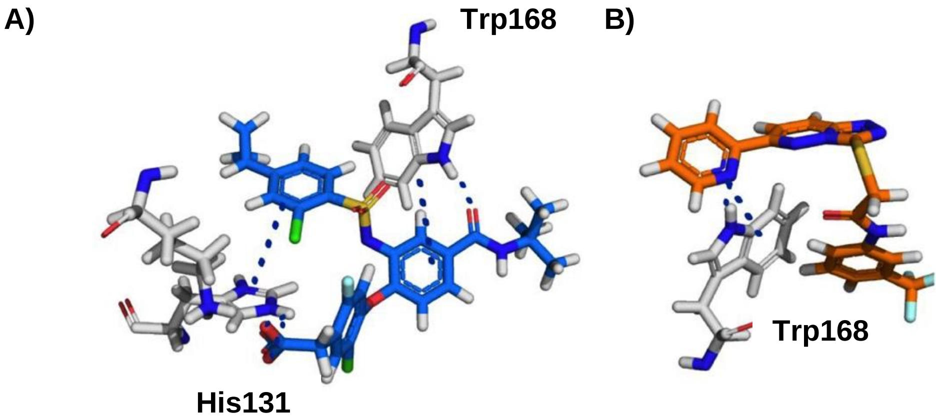


*Figure 5.1.2 - Predicted molecular interactions of compounds **47** and **56** with human calpain-1 PEF(S) of the regulatory subunit using the crystal structure (PDB:4WQ2). **A)** The sulfonamide **47** (blue) showing key binding interactions with PEF(S) residues His131 and Trp168. **B)** [1,2,4]triazolo [4,3-b]pyridazin-6-yl]pyridine compound **56** (orange) showing key binding interactions with PEF(S) Trp168 residue.*[202]

Compound **56** displayed a π-stacking interaction between the pyridine ring and Trp168. Significant interaction with Trp168 was deemed important in the binding of small molecules to the PEF(S) hydrophobic pocket as observed with the diselenides and α-mercaptoacrylic acids. Sulfonamides **48**-**51** were predicted to display possible further interactions with the flexible hydrophilic residues Glu97, Gn100, and Lys172.

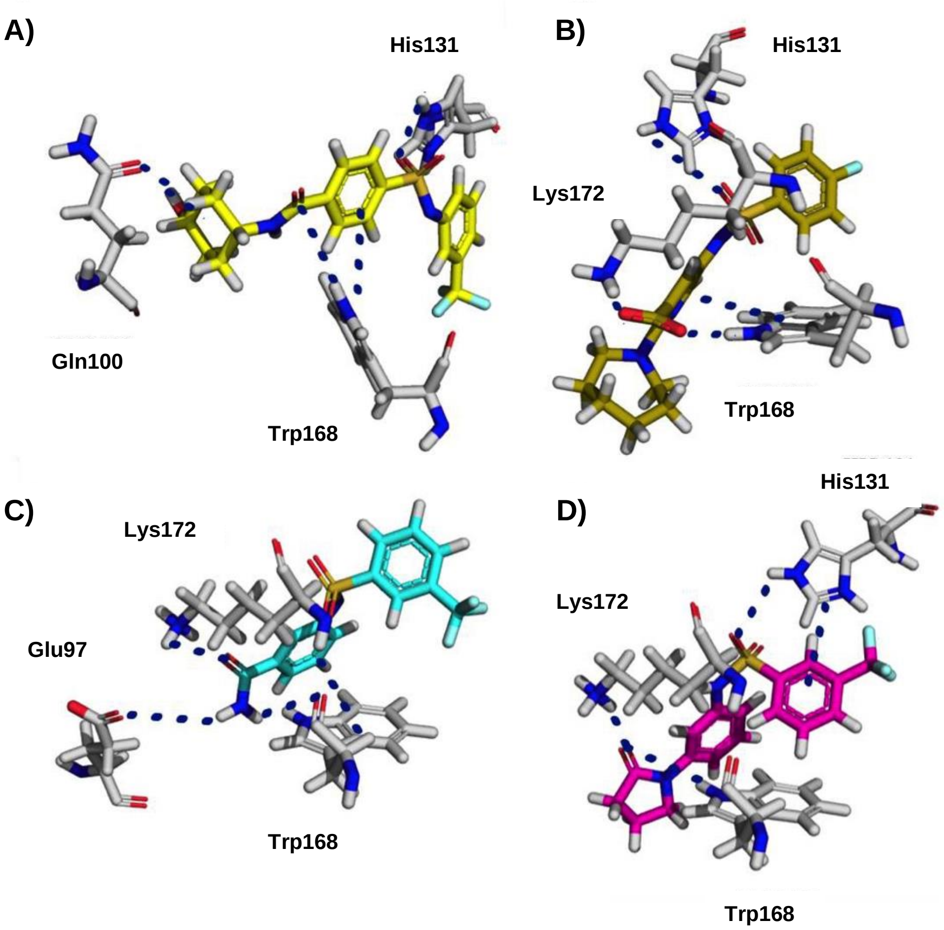


*Figure 5.1.3 - Predicted molecular interactions sulfonamide **48**-**51** with human calpain-1 PEF(S) of the regulatory subunit using the crystal structure (PDB:4WQ2). **A)** The sulfonamide **48** (bright yellow) showing key binding interactions with PEF(S) residues Gln100, His131 and Trp168. **B)** The sulfonamide **49** (dark yellow) showing key binding interactions with PEF(S) residues His131, Trp168 and Lys172. **C)** The sulfonamide **50** (cyan) showing key binding interactions with PEF(S) residues Glu97, His131 and Trp168. **D)** The sulfonamide **51**(magenta) showing key binding interactions with PEF(S) residues His131, Trp168 and Lys172.*[202]

The hydrophobic π-stacking interactions with Trp168 were maintained in all ligand docking experiments, whilst other hydrophilic PEF(S) interactions were also explored in an attempt to exploit other potential binding interactions. Sulfonamide **48** contains a cyclohexyl ring bearing a hydroxyl group, predicted to form a hydrogen bond with Gln100, as well as hydrogen bonding between the sulfonyl group and His131 (

Figure 5.1.3, **A**). Compound **49** was expected to form hydrogen bonds between the carboxyl moiety and Lys172 and Trp168 as well as hydrogen bonds from the sulfonyl group to His131 (

Figure 5.1.3, **B**). Sulfonamide **50** presented hydrogen bonding interactions between its amide group and PEF(S) residues Glu97 and Trp168, whilst the carbonyl group with Lys172 (

Figure 5.1.3, **C**). Hydrogen bonding as predicted to occur between the carbonyl group of the oxopyrrolidin-1-yl ring present on compound **51** and the Lys172 and Trp 168 residues (

Figure 5.1.3, **D**). H-bonding between the sulfonyl moiety and His131 was predicted, as well as π-stacking with the same residue and the aromatic ring of the trifluoromethylphenyl group.

All the docked compounds presented a diverse and novel chemical space for calpain inhibition, with the potential to explore new binding interactions with residues situated around the PEF(S) calpastatin binding site. They also did not contain any readily redox sensitive disulfide or diselenide groups, suspected to complicate the mode of inhibition of those inhibitors. Thus, the TNS displacement and substrate cleavage assays were used to experimentally validate if the docking protocol could identify allosteric inhibitors.

### 5.1.2. TNS displacement assay

The TNS displacement assay, developed with known PEF(S) binder PD150606 (**9**), was also used to determine if the computationally docked compounds exhibited any PEF(S) binding interactions. The assay unveiled that 4 of the sulfonamides bound to PEF(S) with moderate affinity between 38.9 – 48.3 μM. Compounds **51-56** interfered with the wavelength range required for the assay, therefore measuring $K_d$ values was not possible by this method.

| **Compound** | **PEF(S) $K_d$ (μM)** |
|---|---|
| **47** | 46.8 ± 11.9 |
| **48** | 48.3 ± 6.8 |
| **49** | 38.9 ± 3.9 |
| **50** | 39.6 ± 9.5 |
| **51** | *N/A* |
| **52** | *N/A* |
| **53** | *N/A* |
| **54** | *N/A* |
| **55** | *N/A* |
| **56** | *N/A* |

*Table 5.1 - TNS $K_d$ values calculated from TNS displacement from PEF(S) with compounds computationally screened by L. Kalash and A. Bender, University of Cambridge, UK.*

The results from this assay show that some compounds do bind to the allosteric site on the calpain-1 PEF(S) regulatory subunit with moderate affinity.

### 5.1.3. FRET based inhibition assay

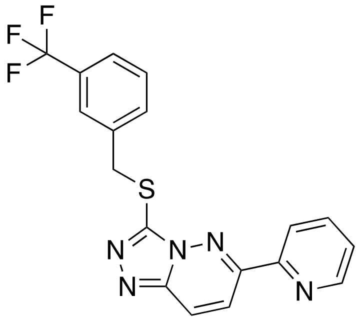
The calpain-1 FRET substrate cleavage assay (Section 2.3) was used to determine the apparent $IC_{50}$ values of compounds **47**-**56** for the full-length calpain-1 heterodimeric complex. Amongst the identified compounds, **47**, **55**, and **56** inhibited the calpain-1 with apparent $IC_{50}$ values of 7.5, 20.5, and 29.7 µM, respectively. The compounds were also tested against calpain-1 CysPC, only **47** showed some weak inhibition with an $IC_{50}$ much greater than 100 µM, suggesting these compounds are allosteric in their mode of action. In contrast, **48** exhibited an $IC_{50}$ value of 41.1 µM for CysPC, compared to the full-length calpain-1 complex, which displayed an $IC_{50}$ of >100 µM, suggesting this compound is more selective for the calpain-1 active site.

| **Compound** | **Full length calpain-1 $IC_{50}$ (µM)** | **Calpain-1 CysPC $IC_{50}$ (µM)** |
|---|---|---|
| **47** | 7.5 ± 1.1 | >100 |
| **48** | -- | -- |
| **49** | >100 | 41.1 ± 15.4 |
| **50** | -- | -- |
| **51** | -- | -- |
| **52** | -- | -- |
| **53** | -- | -- |
| **54** | -- | -- |
| **55** | 20.5 ± 1.9 | -- |
| **56** | 29.7 ± 5.2 | -- |

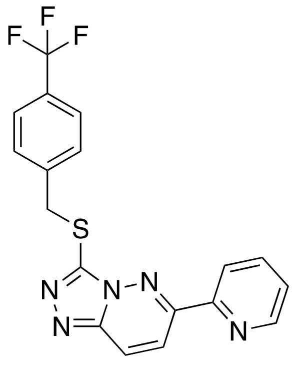
*Table 5.2 – Apparent $IC_{50}$ values for compounds **47-56** for the inhibition of cleavage of FAM substrate by calpain-1 and CysPC under standard assay conditions (Section 6.8.1).*

Compound **47** or AMG 853 has been evaluated in clinical trials for asthma, targetting two GCPRs; prostanoid D receptor and chemoattractant receptor-homologous molecule.[203] It was deemed safe but insufficiently effective at improving asthma symptoms or lung function therefore trials were not continued.[203] In the calpain-1 activity assays, it exhibited a lower $IC_{50}$ than PD150606, by an apparent allosteric mode of action. **47** was reported to have anti inflammatory properties which could possibly arise from being an allosteric inhibitor of calpain-1.[204] Calpain has been reported to play a critical role in inflammation as well as having links to asthma.[22] Potentially, this finding could highlight the importance of considering calpain inhibition for the further development of new anti-asthmatic therapies.

Interestingly **54** did not inhibit calpain-1 whilst **55** displayed an apparent $IC_{50}$ of 20.5 ± 1.9 μM, the only structural difference between these two compounds is the trifluoromethyl group meta (**54**) or para (**55**) to the thioether moiety.



**54**



**55**

*Figure 5.1.4 - Compounds identified by docking studies to bind to PEF(S).*

These compounds interfered strongly with the fluorescence wavelengths used in the TNS binding assay, so $K_d$ values were not determined for PEF(S), however it is evident that the trifluoromethyl position is crucial for the activity of the [1,2,4]triazolo[4,3]pyridazinyl] thioether based inhibitors. Further exploration of the SAR of this class of compounds, supported by structural biological experiments, may elucidate the reason for this difference in activity. Exploitation of this structure-activity relationship may provide a novel pharmacophore for the development of highly specific calpain inhibitors.

## 5.1.4. PEF(L) binding

The inhibitors that displayed allosteric modes of action for calpain-1, **47**, **55** and **56** were tested for binding to PEF(L). As **55** and **56** interfered with the fluorescence assay wavelengths, only **47** could be tested for binding to PEF(L) by this assay.

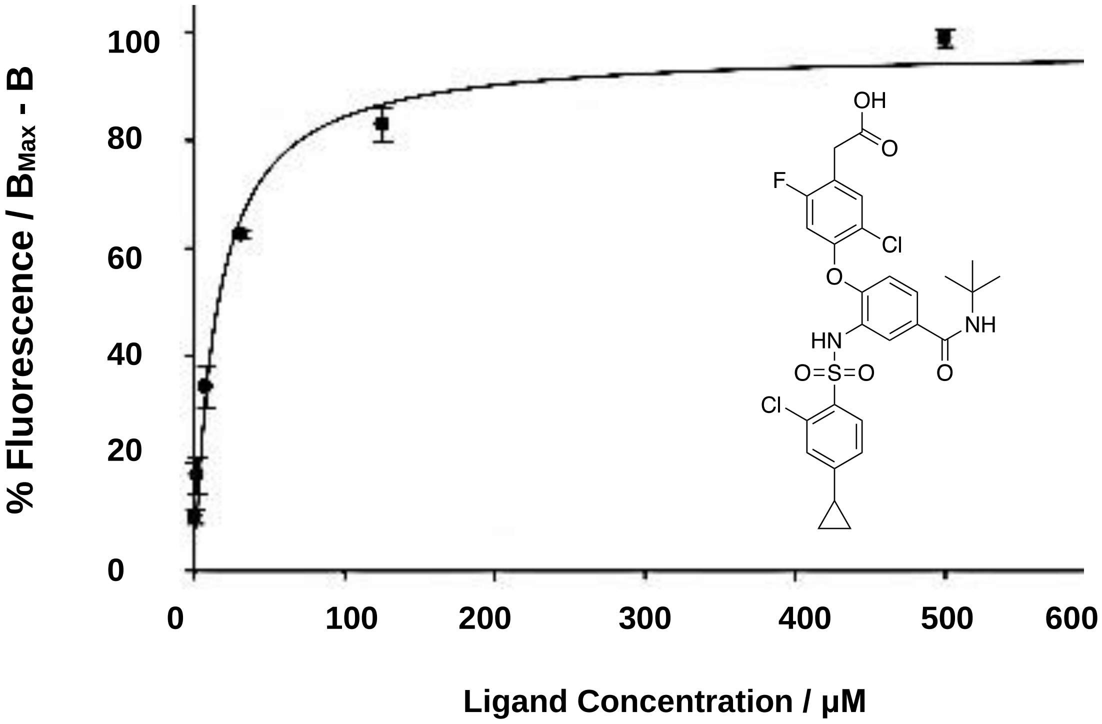


*Figure 5.1.5 - TNS displacement assay on PEF(L) with* ***47*** *measuring the change in fluorescence at different concentrations of ligand, the peak maxima absorbance was normalised and inverted then analysed by non-linear regression using the Ligand Binding module on Sigmaplot (Systat Software, San Jose, CA).*

The calculated $K_d$ value for **47** binding to PEF(L) was 14.9 ± 2.3 µM, which was approximately 3x lower than the $K_d$ of 46.8 ± 11.9 µM for PEF(S). The potent inhibitory allosteric effect observed from this compound could be a result of binding to PEF(L) which is a novel concept in calpain inhibition, and one that should be supported with structural data. Calpastatin fragments A and C bind to PEF(L) and PEF(S) respectively and are reported to be activators of calpain when added to calpain-1 and -2 as separate 19-mer oligopeptides.[96] Whilst an inhibitory effect is observed through PEF(L) and PEF(S) binding of **47**, the potential for both the calcium binding domains to modulate the activity of the full length heterodimeric complex is evident.

### 5.1.5. Computational assessment of CNS permeability of **47** and **56**

Calpain-1 is a therapeutic target for for neurodegenerative disorders and in order to target such diseases, compounds must be able to cross the blood-brain barrier effectively. FAFDrug2 was used to computationally assess the CNS permeability of compounds **47** and **56** to provide an indication of their viability as a potential drug for such diseases.[198] The physiochemical properties of either compound was compared to 5 key parameters linked to a compound's ability to efficiently cross the blood brain barrier. These parameters are molecular weight (MW), hydrogen bond acceptors/donors (HBA/HBD), the topological polar surface area (tPSA) and the logP.

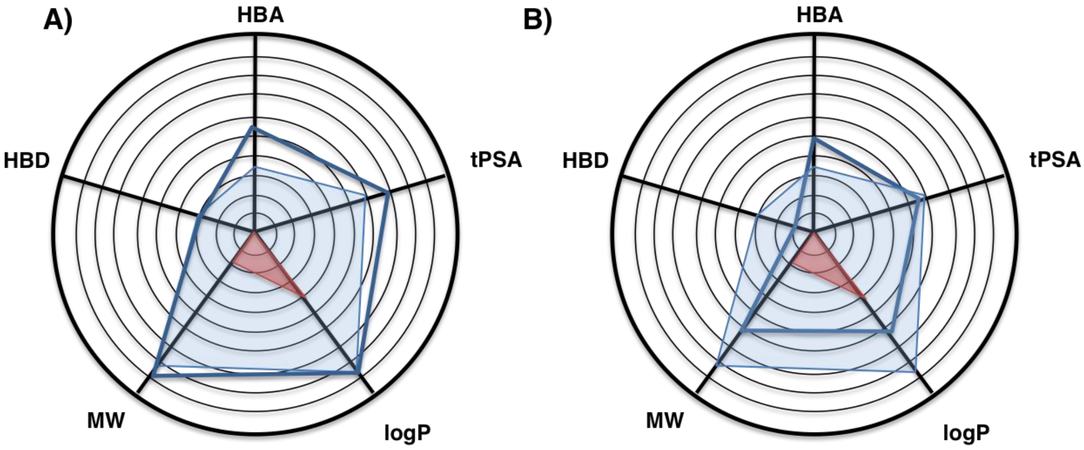


*Figure 5.1.6 – Diagram representing the assessment of sulfonamide **47** (**A**) and [1,2,4]triazole **56** (**B**) for CNS permeability, taking into account molecular weight (MW), hydrogen bond donors/acceptors (HBD/HBA), toplogical polar surface area (tPSA), and the logP. The light blue area represents the CNS filter, inside the red or outside this area compounds are not expected to pass the blood brain barrier.*

Sulfonamide **47** did not pass the CNS filter since all parameters exceeded that of the filter or were on the upper limit of each characteristic, therefore **47** was not expected to exhibit blood brain barrier permeability. [1,2,4]triazole **56** was expected to display medium CNS permeability since all attributes except HBA fell well within the limits. This compound could serve as a starting point for analogue development of allosteric calpain inhibitors, targetting neurodegenerative diseases such as AD.[90]

## 5.2. Calpain-2 Inhibition

### 5.2.1. Calpain-2 AMC substrate cleavage assay

A calpain-2 substrate cleavage assay was used to assess the isoform selectivity of the calpain inhibitors. The calpain-2 substrate used was the 7-amido-4-coumarin (AMC) fluorophore (Figure 5.2.1). The Leu-Leu-Val-Tyr peptide substrate linked to succinic acid at the N-terminus and AMC at the tyrosine C-terminus. Calpain-2 was activated with $CaCl_2$ which hydrolyses the amide bond between the fluorophore and Tyr and relieves the quenched AMC fluorescence, which is then detected at 460 nm.

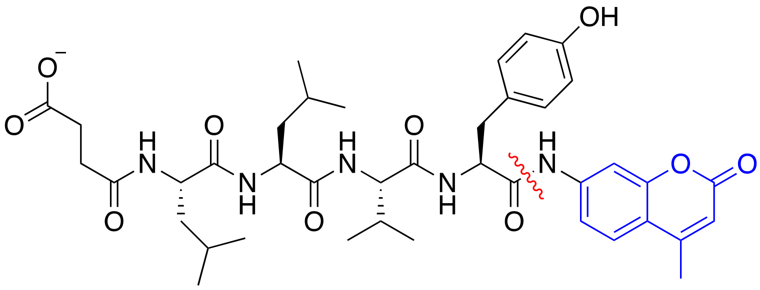


*Figure 5.2.1 - Fluorogenic substrate for calpain-2 assay Suc-Leu-Leu-Val-Tyr-AMC, with the AMC moeity highlighted in blue and the cleavage site highlighted in red.*

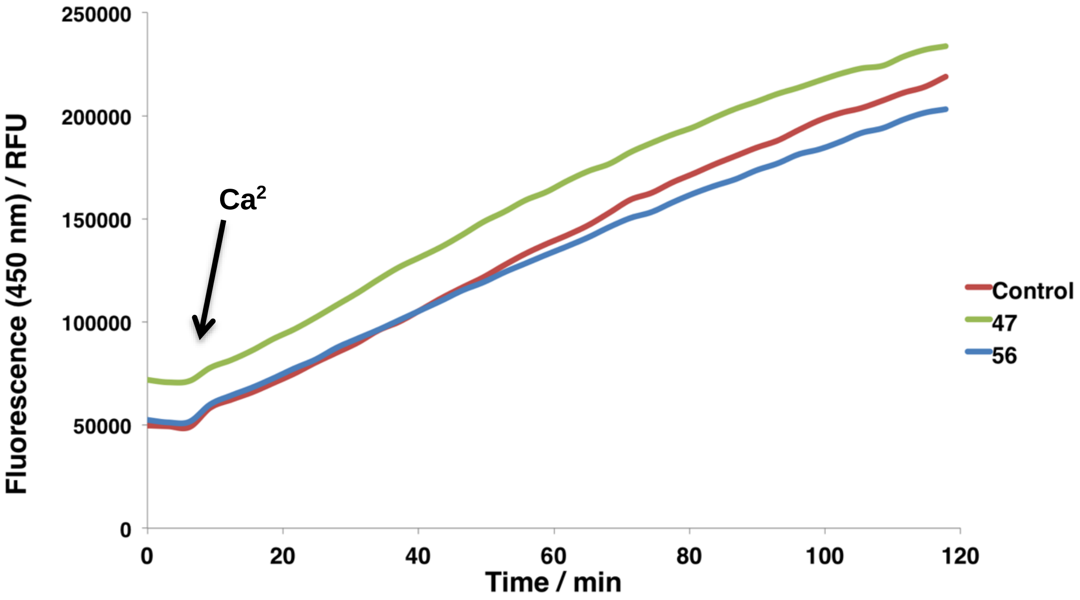


*Figure 5.2.2 - Calpain-2 inhibition assay monitoring the change in fluorescence at 450 nm. Assays were performed with the calpain-2 AMC FRET substrate (1 μM), calpain-2 (25 nM), HEPES (10 μM), glutathione 9:1 (reduced : oxidised) 10 (μM), EDTA (0.5 mM), BSA (0.1%) at pH 7.0. Addition of $CaCl_2$ (5 mM) initiated the reaction, compounds were added to a final concentration of 100 μM.*

The AMC-based calpain-2 assay was used to investigate inhibitory activity of the computationally designed compounds **47**-**56**. Each compound was tested at a concentration of 100 μM, as a preliminary investigation to see if inhibition of the calpain-2 isoform was observed.

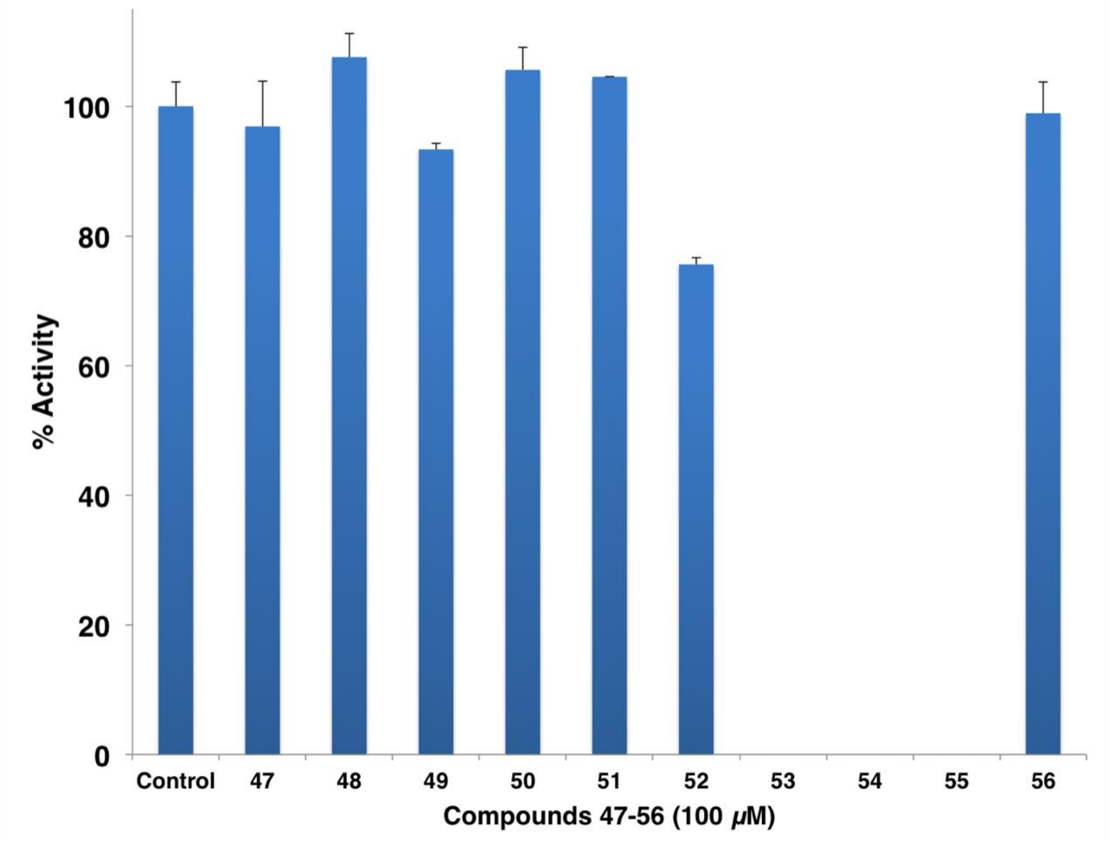


*Figure 5.2.3 - Calpain-2 activity assay with the Suc-AMC substrate showing the relative activities of compounds* ***47-56*** *at 100 μM, tested in triplicate with the initial linear portion of the gradient taken as the relative rate.*

None of the computationally docked compounds displayed significant calpain-2 inhibitory activity at 100 μM concentration. Strong fluorescence of compounds **53**-**55** within the wavelength range of the AMC substrate used in the assay prevented accurate analysis using this protocol. Notably, **47** did not appear to inhibit calpain-2 despite displaying a potent calpain-1 inhibition (apparent $IC_{50}$ of 7.5 ± 1.1 μM), with an apparent allosteric mode of action. Allosteric calpain-1 inhibitor **56** also did not inhibit calpain-2 appreciably, thus displaying impressive isoform selectivity. The differential binding of **47** and **56** to PEF(L) over PEF(S) could be the source of selective inhibition of these compounds, however it is also possible that binding to another calpain domain, such as CBSW, could also be the source of inhibition.

### 5.2.2. Diselenide inhibition of calpain-2

Several of the diselenide inhibitors were also investigated for isoform selectivity using the calpain-2 AMC assay, at a concentration of 100 μM.

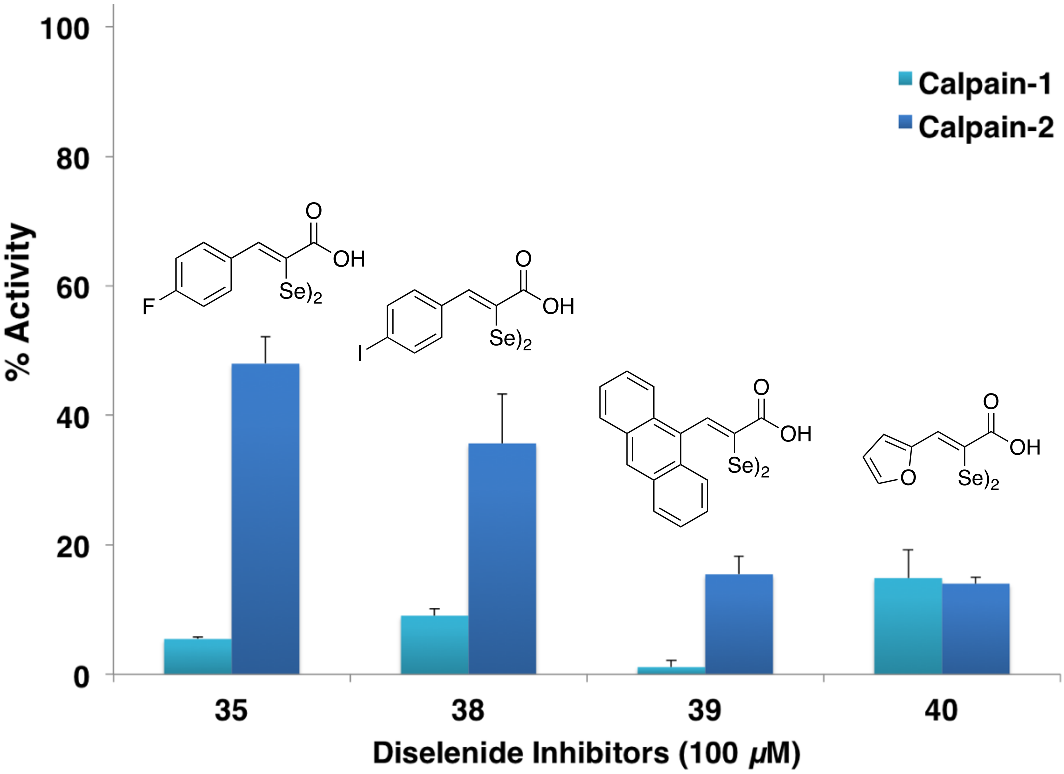


*Figure 5.2.4 – Calpain-1 activity with FAM substrate, calpain-2 activity assay with the Suc-AMC substrate and diselenide inhibitors **35, 38, 39** and **40** at 100 μM, using the assay conditions described (Section 6.8.1).*

The two phenyl based inhibitors **35** and **38,** at a concentration of 100 μM, reduced the activity of calpain-2 to 47.8% and 35.6% respectively. These values were greater than that observed for calpain-1, whereby a reduction of proteolytic activity to 5.5% and 9.1% was acheived respectively. The anthracene based diselenide **39** inhibited calpain-2 more potently than the phenyl based diselenides **35** and **38**. It is evident that these compounds do not inhibit each calpain isoform with the same potency, however remain good inhibitors of calpain-2. Interestingly, furan based diselenide **40** displayed no apparent isoform selectivity, which could be due to the smaller aromatic ring presenting less steric bulk on the active site domain.

### 5.2.3. Evaluation of inhibitors

The synthesised diselenide inhibitors displayed low micromolar apparent $IC_{50}$ values for calpain-1. The compounds bind to the regulatory subunit, *via* the PEF(S) domain, which is present on both calpain-1 and -2, as observed in protein-ligand co-crystal X-ray diffraction data. A decrease in fluroescence at 450 nm, attributed to the displacement of the TNS fluorescent probe, suggests that both the α-mercaptoacrylic acids and diselenoacrylates bind to the calcium binding domain present on the large subunit of calpain-1, PEF(L) with similar affinity to PEF(S).

The diselenides displayed poor isoform selectivity, inhibiting calpain-2 signifcantly. These compounds were also shown to inhibit the active site, CysPC, to a greater extent than the full length enzyme, suggesting that binding to this domain is the source of inhibition. As the CysPC domains are not identical between the calpain isoforms it is likely the mode of inhibition is likely analogous for calpain-2. Despite the observed non-covalent allosteric binding interactions with the PEF(S) and PEF(L) domains, covalent binding to active site cysteine residues is likely to have a more significant effect on the differential inhibition. As this effect is produced by binding to the active site, it would be of little therapeutic interest as interaction with other cysteine proteases would likely produce undesired off target effects.[40]

The *in silico* compounds **47**, **55** and **56** did not significantly inhibit the active site of calpain-1 at concentrations up to 100 μM, thus displaying an apparent allosteric mode of action. Compound **47** was shown to bind to PEF(L) and as well as the PEF(S) domain, **55** and **56** interfered within the wavelengths used for the TNS displacement assay therefore evaluation by this method was not possible. Binding measurements of these compounds could be further explored by other techniques such as isothermal titration calorimetry (ITC), although was not carried out for this project.

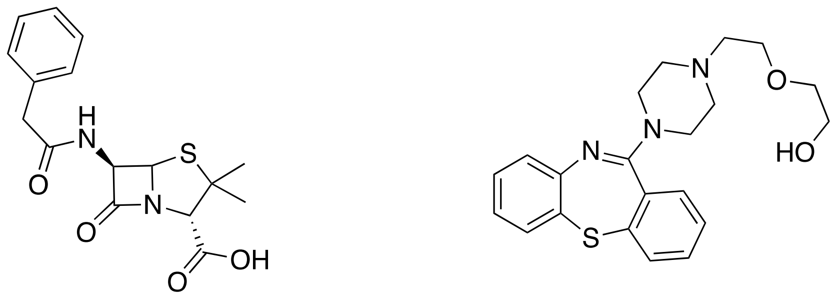
The compounds discovered by docking experiments for PEF(S) displayed much more desirable drug-like characteristics than the diselenide inhibitors as potent, selective allosteric inhibitors of calpain-1. For example instead of a redox sensitive thiol group, compound **47** contained a sulfonamide. Sulfonamides are reported to be relatively unreactive and are present in many drugs including sulphamethoxazole

(**57**, Figure 5.2.5) to treat community-acquired methicillin-resistant *Staphylococcus aureus* infections, and hydrochlorothiazide (**58**, Figure 5.2.5) for hypertension.[205,206]

Sulfamethoxazole **- 57** Hydrochlorothiazide **- 58**

*Figure 5.2.5 - Two sulfonamide containing drugs, sulfamethoxazole (**57**) and hydrochlorothiazide (**58**).*

Thioether compounds are also widely used in pharmacological scaffolds, a survey in 2014 found that of twelve disease categories studied, 8.8% of 1969 drugs studied contained the thioether moeity, of these include the antibiotic penicillin (**59**) and antipsychotic agent Seroquel (**60**).[207]



Penicillin G - **59** Seroquel - **60**

*Figure 5.2.6 - Thioether containing drugs, Penicillin G (**59**) and Seroquel (**60**)*

Drugs containing diselenide groups are not currently FDA approved, they have been explored for prodrug release mechanisms, highlighting their instability in the cytoplasm however further data is required before these compounds could be confidently to treat patients, the selectivity of the compounds would also be a concern. By contrast the calpain inhibitors discovered using the structure guided docking experiments presented novel scaffolds containing established, stable pharmacophores, ideal for structure activity relationship studies to further the development of isoform specific, allosteric calpain inhibitors.

# Chapter 6 - Full Length Calpain-1 Expression

# 6.1. Full Length calpain-1 expression

A complete crystal structure of the heterodimeric complex would enable mechanistic insight into the different $Ca^{2+}$ concentrations required for activation of calpain isoforms, and the physiological consequences.[23] Solving the X-ray crystal structure would also likely accelerate the development of isoform selective inibitors by providing a 3D model of the target protein.[208] The expression and purification of full length calpain-1 was investigated for crystallography experiments.

## 6.1.1. Expression and purification of CAPN1 C115S + CAPNS1ΔGR

Two separate non-covalently bound proteins are required to be simultaneously expressed for the production of heterodimeric full-length calpain-1. CAPN1 C115S was ligated into a pET24b(+) vector with a C-terminal polyhistidine-tag, which also contained a gene coding for kanamycin resistance, whilst CAPNS1ΔGR was ligated into a pAC-pET24 vector with a gene coding for ampicillin resistance.[209] The calpain plasmids were supplied by Hata *et al*. (Department of Advanced Science for Biomolecules, Tokyo, Japan).[209]

C41 (DE3) pLysS cells containing the CAPN1 C115S and CAPNS1ΔGR (CAPNS1 with truncated GR domain as described)[209] genes were grown at 37 °C in kanamycin and ampicillin selective TB(Enhanced) media (1 L) until $OD_{600}$ = 0.9 then induced with IPTG to a final concentration of 1 mM. The protein was expressed overnight at 24 °C and cells harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc. MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 min at 4 °C. The cells were re-suspended in 20 mM HEPES, 0.5 mM EGTA, 0.5 mM TCEP pH 7.6 (buffer A, Section 6.6.8) and lysed by sonication for 5 min (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 min at 30310 rcf in a Sorvall RC6 Plus centrifuge. The supernatant solution was passed through a 0.2 μm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 column volume (CV) buffer A and eluted with 10 CV buffer A supplemented with 250 mM imidazole. The eluted fraction was then applied to a MonoQ HR10/10 column (GE Healthcare) pre-equilibrated with buffer A. The protein was eluted with a linear gradient of 0-0.5 M NaCl in buffer A over 7 CV. The purity of the peak fractions was assessed by SDS-PAGE (10%).

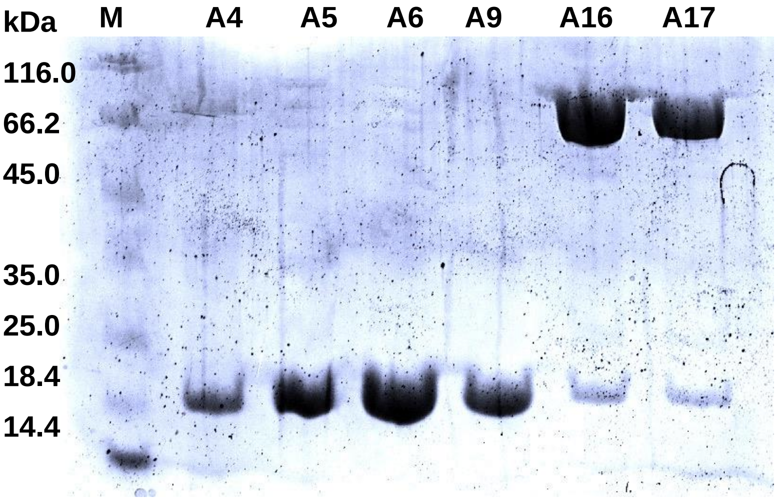


*Figure 6.1.1 – SDS-PAGE of anion exchange chromatography of CAPN1C115S + CAPNS1ΔGR with MonoQ HR10/10 column.*

The peak fractions A16 and A17 showed bands indicating the presence of both CAPN1C115S and CAPNS1ΔGR in comparative ratios to those reported for recombinant heterodimeric calpain.[209]

### 6.1.2. Circular dichroism

A CD spectra of CAPN1C115S + CAPNS1ΔGR (10 μM) was recorded between 190-300 nm with 1 nm steps at 20 °C in $K_iPO_4$ (10 μM, pH 7.4),

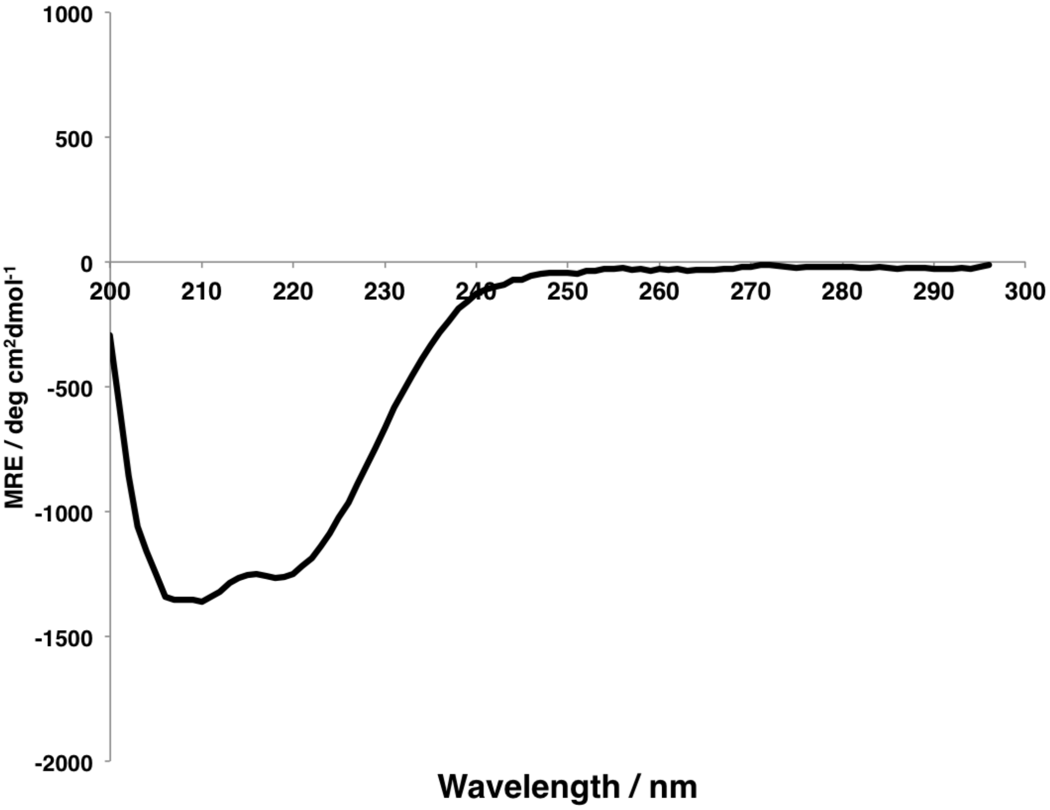


*Figure 6.1.2 - CD spectra of human purified CAPN1C115S + CAPNS1ΔGR, mean residue ellipticity calculation used (Equation 3).*

The CD spectra showed that the protein exhibited primarily α-helical structure in solution after the purifcation steps.

### 6.1.3. Crystallisation attempts of CAPN1 + CAPNS1ΔGR

Crystallisation trials using CAPN1 C115S + CAPNS1ΔGR (10 mg mL$^{-1}$) on PACT-*premier,* JCSG++, LMB and ProPlex screens produced several hits (Table 6.1); using the sitting drop crystallography plate set-up procedure described for CysPC (Section 4.1.8).

| Screen / Well | Precipicant Components | Resolution / Unit cell Parameters of Dataset | Comments |
|---|---|---|---|
| **PACT-*premier* D06** | 0.1 M MMT<br>25 % w/v PEG 1500<br>pH 9.0 | Res(High) = 2.3 Å<br>a = b = c = 147.61 Å | PEF(S)-Hfq |
| **LMB G07** | 0.05 M Sodium cacodylate<br>0.09 M Ammonium sulphate<br>13 % w/v PEG 8000<br>pH 6.5 | Res(High) = 4.5 Å<br>a = 105.44, b = 105.44, c = 552.50 | Unsolved |
| **JCSG++ C06** | 0.1 M Potassium phosphate citrate<br>40 % v/v PEG 300<br>pH 4.2 | Res(High) = 2.9 Å<br>a = 50.90, b = 110.47, c = 111.52 | Unsolved |
| **JCSG++ C09** | 0.1 M Sodium potassium phosphate<br>25 % v/v 1,2-Propanediol<br>10 % v/v Glycerol<br>pH 6.2 | Res(High) = 2.8 Å<br>a = 50.87, b = 110.71, c = 110.69 | Unsolved |

*Table 6.1 – Hits found from crystallisation screening using using the PACT-premier, JCSG++ and LMB screens.*

### 6.1.4. PEF(S) – Hfq co-crystal structure

The resulting crystals from PACT-*premier* well D06 were in the $P2_13$ space group, with a diffraction higher limit of 2.3 Å in resolution, (Table 6.2). The structure was determined by molecular replacement, using PHASER and MOLREP. PDB entry 1DF0 was used as a starting model, however the solution would not reproduce the anticipated large subunit calpain-1 domains.[179,210] The unit cell of the dataset presented was large enough to accommodate the full length structure compared with the previously published calpain-2 structure.[18] After initial modelling, molecular replacement and refinement of various domains from the previously reported μ-like calpain (PDB:1QXP) and calpain-2 structure (PDB:1DF0), the $R_{free}/R_{work}$ remained high, suggesting an incorrect solution.

| Data Reduction Statistics | PEF(S)-Hfq structure |
|---|---|
| **Unit cell dimensions / Å** | a = b = c = 147.61 Å |
| **Unit cell angles / °** | α = 90.00, β = 90.00, γ = 90.00 |
| **Space group** | $P2_13$ |
| **Resolution range / Å** | 2.32-66.02 |
| **Total reflections** | 1,029,720 |
| **Unique reflections** | 46,565 |
| **Completeness / % (last shell)** | 100.0 (100.0) |
| **I / σ (last shell)** | 19.2 (3.0) |
| **R(merge) / % (last shell)** | 11.6 (114.1) |
| **B(iso) from Wilson / Å$^2$** | 42.6 |
| **Refinement Statistics** | |
| **Protein atoms excluding H** | 7869 |
| **Calcium atoms** | 16 |
| **R factor / %** | 0.201 |
| **$R_{free}$ / %** | 0.247 |
| **R.m.s. bond length / Å** | 0.013 |
| **R.m.s. bond angle / °** | 1.855 |

*Table 6.2 – Data reduction and refinement statistics for the PEF(S)-Hfq dataset.*

Only PEF(S) matched the electron density produced by MR, BUCCANEER was used to complete the interpretation of the electron density map resulting from the refinement of the partial solution with the PEF(S) domain.[211,212] The deduced sequence matched that of the small *E.coli* chaperone protein, Hfq (Host Factor for phage Q β-RNA replication). Several rounds of refinement of the model with REFMAC5 confirmed this solution.[180,181] Rounds of graphical adjustment of the model using COOT, and refinement with REFMAC5 led to the final model. Other data analysis was completed with programs in the CCP4 package.[181,182] The electron density map revealed that the unit cell contained 4 copies of the PEF(S) and 4 copies of the Hfq chaperone protein, non-covalently bound, and the structure was refined to $R_{work}$ and $R_{free}$ values of 0.201 and 0.247 respectively. The coordinates were deposited in the PDB with accession code 6QLB.

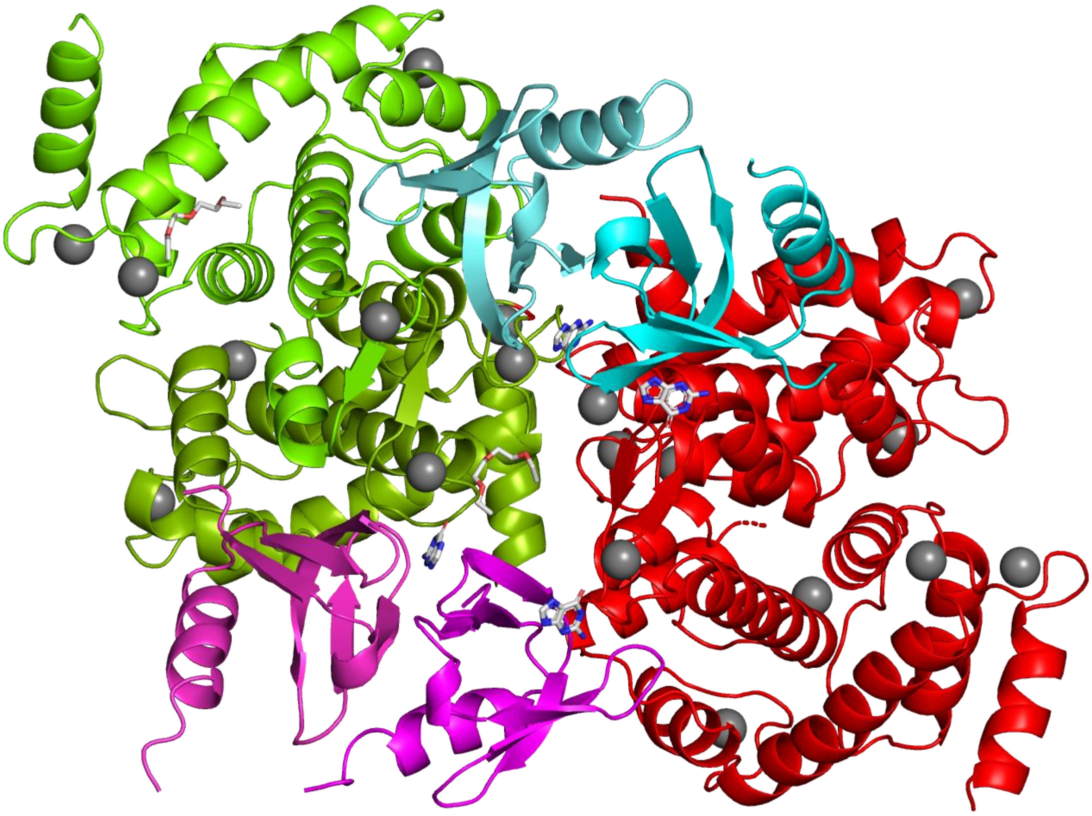


*Figure 6.1.3 - Asymmetric unit of the calpain PEF(S)-Hfq chaperone model in cartoon representation (PEF(S) – green, red shades, Hfq – cyan/magenta shades, $Ca^{2+}$ - grey spheres, guanine (GUN) and polyethylene glycol (PEG) ligands as sticks.*

Crystallisation trials using PACT-*premier* and JCSG++ screens produced several hits (Table 6.1). The impurity (Hfq) was likely overlooked in SDS-PAGE due to the small, faint band in the gels, which is poorly resolved on a 10% acrylamide gel (

Figure 6.1.1). The PEF(S) – Hfq interaction may have been strong enough to carry it through purification however it has been observed that the calpain-1 heterodimer can readily undergo subunit dissociation under mild conditions,[50] the stability of a PEF(S) homodimer in solution and tendency for the large subunit to precipitate likely contributed to the formation of the PEF(S)-Hfq bound structure in the crystallisation conditions. Purification of this protein in the future should involve a size exclusion chromatography step to ensure the protein of the correct size oligomeric state is isolated free from chaperone proteins.

### 6.1.5. Hfq-PEF(S) interactions

The Hfq-PEF(S) interactions show 1:1 stoichiometry, and are primarily hydrophilic between the 4th EF hand in the PEF(S) monomer and the Hfq nucleotide binding site. Hydrogen bonding from the PEF(S) backbone G153 oxygen and the Hfq Q52 amino group at a distance of 2.95 Å is favourable. The interface area between each PEF(S) dimer pair (chains A-B, C-D) is 2093 $Å^2$ and 2017 $Å^2$. The Hfq dimer pairs (chains E-F, G-H) share interface areas between 668 -717 $Å^2$.[213] The mixed PEF(S)-Hfq interface areas are smaller than each homodimer interface between 444 – 506$Å^2$

The strength of interaction between PEFS(S):PEF(L) ( $\Delta$G = -27.2 kJ mol-1 ) estimated using PISA was found to be higher than that between PEF(S) and Hfq ( $\Delta$G = -12.6 kJ mol-1) respectively.[186] The PEF(S)-Hfq oligomer may be an appropriate description of how PEF(S) is retained separately from the rest of the full length calpain-I during ER folding, until the counterpart PEF(L), is ready to displace the chaperone and latch on to its PEF(S) and complete the folding during expression in this E. coli strain.

Sufficient density for a small molecule was found in the nucleotide binding site of all Hfq monomers in a hydrophobic pocket at Y25, with leucines and isoleucines further increasing the hydrophobicity, the ligand was modelled as guanine, density for the ribose or phosphate groups of a nucleoside or nucleotide were not detectable at this resolution.

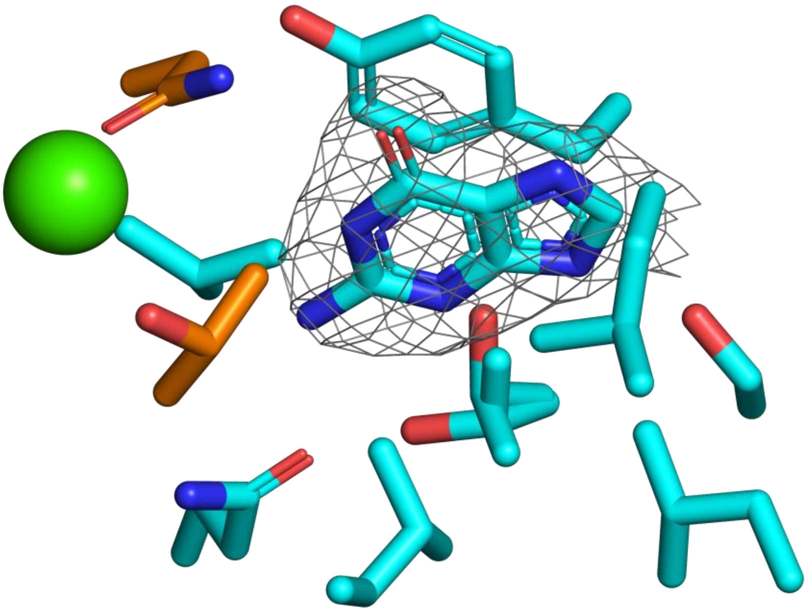


*Figure 6.1.4 - Guanine bound in hydrophobic pocket between Hfq (cyan) and PEF(S) (orange) chain residues selected within 5 Å of the guanine (GUN) ligand and $Ca^{2+}$ ion (green sphere), the map is contoured to 1.0 σ. A water molecule was modelled in to satisfy the density above the GUN ligand,*

While this structure was not the original goal for these crystallographic experiments, it does provide insight into the stress that overexpression of proteins can have on the host bacteria, and the subsequent genetic response to these processes. The primary focus of Hfq studies has been on RNA binding capabilities, having been shown to be involved in distinct metabolic pathways including sugar transport, membrane remodelling and quorum sensing.[214215] Hfq protein-protein interactions have been studied to a lesser extent, so the structure presented here suggests that Hfq could play a role in cellular systems via protein binding interactions.[216] Prokaryotes primarily express the Sm family member Hfq, while eukaryotic cells express over 20 multimeric Sm-like chaperone complexes.[217] The endoplasmic reticulum functions as a folding compartment for protein production as well as a highly dynamic calcium store capable of sensing cellular stress, triggering biochemical cascades that result in cell apoptosis.[218] Upon accumulation of misfolded proteins, calcium homeostasis can be disrupted. Calpains are known to be modulator proteases, playing an extensive proteolytic role in cell apoptosis after activation *via* calcium release, from the endoplasmic reticulum stress response.[218] It is possible that the complex observed in this crystal structure suggests a novel role of human Sm-like chaperone proteins, capable of binding to calpain calcium binding EF domains, highlighting a potential involvement for Sm-like chaperones in the stress response. The formation of human Sm-like chaperones complexed to calpain regulatory domains, as a result of *in vivo* endoplasmic reticulum induced cell stress, could have major implications in the mechanism and treatment of ischemia-based diseases.

### 6.1.6. Calpain-1 expression summary

Both CAPN1C115S and CAPNS1ΔGR were expressed and purified for crystallography trials, crystallisation of the protein isolated from anion exchange purification produced crystals containing only the small subunit and an *E.coli* chaperone, Hfq. Further purification of the construct with size exclusion chromatography whilst analysing the protein samples using an SDS gel with a higher acrylamide content >10 % or by tricine-SDS-PAGE should be used in this purification. This would facilitate the isolation of only calpain samples that do not contain the chaperone whilst ensuring the calpain complex is of the correct size in solution. It was not possible to obtain reliable solutions for the unsolved X-ray diffraction data with molecular replacement techniques using published calpain-1 domains or Hfq models from the PDB (PDB:1QXP, 2ARY, 4PHJ, 4PNO).[30,48,104,216] The large unit cells of these datasets are able to contain the entire calpain heterodimeric complex, but an exact solution is currently unknown. Preparation of fresh crystals incorporating anomalous diffraction techniques may be able to produce a solution for these datasets. Addition of heavy metal ions, selenomethionine replacement or addition of heavy ligands such as organomercury compounds can provide an anomalous signal able to solve the phasing of the data by multiple isomorphous replacement, however was not carried out in this project.

# Summary and Outlook

## 6.2. Summary

Novel diselenide compounds were synthesised and evaluated as redox stable allosteric inhibitors of calpain-1. These compounds displayed apparent $IC_{50}$ values for calpain-1 in the low micromolar range, similar in potency to the previous α-mercaptoacrylic acid inhibitors in reduced form. A live neutrophil spreading assay was used to show that the cell spreading process was inhibited by the presence of a diselenide at 10 μM concentration, showing these compounds are cell permeable and displayed similar efficacy in a cellular environment as the α-mercaptoacrylic acids.

A fluorescence assay was used to show the diselenides bind to the calpain calcium binding domains PEF(S) and PEF(L) with $K_d$ values in the low micromolar range. Co-crystal structures obtained of the diselenides bound to PEF(S) showed that the dihedral geometry of the diselenide bond facilitated analogous binding to the hydrophobic CAST binding groove as reported for the α-mercaptoacrylic acids. Multiple observed ligand geometries suggest that the binding to these sites is not optimal, therefore optimisation of the diselenide topology by SAR may produce more potent inhibitors.

The calpain-1 active site domain, CysPC, was expressed and purified as an active truncated form of calpain-1. The diselenide compounds were shown to inhibit this domain with greater potency than the full-length enzyme, confirming that they are not allosteric in their mode of action. The presence of several other non-covalent binding sites during the full-length calpain inhibition assays likely resulted in a lower effective free concentration of the inhibitor, therefore when inhibiting the isolated CysPC domain, a lower $IC_{50}$ was observed. CysPC was co-crystallised with the active site and soaked with diselenide **35**. Large volumes of difference density in the electron density map on the solvent exposed cysteine residues C49 and C108 was observed and this was proposed to be from the presence of a selenyl-sulfide adduct. The catalytic residue C115 is buried in the active site cleft and thus not believed to be exposed until substrate binding, in this case it was therefore was not able to react with the diselenide to form the selenyl-sulfide adduct.

Some of the diselenides were also tested for calpain-2 inhibition at a concentration of 100 μM, phenyl-based compounds **35** and **38** and anthracene **39** based inhibitor displayed some selectivity for calpain-1 over calpain-2, however the furan containing diselenide **40** resulted in the reduction in activity, at 100 μM concentration, indicating that this smaller moiety did not impart any isoform selectivity.

The computational design of novel calpain-1 inhibitors was explored using PEF(S) domain crystal structure as a 3D model for docking compounds into the binding site. This technique was validated using the TNS displacement assay to determine $K_d$ values for 4 out of the 10 compounds between 21.6 – 48.3 μM. Of these binders, three displayed an allosteric mode of action by inhibiting the full-length enzyme and not the isolated active site domain, with apparent $IC_{50}$ values ranging from 7.5 – 29.7 μM. The compounds were also tested for calpain-2 inhibition at a concentration of 100 μM, interestingly no significant inhibitory effect was observed by any of these compounds. The allosteric inhibitors were further investigated by TNS displacement assay on calpain-1 PEF(L) domain, **47** exhibited three-fold more potent binding for PEF(L) than to PEF(S). The observed isoform selectivity from these inhibitors may arise from binding to PEF(L), as this domain is unique between calpain isoforms.

The Hfq-PEF(S) structure obtained was not the goal of the calpain-1 crystallisation experiments however it does provide information into the possible role of Sm-like chaperones in protein folding, and highlights a potential link between the production of calpain at the ER and the role of Hfq has on the formation of this heterodimer.

## 6.3. Outlook

The diselenides tested were membrane permeable, inhibitors of calpain-1 that displayed moderate selectivity for calpain-1 over calpain-2. They were shown to bind non-covalently to the calcium binding domains PEF(S) and PEF(L), however when tested on the active site of calpain-1 a lower apparent $IC_{50}$ value was calculated, indicating that they are not allosteric in their mode of action. Structural data suggests that the formation of selenyl-sulfide adducts on the active site domain were the source of the inhibition. The complexity of the binding of these compounds to calpain at several sites in multiple ways strongly suggests they are unsuitable as *in vivo* inhibitors for the treatment of calpain related conditions. As these compounds contain selenium, which as a highly toxic but essential trace element, as a redox reactive group, it is unlikely to be considered a suitable drug candidate to inhibit calpain-1 *in vivo*.[219]

Mutation of the C49 and C108 residues to serine or alanine and repeating the inhibition assays using PD150606 and diselenides would provide confirmation if these residues are responsible for the inhibition of CysPC and full-length calpain-1. However as the diselenide compounds proved to be unselective and able to inhibit the active site alone, are therefore not good candidates for calpain inhibition.

Sulfonamide **47** and [1,2,4]triazoles **56** and **57** exhibited excellent selectivity for calpain-1 and -2 with little effect on the active site domain. These compounds displayed novel calpain inhibitor scaffolds, obtaining crystal structures of compounds bound in the PEF(S) and PEF(L) domains would enable functional SAR studies to optimise the potency and selectivity of the compounds. The crystal structure of heterodimeric PEF(L)/PEF(S) bound to **47**, **56** or **57** would likely provide mechanistic insight into the difference of binding observed for both proteins. This may also elucidate how the proposed binding to PEF(L) is able to impart isoform selective inhibition. Computational physiochemical calculations of the properties of [1,2,4]triazole (**56**) suggested that **56** may be permeable to the blood-brain barrier. Further experimental testing of **56** for CNS permeability, combined with SAR studies, may render this compound a potential candidate for allosteric calpain-1 inhibition in the CNS as a treatment for pathological conditions, such as AD. Physiochemical calculations of sulfonamide **47** predicated a lower CNS permeability, however this compound was more potent than **56** towards full length calpain-1. *In vivo* testing of this compound in cell spreading and migration assays should be carried out to determine if this compound displays potent *in vivo* calpain-1 inhibition.

The crystal structure of the full length calpain-1 heterodimer remains invaluable structural data, solving this would elucidate mechanistic consequences of activated calpain in physiological roles, as well as providing a 3D molecular model of a valuable therapeutic target. The reported protocol for expression of calpain-1 in *E. coli* produces the majority of the protein in the insoluble form, which cannot be recovered. The production of the heterodimeric protein using other expression systems such as yeast or mammalian cells may produce more of the target protein in soluble form which would be more suitable for structural studies. Purification of the protein using size exclusion chromatography and confirmation of the purity using tricine-SDS-PAGE gels to ensure small chaperones are not present would aid crystallisation experiments.

# Materials and Methods

# 6.4. Protein Expression and Purification

All chemicals were purchased from Sigma-Aldrich and Fisher Scientific, unless otherwise stated. IPTG, ampicillin and kanamycin were purchased from Melford. TCEP was purchased from Fluorochem. All chemicals were used in accordance to their material safety sheets, COSHH, and risk assessments

## 6.4.1. Luria-Bertani (LB) growth medium

Sodium chloride (10.0 g), tryptone (10.0 g), and yeast extract (5.0 g) were dissolved in 1 L of deionized water then autoclaved at 121 ℃ and 15 lb $in^{-2}$ for 20 minutes.

## 6.4.2. Terrific broth "enhanced"

Tryptone (12.0 g), yeast extract (24.0 g) and glycerol (4 mL) were dissolved in 900 mL of deionised water. The solution, 1 M $KH_2PO_4$ (136.08 g / L) and 1 M $K_2HPO_4$ (174.17 g / L) were sterilised in an autoclave at 121 ℃ and 15 lb $in^{-2}$ for 20 minutes. Upon cooling the solution was cooled and 1 M $KH_2PO_4$ (20 mL), 1 M $K_2HPO_4$ (80 mL), 1.0 mL sterile 1000x metals mix and sterile 2.0 M $MgSO_4$ (1.0 mL) were added.

1000x metals mix (50 mL) was prepared by dissolving $FeCl_3.6H_2O$ (1.35 g) in 50 mL 0.1 M HCl and filter sterilised. $MnCl_2.4H_2O$ (197.9 mg), $ZnSO_4.7H_2O$ (287.6 mg), $CoCl_2.6H_2O$ (47.5 mg) and $NiCl_2.6H_2O$ (47.5 mg) were dissolved in 50 mL $H_2O$ and filter sterilised. 50 mL of the $FeCl_3$ stock was mixed with the 50 mL dissolved metals to make 100 mL of 1000x metals mix to final concentrations 50 µM $FeCl_3$, 10 µM $MnCl_2$, 10 µM $ZnSO_4$, 2 µM $NiCl_2$ and 2 µM $CoCl_2$.

## 6.4.3. Antibiotics

Ampicillin (500 mg) was dissolved in 5 mL of deionized water to a concentration of 100 mg/mL, sterilised through a 0.2 µm filter and used at a working concentration of 0.1 mg/mL.

Kanamycin (100 mg) was dissolved in 2 mL of deionized water to a concentration of 50 mg/mL, sterilised through a 0.2 μm filter and used at a working concentration of 0.05 mg/mL.

### 6.4.4. LB agar

Tryptone (10.0 g), yeast extract (5.0 g) and sodium chloride (10.0 g) and agar (15.0 g) were dissolved in 1 l of deionised water for preparation of culture plates. The solution was sterilised in an autoclave at 121 °C, 15 lb.(sq. in)$^{-1}$ for 20 minutes then poured into sterile petri dishes and allowed to set at room temperature.

### 6.4.5. Transformation of plasmid DNA into bacteria

XL1-Blue ® competent cells (Stratagene, CA, USA) and the required DNA solution were thawed on ice. Once thawed, the DNA solution (1.0 μL) was pipetted into the cell suspension and mixed by flicking the Eppendorf under sterile conditions. The cell-DNA mixture was left on ice for 20 minutes before it was heat shocked in a water bath at 42 °C for 60 seconds and then allowed to cool on ice for 15 minutes. LB medium (1.0 mL) was then added to the solution and the cells were incubated for 1 hour at 37°C with shaking at 220 rpm in an Innova® 44 shaker (New Brunswick Scientific, Hertfordshire, UK). The cells were harvested *via* centrifugation (Eppendorf centrifuge 5424) at 16100 RCF for 1 minute. The supernatant solution was discarded and the pellet was resuspended in the residual fluid. The resuspended pellet was then plated onto selective LB agar plates. The plated cells were incubated at 37 °C overnight.

### 6.4.6. Growth and purification of plasmid DNA

One colony from the ampicillin selective (0.1 mg/mL) LB agar plates (*vide supra*) was extracted and placed into LB medium (20 mL) containing ampicillin (0.1 mg/ml).

The culture was incubated overnight at 37 °C while shaking at 220 rpm in an Innova® 44 shaker (New Brunswick Scientific, Hertfordshire, UK). The cells were then harvested by centrifugation (Eppendorf centrifuge 5424) for 15 minutes at 3220 RCF. The supernatant solution was discarded and the plasmids were isolated using the QIAprep spin miniprep kit. The pellet was first resuspended and lysed in P1-buffer (250 µL). P2-buffer (250 µL) was added before N3-buffer (350 µL) was added. The solution was mixed by inversion upon both additions. The cells were harvested through centrifugation (Eppendorf centrifuge 5424) for 10 minutes at 16100 RCF. The supernatant solution was transferred into a spin column where the plasmid then binds to the filter.

The centrifugation step was repeated for 1 minute at 16100 RCF and the flow through was discarded. PE-buffer (750 µL) was added into the spin column to wash the plasmids. The centrifugation step was repeated and the flow through was discarded. This step was repeated once more to ensure that all the residual ethanol had been removed. The miniprep column was then placed into a clean Eppendorf tube and sterile water (50 µL) was added. The solution was then centrifuged twice for 1 minute at 16100 RCF, and DNA concentration measured using NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, MA, USA).

### 6.4.7. DNA sequencing

Upon isolation of the plasmid DNA (*vide supra*), a UV absorbance spectrum of each solution was analyzed at 260 and 280 nm on the NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific, MA, USA) to give an estimate of the concentration of the DNA obtained. Deionized water was used as a blank and 1.5 µL of each solution was used for analysis. 15 µL of DNA between 50-150 ng/µL concentrations was sent for sequencing (Eurofins Genomics).

### 6.4.8. DNA miniprep buffers

*P1-Buffer (suspension buffer)*

Tris-HCl (158 mg) and EDTA (59 mg) were dissolved in 100 mL of deionized water to final concentrations of 50 mM Tris-HCl and 10 mM EDTA. The pH of the solution was adjusted to 8.0. RNAse was dissolved to a final concentration of 50 µg/ml. The resulting solution was stored at 4°C.

*P2-Buffer (lysis buffer)*

NaOH (4.0 g) and SDS (5.0 g) were dissolved separately in 250 mL of deionized water, to final concentrations of 0.4 M and 2% (w/v). These solutions were then mixed at a 1:1 ratio and stored at room temperature.

*N3-Buffer (neutralization and binding buffer)*

Guanidine hydrochloride (7.6 g) and potassium acetate (980 mg) were dissolved in 20 mL of deionized water, to concentrations of 4M and 0.5 M respectively. The pH of the solution was adjusted to 4.2 and the resulting solution was stored at room temperature.

*PE-Buffer (wash buffer)*

NaCl (0.117 g) and Tris-HCl (3.1 mg) were dissolved in 20 mL of deionized water. The pH was adjusted to 7.5 before ethanol (80 mL) was added to the solution. The resulting solution of 20 mM NaCl and 3 mM Tris-HCl was stored at room temperature.

### 6.4.9. Cloning strains

XL1-Blue® competent cells (Stratagene, CA, USA) and were used for the preparation of new plasmid constructs as they are highly proficient in the replication of DNA, creating high concentrations of the desired plasmid with an efficiency of up to ≥1 x 108 colony forming units per microgram.

### 6.4.10. PCR ampliciation of DNA

0.5 µL of template DNA, 0.5 µL of 50 mM forward and reverse primers and were added to 25 µL PrimeSTAR® HS DNA Polymerase 2x master mix (Takara Bio) and made up to 50 µL final volume. The mixture was placed in a Techne TC-512 thermocycler and cycle protocol as follows:

Step 1: 98 °C – 4 min
Step 2: 98 °C – 10 s
Step 3: 55 °C – 5 s
Step 4: 72 °C – 6 m 30 s
Repeat Step 2 – 4 for 30 cycles
Step 5: 72 °C – 10 min
Step 6: 4 °C

1.0 μL FastDigest Dpn1 (Thermo Fisher Scientific) was placed into the reaction mixture and incubated at 37 °C for 1 h. 1 μL of the reaction mixture was then transformed into XL1-Blue cells as described (Section 6.4.5).

### 6.4.11. Golden gate cloning

1.5 μL T4 ligase buffer, 0.15 μL BSA 100x stock, 5 μL acceptor plasmid, 5 μL gene fragment, 1.5 μL BsaI and 1 μL T4 ligase were added to a PCR tube and made up to 15 μL with deionised $H_2O$.

The mixture was placed in a Techne TC-512 thermocycler and cycle protocol as follows:

Step 1: 37 °C – 3 min
Step 2: 16 °C – 4 min
Repeat for 25 Cycles
Step 3: 50 °C – 5 min
Step 4: 80 °C – 5 min

5 μL of the reaction was then transformed into XL1-Blue cells as described (Section 6.4.5).

## 6.5. Protein Expression

### 6.5.1. Expression strains

BL21 (DE3) and BL21-CodonPlus-(DE3)-RP competent cells (Stratagene, CA, USA) were used for the expression of the human PEF(S) gene. The CodonPlus-(DE3)-RP strain of *E. coli* contains extra genes that encode for arginine and proline tRNA. The excess tRNA allows for the efficient expression of proteins that contain a large quantity of these amino acids.

### 6.5.2. Expression and purification of TEV protease

*BL21 (DE3) cells containing the TEV gene codon optimised for E. coli expression were obtained from Professor N. Richards (Cardiff University).*

The cells containing the TEV protease gene were grown at 37 °C in ampicillin selective LB media until $OD_{600}$ = 0.6-0.8 then induced with 1 mM IPTG (238 mg). The protein was expressed overnight at 20 °C and cells harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C. The cells were resuspended in 10 mL CysPC Buffer A (20 mM HEPES, 100 mM NaCl, 0.5 mM TCEP pH 7.6 and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 RCF in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 µm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 CV CysPC Buffer A containing 20 mM imidazole, and eluted with 10 CV buffer A containing 250 mM imidazole. The eluent was mixed with 20% v/v glycerol, aliquoted (1 mL) and stored at -80 °C.

### 6.5.3. Expression of the PEF(S) gene

BL21-CodonPlus-(DE3)-RP competent cells were transformed with PEF(S) DNA (1 µL) as previously described. The cells were incubated overnight at 37°C on the ampicillin selective (0.1 mg/mL) LB agar plates. A single colony was harvested for inoculation in 100 mL of ampicillin selective (0.1 mg/ml) LB media, which was incubated overnight at 37 °C while shaking at 150 rpm in an Innova® 44 shaker (New Brunswick Scientific, Hertfordshire, UK). 4 x 20 mL of the overnight starter culture was used to inoculate 4 x 1 L of ampicillin selective (0.1 mg/ml) LB media. The cultures were incubated at 37 °C with shaking as previously described until they reached an O.D. of 0.6 at 600 nm; at which point, they were induced with IPTG (126 mg, 0.5 mM) and incubated for a further 4 hours. Cells were harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, MA, USA) using a F21 rotor at 6080 RCF for 20 minutes. The supernatant solution was discarded and the pellets were stored at -20 °C.

### 6.5.4. Expression of the PEF(L) gene

PEF(L) was grown from a frozen glycerol stock (plasmid library entry 1121, Allemann group, Cardiff) of BL21 (DE3) CodonPlus RP cells containing the plasmid, in ampicillin selective LB media overnight at 37 °C. 10 mL of this culture was then grown in 1 L of TB 'Enhanced' media until $OD_{600}$ = 0.8-1.0, IPTG (238 mg) was added to reach a final concentration of 1 mM and the culture was grown for 20 hours at 20 °C, the cells were then harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C, then pelleted and frozen.

### 6.5.5. Expression and purification of CysPC

BL21-CodonPlus (DE3) RP cells containing the human calpain-1 CysPC gene were grown at 37 °C in kanamycin selective LB media until $OD_{600}$ = 0.6-0.8 then induced with 1 mM IPTG (238 mg). The protein was expressed overnight at 20 °C and cells harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C. The cells were resuspended in CysPC Buffer A (20 mM HEPES, 100 mM NaCl, 0.5 mM TCEP pH 7.6) and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 RCF in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 µm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 CV CysPC Buffer A containing 20 mM imidazole and eluted with 10 CV CysPC Buffer A containing 250 mM imidazole which was further dialysed in CysPC Buffer A overnight in a 10 kDa membrane containing 1 mL aliquot of purified TEV protease. The cleavage product was then passed back through a Ni-NTA column to remove the 6xHis-GB1 solubility tag and TEV protease, with the flow through containing CysPC.

### 6.5.6. CAPN1 C115S + CAPNS1ΔGR expression

C41 (DE3) pLysS cells containing the CAPN1 C115S and CAPNS1ΔGR (CAPNS1 with truncated GR domain[209]) genes were grown at 37 °C in kanamycin and ampicillin selective TB(Enhanced) media until $OD_{600}$ = 0.9 then induced with 1 mM IPTG. The protein was expressed overnight at 20 °C and cells harvested by centrifugation in a Sorvall RC6 Plus centrifuge (Thermo Fisher Scientific, Inc, MA, USA) using an SLA-3000 rotor at 6080 RCF for 20 minutes at 4 °C. The cells were re-suspended in 20 mM HEPES, 0.5 mM EGTA, 0.5 mM TCEP pH 7.6 (buffer A) and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 rcf in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 µm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 column volume (CV) buffer A and eluted with 10 CV buffer A supplemented with 250 mM imidazole. The eluted fraction was then applied to a MonoQ HR10/10 column (GE Healthcare) pre-equilibrated with buffer A. The protein was eluted with a linear gradient of 0-0.5 M NaCl in buffer A over 7 CV. The purity of the peak fractions was assessed by SDS-PAGE (10%). The protein was then concentrated to 10 mg $mL^{-1}$ in a 30 kDa MWCO Vivaspin concentrator.

## 6.6. Protein Purification

Proteins were purified using ATKA UPC-900 (Amersham Biosciences) or Bio-Rad NGC Medium-Pressure Liquid Chromatography System (California, USA) FPLC apparatus monitoring the UV absorbance at 280 nm and checking the peak fractions by SDS-PAGE.

### 6.6.1. PEF(S) buffer A (Low Salt)

Tris base (4.85 g) and EDTA (1.66 g) were dissolved in 2 L of deionized water to final concentrations of 20 mM Tris base, 2 mM EDTA, and 5 mM β-ME. The pH of the solution was adjusted to 8.0 using 6 M hydrochloric acid (HCl). The solution was filtered and thoroughly degassed *via* vacuum pump, 0.7 mL β-mercaptoethanol was then added to give a final concentration of 5 mM.

### 6.6.2. PEF(S) buffer B (High Salt)

Tris base (4.85 g), EDTA (1.66 g) and NaCl (116.9 g) were dissolved in 2 L of deionized water to yield total concentrations of 20 mM Tris base, 2 mM EDTA, 5 mM β-ME, and 1 M NaCl. The pH of the solution was adjusted to 8.0 using 6 M HCl. The solution was filtered and thoroughly degassed *via* vacuum pump; 0.7 mL β-mercaptoethanol was then added to give a final concentration of 5 mM.

### 6.6.3. PEF(S) buffer C (Size Exclusion)

Tris base (4.85 g), EDTA (1.66 g) and NaCl (17.5 g) were dissolved in 2 L of deionized water to yield total concentrations of 20 mM Tris base, 2 mM EDTA, and 0.15 M NaCl. The pH of the solution was adjusted to 8.0 using 6 M HCl. The solution was filtered and thoroughly degassed *via* vacuum pump, 0.7 mL β-mercaptoethanol was then added to give a final concentration of 5 mM.

### 6.6.4. PEF(L) buffer A

HEPES (2.38 g), NaCl (2.92 g), TCEP (42 mg), 5 % glycerol (v/v) were dissolved in 500 mL of deionized water to final concentrations of 20 mM HEPES, 100 mM NaCl, 0.3 mM TCEP and 5% (v/v) glycerol. The pH of the solution was adjusted to 7.6.

### 6.6.5. PEF(L) buffer B

HEPES (2.38 g), NaCl (2.92 g), TCEP (42 mg), imidazole (17.02 g) and 5 % glycerol (v/v) were dissolved in 500 mL of deionized water to final concentrations of 20 mM HEPES, 100 mM NaCl, 0.3 mM TCEP, 500 mM imidazole and 5% (v/v) glycerol. The pH of the solution was adjusted to 7.6.

### 6.6.6. PEF(L) buffer C

HEPES (2.38 g), TCEP (42 mg), 5 % glycerol (v/v) were dissolved in 500 mL of deionized water to final concentrations of 20 mM HEPES, 0.3 mM TCEP and 5%

(v/v) glycerol. The pH of the solution was adjusted to 7.6. The solution was filtered and thoroughly degassed *via* vacuum pump and stored at room temperature.

### 6.6.7. PEF(L) buffer D

HEPES (2.38 g), TCEP (42 mg), NaCl (14.61 g) 5 % glycerol (v/v) were dissolved in 500 mL of deionized water to final concentrations of 20 mM HEPES, 0.3 mM TCEP, 1 M NaCl and 5% (v/v) glycerol. The pH of the solution was adjusted to 7.6. The solution was filtered and thoroughly degassed *via* vacuum pump and stored at room temperature.

### 6.6.8. CysPC buffer A

HEPES (2.38 g), NaCl (2.92 g), TCEP (42 mg) were dissolved in 500 mL of deionized water to final concentrations of 20 mM HEPES, 100 mM NaCl, and 0.3 mM TCEP. The pH of the solution was adjusted to 7.6.

### 6.6.9. CysPC buffer B

HEPES (2.38 g), NaCl (2.92 g), TCEP (42 mg) and imidazole (17.02 g) were dissolved in 500 mL of deionized water to final concentrations of 20 mM HEPES, 100 mM NaCl, 0.3 mM TCEP and 500 mM imidazole. The pH of the solution was adjusted to 7.6.

### 6.6.10. SDS-PAGE buffers

*10% (w/v) Sodium dodecyl sulfate (SDS)*

Sodium dodecyl sulfate (20 g) was dissolved in 200 mL of water and stored at room temperature.

*10% (w/v) Ammonium persulfate (APS)*

Ammonium persulfate (50 mg) was dissolved in 0.5 mL of deionized water. This solution was stored at 4 °C.

*SDS Resolving Buffer*

Tris base (27.7 g) was dissolved in 150 mL of deionized water to give a total concentration of 1.5 M. The pH of the resulting solution was adjusted to 8.8 and stored at room temperature.

*SDS Stacking Buffer*

Tris base (6.0 g) was dissolved in 100 mL of deionized water to give a final concentration of 0.5 M. The pH of the solution was adjusted to 6.8 with 6 M HCl and the resulting solution was stored at room temperature.

*SDS Electrode Running Buffer (10x)*

Tris base (30.3 g), glycine (144.0 g), SDS (10.0 g) were dissolved in 1 L of deionized water giving total concentrations of 0.25 M Tris base, 1.92 M glycine and 1% (w/v) SDS. The solution was stored at room temperature and was diluted 1:9 with deionized water immediately before use.

*SDS Protein Sample Loading Dye*

Tris-HCl (1.25 mL, pH 6.8) (SDS Stacking Buffer), glycerol (3.0 mL), bromophenol blue (1.0 mL), SDS (2.0 mL) and β-ME (1.0 mL) were mixed. Deionized water was added to give a total volume of 10 mL giving total concentrations of 62.5 mM Tris-HCl, 30% glycerol, 0.6% (w/v) bromophenol blue, 10% SDS and 10% β-ME. The resulting solution was stored at room temperature.

*Gel Staining Dye*

Coomassie Brilliant Blue (CBB, 80 mg) was dissolved in ethanol (2 x 10 mL). Deionized water was added to make up a total volume of 500 mL, HCl (2.0 mL) was added and the solution was stirred for an hour. The resulting solution was stored at room temperature.

### 6.6.11. Large scale PEF(S) purification

Purification of the protein was carried out utilizing an ÄKTA fast protein liquid chromatography (FPLC) instrument (GE Healthcare Life Sciences, Buckinghamshire, UK). A UP- 920 pump was used for the buffers and the buffers were mixed to achieve the desired gradient with a M-925 gradient controller. The samples were injected into the column using an INV-907 injection valve and UV-analysis was carried out on a UPC- 900 detector. The nickel-NTA agarose resin used was purchased from Qiagen.

The frozen cell pellet was re-suspended in 20 mM HEPES, 100 mM NaCl, 0.5 mM TCEP and 5% glycerol pH 7.6 (buffer A) and lysed by sonication for 5 mins (pulsed 5 s on, 10 s off). The lysate was clarified by centrifugation at 4 °C for 40 minutes at 30310 RCF in a Sorvall RC6 Plus centrifuge. The supernatant was passed through a 0.2 μm syringe filter and applied to a Ni-NTA column. The bound protein was washed with 15 CV buffer A and eluted with 10 CV buffer A containing 250 mM imidazole. The fractions were analysed by SDS-PAGE.

### 6.6.12. PEF(S) anion exchange chromatography

Cell pellets were defrosted on ice before suspension in protein purification buffer A. The Q-Sepharose anion exchange column (column volume (CV) = 75 mL), which was attached to a FPLC instrument (ÄKTA FPLC), the column washed into protein purification buffer A until the UV absorbance at 280 nm approximately reached the baseline absorbance. Meanwhile, the thawed cells were then lysed by sonication in an ice bath (5 second sonication pulses and 10 seconds off) for 5 minutes.

The solution was centrifuged for 30 minutes at 30310 RCF and 4 °C in a Sorvall RC6 Plus centrifuge. The supernatant solution was transferred into a beaker and loaded into the sample loop of the FPLC instrument through a 0.2 μm syringe tip filter. A two-stage NaCl gradient was applied to the column, with 0.0–0.5 M NaCl (50%-100% PEF(S) protein purification buffer B, over 5 CV, followed by 0.5-1.0 M NaCl (50-100% buffer B) over 2 CV. Fractions (10 mL) were collected from 0-100% buffer B. The column was washed with protein buffer B for 3 CV and stored under high salt (1 M) conditions at 4 °C. The flow rate was set to 5 mL/min throughout the purification process and the absorbance was monitored at 280 nm. The fractions that showed high absorbance in the UV-spectrum were analyzed by SDS-PAGE.

The fractions containing the desired protein as indicated by SDS-PAGE analysis were combined in preparation for further purification *via* size exclusion chromatography (*vide infra*).

### 6.6.13. PEF(S) size exclusion chromatography

The eluent from the Q-Sepharose column that contained the desired protein was concentrated *via* Amicon® through a 30 kDa membrane to approximately 5 mL. The concentrated sample was loaded onto a HiLoad SuperDexTM 75 (GE Healthcare Life Sciences, Buckinghamshire, UK) column (320 mL CV) which had been equilibrated with PEF(S) protein purification buffer C . The column was eluted with 1.5 CV (480 mL) of buffer C at 2.5 mL/min and the absorbance was monitored at 280 nm. 10 mL fractions were collected and analyzed by SDS-PAGE to determine which fraction the desired protein was eluted in.

The fractions containing PEF(S) protein were pooled and buffer exchanged into crystallography buffer (Secton 6.7.4) using a VivaSpin concentrator 30 kDa MWCO to wash the protein 5x with the new buffer. This was carried out by concentrating 5 mL of protein solution of known concentration to 200 µL, then adding 4.8 mL crystallography buffer and repeating 5 times.

### 6.6.14. PEF(L) anion exchange chromatography

A Resource™Q column (7 mL, GE Healthcare Life Sciences, Buckinghamshire, UK) was equilibrated with PEF(L) Buffer C (Section 6.6.6) at a flowrate of 3 mL / min using a Bio-Rad (California, USA) NGC FLPC instrument. PEF(L) elution from Ni-NTA column was diluted to 50 mM NaCl using PEF(L) Buffer C) and loaded to the column. A gradient of 0 to 0.5 M NaCl was run over 10 CV with 2 mL fractions collected whilst monitoring the UV-absorbance at 280 nm. The peak fractions containing protein were analysed by SDS-PAGE.

### 6.6.15. Protein concentration determination

*Bradford Reagent*

50 mg Coumassie Brilliant Blue G-250 was dissolved in 50 mL of methanol and 100 mL 85% w/v phosphoric acid was added. The mixture was slowly poured into 850 mL $H_2O$ and stirred for 1 hour at room temperature. The solution was filtered before use.

Protein concentration was determined by Bradford assay. 1% BSA was diluted with deionised $H_2O$ to a final volume of 200 µL (Table 0.1). Samples of protein were typically taken as 1, 5 and 10 µL with 199 , 195 and 190 µL deionised water added respectively. 800 µL filtered Bradford reagent was added to each dilution to a final volume of 1 mL.

| **1% BSA (1%) / µL** | **0** | **2** | **4** | **6** | **8** | **10** | **12** | **14** | **16** | **18** | **20** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **$H_2O$ / µL** | 200 | 198 | 196 | 194 | 192 | 190 | 188 | 186 | 184 | 182 | 180 |

*Table 0.1 - Bradford assay BSA dilution series for in platereader.*

200 µL of each dilution was added to a Greiner CELLSTAR 96 well clear flat bottom plate, allowing triplicate measurement of each dilution concentration. Absorbance at 450 and 595 nm was recorded FLUOstar Omega plate reader at 25 °C. The data values were plotted as the triplicate average value of 595 nm / 450 nm and concentration (µg / mL) to obtain the concentration of the unknown sample.

# 6.7. Protein crystallography

## 6.7.1. Crystal tray setup

All consumables for protein crystallography were purchased from Molecular Dimensions Ltd (Suffolk, UK). The crystallography precipitant solutions were dispensed to the SwissCI 96 well 2 drop MRC plates with an Integra Viaflo 96 multichannel pipette (Berkshire, UK). A Douglas Instruments Ltd Oryx 4 (Hungerford, UK) dispensing robot was used to set up the trays, which was initially washed with 0.05% sodium azide solution. A 50 µL aliquot of protein sample was centrifuged at 3381 RCF for 30 seconds then placed into the robot along with the screening solutions in the respective tray (50 µL). The robot dispensed protein 0.2 µL aliquots of protein and precipitant solution into the sitting drop plates simultaneously which were then sealed with ClearVue sheets (Molecular Dimensions Ltd, Suffolk, UK). The trays were stored at 20 °C in a Molecular Dimensions Ltd (Suffolk, UK) floor standing vibration free incubator. Crystal monitoring was carried out using a microscope (Leica Microsystems, Buckinghamshire, UK) after 24 h, 72 h, 1 week and 2 weeks to analyse crystal growth.

### 6.7.2. Crystal soaking

For crystal soak experiments, precipitant from the large well was used to dilute inhibitors from 50 mM DMSO solution to 5 mM – 9 µL of precipitant, 1 µL inhibitor. The cover was removed from the tray and 0.5 µL of the dilution was then added to the drop containing the crystals. The wells were then re-sealed and the plates stored at 20 °C for 24 hours until harvesting.

### 6.7.3. Crystal harvesting

The cover for the well was removed where crystals had formed, 100 – 300 µm loops were used to harvest the crystals, which were then flash frozen in liquid nitrogen and placed into the transport puck, submerged in liquid nitrogen. Where cryo-protection was used, ethylene glycol was added to the large well of the crystallization tray to a final concentration of 20% v/v, 0.5 µL of this solution was then added to the crystal prior to harvesting. The crystals were stored in liquid nitrogen until the samples were transported to the synchrotron (Diamond Light Source, Oxford, UK).

Data was collected at Diamond Light Source (Oxfordshire, UK) at 100 K, using a wavelength of 0.976 Å with a Pialtus pixelated detector. The diffraction images were processed with the xia-2 data reduction system, the data were scaled, reduced and analysed using the Scala program from the CCP4i package (Collaborative Computational Project number 4).[178,182]

The structures solved in this project were solved by molecular replacement, using PHASER and MOLREP.[179,210] PDB entry 4PHJ was used as a starting model for PEF(S) strutures and PDB entry 4ARY for CysPC. PDB entry 1DF0 was used as a starting model for the PEF(S)-Hfq structure, however the solution would not reproduce the anticipated large subunit domains of calpain-1. Only domain VI, PEF(S), matched the electron density produced by MR. BUCCANEER was used to complete the interpretation of the electron density map resulting from the refinement of the partial solution with domain VI, (PEF(S)).[211,212] The sequence was for *E.coli* chaperone, Hfq, which was confirmed in the successful refinement of the model with REFMAC5.[180,181] The model was adjusted graphically using COOT, and rounds of refinement with REFMAC5 led to the final model. Other data analysis such as QtPISA was completed with programs in the CCP4 package.[181,182,186]

### 6.7.4. PEF(S) crystallisation precipitant 1

Sodium cacodylate (0.54 g), PEG6000 (6.25 g) and $CaCl_2$ (0.11 g) were dissolved in 50 ml of deionised water, resulting in final concentrations of 50 mM sodium cacodylate, 12.5% (w/v) PEG6000 and 20 mM $CaCl_2$. The pH of the solution was adjusted to 7.4 with 5 M NaOH. The solution was filtered with a 0.22 μm syringe tip filter and stored at 20 °C.

### 6.7.5. Protein preparation

*PEF(S) Protein Crystallisation Buffer*

Sodium cacodylate (1.07 g), EDTA (0.65 g), β-ME (350 μL) and $CaCl_2$ (0.11 g) were dissolved in 1 l of deionised water, resulting in the final concentrations of 5 mM sodium cacodylate, 2 mM EDTA, 5 mM β-ME and 1 mM $CaCl_2$ . The pH of the solution was adjusted to 7.0 with 5 M NaOH. The solution was then stored at 20 °C. PEF(S) was purified and concentrated using a protein spin column with a molecular weight cut-off of 30,000. The protein was then washed (4 times) with the PEF(S) protein crystallization buffer to ensure that the buffer was fully exchanged. The protein was concentrated to ~10 mg/ml.

### 6.7.6. Mass spectrometry

Liquid chromatography mass spectrometry (LC-MS) was performed on a Waters Synapt G2-Si quadrupole time of flight mass spectrometer coupled to a Waters Acquity H-Class UPLC system. The column was an Acquity UPLC protein BEH C4 (300 Å 1.7 µm x 2.1 mm x 100 mm) operated in reverse phase and held at 60 °C. The gradient employed was 95% A to 35% A over 50 minutes, where A is $H_2O$ with 0.1% $HCO_2H$ and B is acetonitrile (ACN) with 0.1% $HCO_2H$. Data was collected in positive electrospray ionization mode and analyzed using the Waters MassLynx software version 4.1. Deconvolution of protein charged states was obtained using the maximum entropy 1 processing software (MaxEntropy1).

### 6.7.7. Circular dichroism spectroscopy

All circular dichroism spectroscopy was carried out using the Chirascan™ apparatus (Applied Photophysics Limited, UK)

A 10 mM KiPO4 buffer at pH 7.4 was filter sterilised through a 0.2 µm syringe then degassed *in vacuo*. A blank measurement was taken of the buffer prior to recording the experiment.

### 6.7.8. Calculation of mean residue ellipticity (MRE)

MRE was calculated using the signal obtained from the CD spectrometer as Θ.

$$\Theta\mathrm{MRE} = \theta / (10 \times n \times c \times l)$$

*Equation 3 - The equation used for calculating CD mean residue ellipticity (MRE); where $\theta$ = CD signal (millidegrees), n = number of back bone peptide bonds, c = molar concentration and l = the path length of the cuvette used (cm).*

The pathlength of the cuvette was 0.1 cm and had a total volume of 350 µL, typically CD measurements were taken on 200 µL of sample. The CD spectra were collected at 20 °C and recorded as an average of 5 spectra.

# 6.8. Bioassay Methods

## 6.8.1. FRET substrate (FAM) assay

100 mM HEPES (476 mg), 5 mM EDTA (38 mg) was dissolved in 19 mL deionised water and pH adjusted to 6.8, then 1 mL of 1% BSA was addded, sterilised by passing through 0.2 µm syringe filter and stored as 1 mL aliquots. 90 mM reduced glutathione (307 mg) and 10 mM oxidised glutathione (123 mg) were dissolved in 20 mL deionised water and stored as 1 mL aliquots. These 10x buffer components were defrosted on ice before and diluted accordingly to the assay requirements in deionised water. The layout of the 96 well plate for each inhibitor was typically as shown.

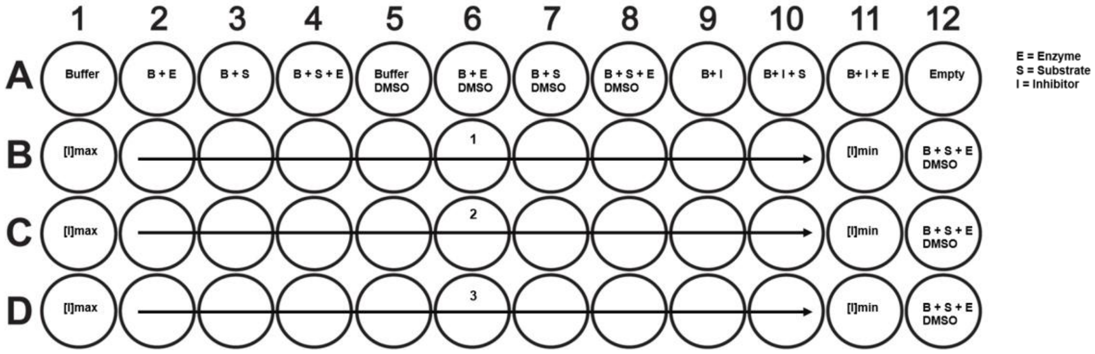


*Figure 6.8.1 – Typical layout of 96 well plate used for calpain inhibition assay with each set of inhibitors in triplicate.*

50 µL FAM assay buffer (10 mM HEPES, 1 mM EDTA, 0.1 % BSA, 10 mM glutathione 9:1 reduced:oxidised, pH 6.8 was placed into a Greiner CELLSTAR 96 well black flat bottom plate. 50 mM inhibitor in DMSO (or DMSO to the same volume as the control) was diluted over a range by serial dilution. Calpain-1 (Merck) was stored as a 2.5 µM stock in assay buffer, requiring 1 µL enzyme per well used during the assay, the FAM substrate was stored as a 1000x DMSO stock. 40 µL FAM assay buffer containing human calpain-1 and FAM substrate to the final concentration was then added to each well. The excitation wavelength was 490 nm and the emission wavelength was 520 nm as detected by the platereader and the assay was carried out at 25 °C. The fluorescence of the assay conditions in each well was monitored prior to calpain-1 activation. 5 mM $CaCl_2$ was added to initiate the reaction. The apparent $IC_{50}$ values were obtained by fitting the gradients with non-linear regression

with the SigmaPlot (Systat Software, San Jose, CA) and reported results are the mean +/- standard deviations of three independent experiments.

### 6.8.2. TNS displacement assay

*TNS Fluorescence Buffer*

Tris base (242 mg), $CaCl_2$ (12 mg) and EDTA (38 mg) were dissolved in 100 mL deionised water to final concentrations of 20 mM Tris, 1.1 mM $CaCl_2$ and 1 mM EDTA then adjusted to pH 7.4, filtered and degassed *in vacuo*.

Prior to the assay the protein was then washed (5 times) with the TNS fluorescence buffer in a Vivaspin concentrator (10 kDa MWCO) to ensure that the buffer was fully exchanged.

10 µM protein (PEF(S) or PEF(L)) in TNS fluorescence buffer was incubated with TNS (46.7 µM final from 1mM stock in 40% ethanol) in a Greiner CELLSTAR 96 well black flat bottom plate for 5 minutes at 25 °C. Compounds were stored as 50 mM stock solutions in DMSO, and then diluted from 500 µM to 500 nM by serial dilution over 10 wells with an Integra Viaflow 96 multichannel pipette in triplicate.

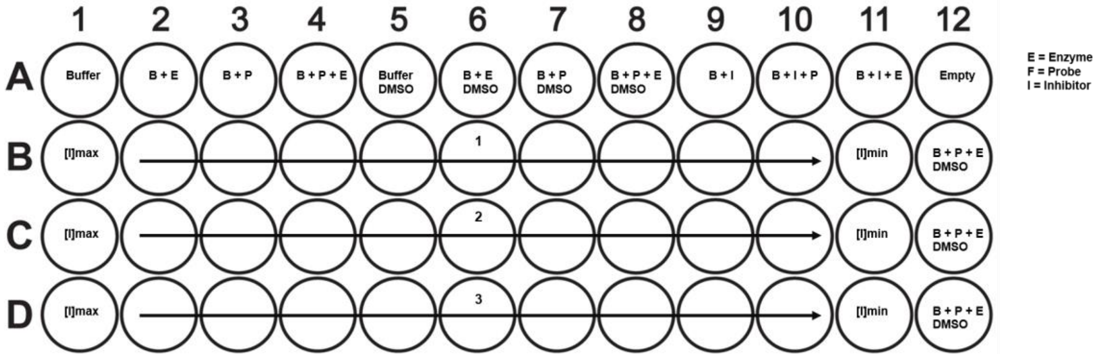


*Figure 6.8.2 - Typical layout of 96 well plate used for calpain domain binding assay with each set of inhibitors in triplicate.*

The plates were then incubated in a FLUOstar Omega plate reader at 25 °C for 5 minutes then were analysed using an excitation wavelength of 355 nm and an emission wavelength of 450 nm with 10 flashes per well with orbital averaging. The baseline (just TNS) was subtracted from the fluorescence (B). B was then subtracted from the $B_{max}$ (with no inhibitor) to invert the data. The ligand-binding module on

Sigmaplot (Systat Software, San Jose, CA) was used to calculate the $K_d$ values for the inhibitors.

### 6.8.3. Neutrophil spreading assay

Human neutrophils were isolated from the blood of healthy volunteers as described previously[220] as approved by SMREC ethics board (Cardiff University School of Medicine Research Ethics Committee, Cardiff, UK) with the informed consent of participants. The neutrophils were then suspended in Krebs medium (120 mM NaCl, 4.9 mM KCl 1.2 mM KH2PO4, 1.2 mM MgSO4, 1.3 mM CaCl2, 25 mM HEPES and 0.1% BSA, adjusted to pH 7.4 with NaOH). The neutrophils were loaded with the fluorescent $Ca^{2+}$ probe Fluo4 acetoxylmethyl ester (Fluo4-AM, 1 mM) (Molecular Probes). Extreme care was taken during the isolation and loading procedures to avoid mechanical activation of the cells, which can cause them to adhere rapidly and spontaneously to uncoated glass. Under the isolation conditions used here, neutrophils adhere to glass, but remain in the apparently ‘spherical’ (i.e. non-spread) morphology. $Ca^{2+}$.

The cells were monitored using a resonant scanning head of the Leica RS confocal microscope with a frequency of 40 MHz. Neutrophils were allowed to adhere to the glass coverslip mounted onto a thermostatically controlled stage (37.0˚C) During image acquisition, spreading activation was achieved by addition of fMLP. The cell morphology change during neutrophil spreading was dramatic and quantification of diameter was measured in imageJ software.

# 6.9. General Synthetic Methods

*All chemicals were purchased from Sigma-Aldrich unless otherwise stated. Reactions were stirred at room temperature in air unless otherwise stated. All glassware was clean and dry before use. Dry ethanol was prepared via freeze-pump-thaw degassing method.*

$^1$H and $^{13}$C NMR spectra were measured on a Bruker Avance 500 NMR spectrometer, a Bruker Avance DPX400 NMR spectrometer or a Bruker Fourier 300 and are reported as chemical shifts in parts per million downfield from tetramethylsilane, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constant (to the nearest 0.5 Hz) and assignment, respectively. $^{19}$F NMR spectra were measured on a Bruker Avance 500 NMR spectrometer and a Bruker Avance DPX400 NMR spectrometer and are reported in chemical shift downfield from $CFCl_3$, followed by multiplicity and coupling constant (to the nearest 0.5 Hz) if appropriate. IR spectra were recorded on a Shimadzu IRAffinity-1S Fourier Transform Infra-Red spectrometer and peak frequencies (in wavenumbers) are listed in descending numerical order. Mass Spectra were recorded by the School of Chemistry Analytical Service in Cardiff University. Purity of some compounds was assessed by LC/MS.

## 6.9.1. GC-MS analysis

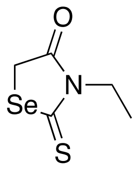
GCMS was performed on a Perkin Elmer Clarus 680 GC fitted with a Perkin Elmer Elite-1 column (30 m x 0.25 mm internal diameter) and a Perkin Elmer Clarus SQ 8 C mass spectrometer. The program uses an injection port temperature of 100 °C; split ratio 19:1; initial temperature 80 °C hold 2 min, ramp of 10 °C/min to 200 °C, then 20°C/min to 280 °C, hold 3 min.

## 6.9.2. UV-Vis analysis

1 µL of a stock solution of , (2Z,2'Z)-2,2'-diselanediylbis(3-(furan-2-yl)acrylic acid) (**40**) 50 mM in DMSO, was added to 1 ml of a buffered solution containing

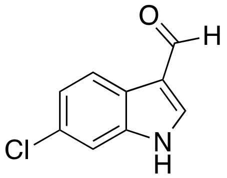
HEPES (10 mM), (pH 7.4). A UV-Vis spectrum was measured over the range 450 nm to 200 nm at 20 °C. A Jasco V-660 Spectrophotometer (Jasco-UK, Essex, UK) was used with the temperature regulated with a Peltier element.

### 6.9.3. 3-ethyl-thioxo-1,3-selenazolidin-4-one (**12**)



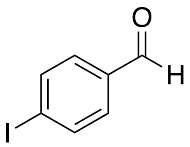
A flame dried, round bottom flask in an argon atmosphere was added degassed anhydrous ethanol (200 mL) at 5˚C. To this was added powdered selenium (6.00 g, 75.9 mmol) followed by sodium borohydride (3.00 g, 79.2 mmol). After 15 min the initial vigorous gas evolution had subsided and the ice bath was removed. The dark red solution was left to stand at RT for 0.5 h with stirring to give a colourless ethanolic solution of sodium hydrogen selenide. A solution of chloroacetic acid (7.20 g, 75.9 mmol) in degassed anhydrous ethanol (7 mL) was added dropwise to the freshly prepared sodium hydroselenide solution in ethanol and was heated for 15 min at 35 °C to give a grey-white suspension. Ethylisothiocyanate (8.1 mL, 91.2 mmol) was added dropwise and the yellow solution was heated for 1 h at 50 °C. The resulting green solution was heated to 80 °C for 4 h and further ethylisothiocynate (1 mL) was added after 1 h and 3 h. The reaction mixture was cooled, and cold water (30 mL) was added to quench the reaction. The reaction product was extracted with chloroform (3 x 50 mL). The chloroform layer was washed with water and dried over anhydrous magnesium sulfate. The solution was filtered and concentrated under reduced pressure to give an orange oily liquid. The oil was further purified in vacuo in a short path Kügelrohr distillation apparatus (b.p. 120 °C – 125 °C / 12 mm » 16 mbar) and a clear distillate containing ethanol was removed in the first two flasks to give the product, mostly present in the original flask, as an orange coloured oil (5.077 g, 24.39 mmol, 32%).$\delta_H$ (400 MHz, $CDCl_3$): 4.01 (2 H, q, *J* = 7.1 $CH_2$), 2.04 (2 H, s, ArCH), 1.14 (3 H, t, *J* = 7.3, $CH_3$) $\delta_C$(100 MHz, $CDCl_3$): 199.07 (CS), 176.26 (CO), 41.30 ($CH_2$), 28.88 ($CH_2$), 11.96 ($CH_3$). Mass spectrum: HRMS ($EI^+$) found 204.9436, $C_5H_7NOS^{76}Se$ requires 204.9440. υ $cm^{-1}$: 2974.2, 2933.7, 1705.1, 1620.2, 1531.5, 1367.5, 1342.5, 1309.67, 1234.4, 1116.8, 825.5, 769.6.

### 6.9.4. 6-chloro-1H-indole-3-carboxylic acid (**13**)



6-chloroindole (172 mg, 1.13 mmol), and *N,N*-dimethylformamide (3 mL), were stirred at 0 °C. Phosphorus oxychloride (0.12 mL, 1.25 mmol) was added dropwise and the solution was allowed to stir at room temperature for 15 mins before heating to 40 °C for 1 hour. The reaction was then cooled to room temperature and ice cold 2 M NaOH (2 mL) was added. The mixture was then heated under reflux for 1 hour then cooled on ice and the precipitate collected by filtration, the solid was washed with water (3 x 5 mL) and air dried under reduced pressure to yield the desired product as an orange solid (182 mg, 1.01 mmol, 89%). $\delta_H$ (400 MHz, DMSO-$d_6$): 12.24 (1 H, s, NH), 9.96 (1 H, s, CHO), 8.35 (1 H, s, ArCH), 8.09 (1 H, d, *J*= 8.4), 7.59 (1 H, d, *J* = 1.9, ArCH), 7.27 (1 H, dd, *J* = 8.4, 1.9, ArCH). $\delta_C$ (100 MHz, $CDCl_3$): 185.5 (CHO), 139.7 (ArCH), 137.9 (ArC), 128.3 (ArC), 123.3 (ArCH), 122.9 (ArCH), 122.5 (ArC),118.4 (ArCH), 112.6 (ArC). Mass spectrum: HRMS ($EI^+$) found 180.0213, $C_9H_7NOCl$ requires 180.0216. υ $cm^{-1}$: 3101.5, 3032.1, 2939.5, 2885.5, 2825.7, 1637.6, 1618.3, 1525.7, 1498.7, 1450.5, 1421.5, 1388.8, 1340.5, 1274.9, 127.9, 1205.5, 1161.2, 1128.4, 1111.0, 889.2, 848.7, 777.3, 738.7, 657.7, 609.5, 572.9.

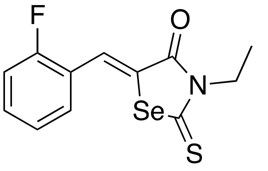
### 6.9.5. 4-iodobenzaldehyde (**15**)



1,4-diiodobenzene (0.7277 g, 2.2 mmol) was dissolved in 10 mL dry THF under an argon atomosphere then cooled to -78 °C. 0.84 mL n-butyllithium was added dropwise and stirred for 30 mins. DMF (0.2 mL) was then added dropwise and the yellow solution was stirred for a further 30 mins. The reaction was then quenched with water and the solution was extracted with diethyl ether, the combined organic fractions were then washed with water, brine, dried with $MgSO_4$ then filtered and the solvent removed *in vacuo*. The yellow oil was purified by flash chromatography hexane/ethyl acetate (7:3). The solvent was removed *in vacuo* and the product was isolated as a yellow solid (0.2693 g, 1.2 mmol, 53%). $\delta_H$ (400 MHz, $CDCl_3$): 7.85 (1 H, s, CHO), 7.85 (2H, dt, *J* = 8.1, *J* = 1.8, ArCH), 7.50 (2 H, dt, *J* = 8.1, *J* = 1.8, ArCH), $\delta_C$ (100 MHz, $CDCl_3$): 191.5 (CHO), 139.3 (ArCH), 138.4 (ArC), 130.8 (ArCH), 93.3 (CI), Mass spectrum: HRMS (EI$^+$) found 231.9384, $C_7H_5O^{127}I$ requires 231.9385. $\upsilon$ cm$^{-1}$: 2825.7 2796.8, 2735.1, 1685.8, 1579.7, 1402.3, 1379.1, 1203.6, 1161.2, 1049.3, 1022.9, 827.5, 800.5, 675.1, 623.0.

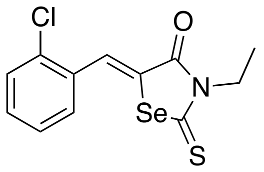
# 6.10. General Method for Condensation Reactions

## 6.10.1. (Z)-3-ethyl-5-(2-fluorobenzylidene)-2-thioxo-1,3-selenazolidin-4-one (**16**)



2-Fluorobenzaldehyde (25 mg, 0.20 mmol), *N*-ethylselenorhodanine (50 mg, 0.24 mmol), ammonium acetate (46 mg, 0.59 mmol) were dissolved in acetic acid (2 mL) and refluxed for 2 hours. The reaction mixture was then cooled on and ice deionised water was added (10 mL). The resultant yellow/brown crystals were filtered, washed with cold deionised water (2 x 10 mL) and dried *in vacuo* to yield the product as a yellow/brown solid (21.6 mg, 0.069 mmol, 34%). $\delta_H$ (500 MHz, $CDCl_3$): 8.20 (1 H, s, C*H*=CSe), 7.51-7.75 (1 H, m, ArCH), 7.50-7.33 (3 H, m, ArCH), 4.12 (2 H, q, *J* = 7.1, $CH_2$) 1.19 (3 H, t, *J* = 7.1, $CH_3$). $\delta_C$ (100 MHz $CDCl_3$) 193.2 (SeCS), 169.4 (CO), 161.0 (d, *J* = 254.8, CF), 132.4 (d, *J* = 8.7, Ar*C*=C) 129.4 (d, *J* = 5.6, ArCH), 128.6 (CSe), 125.7 (ArC), 124.9 (d, *J* = 3.5, ArCH) 122.6 (d, *J* = 11.8, ArCH), 116.4 (d, *J* = 22.1, ArCH), 41.1 ($NCH_2$), 12.2 ($CH_3$) $\delta_F$ (500 MHz $CDCl_3$): -113.01. Mass spectrum: HRMS (ASAP$^+$) found 315.9805, $C_{12}H_{10}FNOSSe$ requires 315.9632. $\upsilon$ cm$^{-1}$: 2972.3, 2927.9, 1697.4, 1597.1, 1475.5, 1450.5, 1384.9, 1375.3, 1348.2, 1307.7, 1288.5, 1246.0, 1224.8, 1124.5, 1033.9, 997.2, 979.8, 837.1, 744.5, 707.9, 644.2, 584.4, 551.6, 536.2, 501.5.

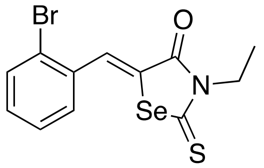
### 6.10.2. (Z)-3-ethyl-5-(2-chlorobenzylidene)-2-thioxo-1,3-selenazolidin-4-one (**17**)



Yellow/brown solid, 20%

δ$_H$ (500 MHz, CDCl$_3$): 8.39 (1 H, s, CH=CSe), (1 H, d, *J* = 8.1, ArCH), 7.49-7.37 (1 H, m, ArCH), 7.31-7.19 (3 H, m, ArCH), 4.16 (2 H, q, *J* = 7.1, CH$_2$) 1.22 (3 H, t, *J* = 7.1, CH$_3$). δ$_C$ (100 MHz CDCl$_3$): 193.3 (CS), 169.1 (CO), 136.1 (CCl), 133.4 (C*H*=CSe), 131.3 (ArCH), 130.5 (ArCH), 128.2 (ArCH), 127.4 (ArCH), 126.8 (ArC), 119.6 (CSe), 41.0 (CH$_2$), 12.1 (CH$_3$). Mass spectrum: HRMS (ASAP$^+$) found 327.9454, $C_{12}H_{10}ClNOS^{76}Se$ requires 327.9442. υ cm$^{-1}$: 3039.8, 2976.2, 2947.2, 1697.4, 1581.6, 1433.1, 1371.4, 1340.5, 1311.6, 1240.2, 1209.4, 1120.6, 1095.6, 1078.2, 1037.7, 754.2, 700.2, 677.0, 644.22, 584.4.

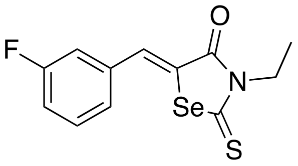
### 6.10.3. (Z)-5-(2-bromobenzylidene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**18**)



Yellow/brown solid, 64%

δ$_H$ (400 MHz, CDCl$_3$) 8.32 (1 H, s, CH=CSe), 7.61 (1 H, d, *J* = 8.1, ArCH), 7.31 (1 H, d, *J* = 8.1, ArCH), 7.26-7.14 (2 H, m, ArCH), 4.16 (2 H, q, *J* = 7.1, CH$_2$) 1.22 (3 H, t, *J* = 7.1, CH$_3$) δ$_C$ (100 MHz CDCl$_3$) 192.8 (SeCS), 169.4 (CO), 139.2 (Ar*C*=C), 138.8 (ArC) 136.3 (ArC), 135.1 (ArCH), 130.8 (ArCH) 128.6 (ArCH) 125.1 (CSe), 95.0 (CBr), 41.1 (NCH$_2$), 12.18 (CH$_3$) Mass spectrum: HRMS (EI$^+$) found 370.8866, $C_{12}H_{10}BrNOSSe$ requires 370.8859. υ cm$^{-1}$: 3049.5, 2976.2, 2929.9, 1685.8, 1653.0, 1591.3, 1558.5, 1541.6, 1438.4, 1381.0, 1367.5, 1336.7, 1305.8, 1274.9, 1226.7, 1203.5, 1116.8, 1064.7, 989.5, 925.8, 888.7, 829.4, 794.7, 765.7, 729.1, 665.4, 621.1, 574.8, 524.0.

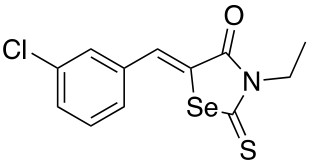
### 6.10.4. (Z)-3-ethyl-5-(3-fluorobenzylidene)-2-thioxo-1,3-selenazolidin-4-one (**19**)



Orange/brown solid, 36%

$\delta_H$ (400 MHz, DMSO-$d_6$): 8.22 (1 H, s, CH=CSe), 7.68-7.59 (1 H, m, ArCH), 7.52-7.39 (1 H, m, ArCH), 7.37-7.32 (2 H, m, ArCH), 4.10 (2 H, q, *J* = 7.1, $CH_2$) 1.18 (3 H, t, *J* = 7.1, $CH_3$) $\delta_C$ (100 MHz DMSO-$d_6$) 192.9 (SeCS), 169.5 (CO), 163.0 (d, *J* = 233.7, CF), 136.3 (d, *J* = 8.15, ArC), 135.5 (Ar*C*H=C), 131.0 (d, *J* = 8.22, ArCH),), 125.6 (ArCH), 125.2 (CSe), 117.5 (d, *J* = 21.9, ArCH), 116.4 (d, *J* = 22.3, ArCH), 41.0 ($NCH_2$), 12.1 ($CH_3$). $\delta_F$ (400 MHz, DMSO-$d_6$): -111.00. Mass spectrum: HRMS ($EI^+$) found 310.9661, $C_{12}H_{10}FNOSSe$ requires 310.9659. $\upsilon$ cm$^{-1}$: 2980.0, 2937. 6, 1699.3, 1681.9, 1573.9, 1431.2, 1371.4, 1346.3, 1315.5, 1286.5, 1232.5, 1157.3, 1122.6, 948.9, 927.8, 885.33, 773.5, 760.0, 736.8, 671.2, 638.4, 580.6, 518.9.

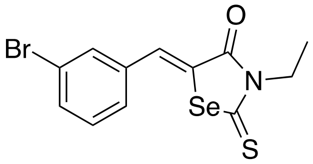
### 6.10.5. (Z)-5-(3-chlorobenzylidene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**20**)



Orange solid, 85%

$\delta_H$ (400 MHz, DMSO-d6): 8.22 (1 H, s, CH=CSe), 7.71 (1 H, s, ArCH), 7.64-7.55 (2 H, m, ArCH), 7.49-7.38 (1 H, m, ArCH), 4.11 (2 H, q, *J* = 7.0, CH2) 1.18 (3 H, t, *J* = 7.0 CH2). $\delta_C$ (100 MHz DMSO-d6) 192.8 (SeCS), 169.4 (CO), 135.9, (CCl), 135.4 (ArC), 135.2 (C*H*=CSe), 130.5 (ArCH), 130.4 (ArCH) 129.8 (ArCH) 127.7 (ArCH), 125.2 (CSe), 41.1 (NCH2), 12.2 (CH3). Mass spectrum: HRMS (ASAP+) found 331.9424, $C_{12}H_{10}ClNOSSe$ requires 330.9337. $\upsilon$ cm$^{-1}$: 2974.2, 2935.7, 1683.9, 1591.3, 1577.8, 1560.4, 1541.1, 1506.3, 1458.2, 1369.5, 1294.2, 1230.6, 1157.3, 1126.4, 1095.6, 825.5, 767.7, 754.2, 609.5, 582.5, 518.9.

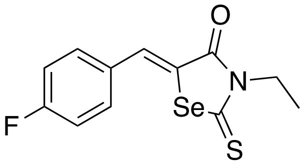
### 6.10.6. (Z)-5-(3-bromobenzylidene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**21**)



Brown solid, 60 %

$\delta_H$ (400 MHz, DMSO-$d_6$): 8.21 (1 H, s, CH=CSe), 7.85 (1 H, t, *J* = 1.8, ArCH), 7.71 (1 H, dt, *J* = 7.5, *J* = 1.8, ArCH), 7.52-7.46 (2 H, m, ArCH), 4.10 (2 H, q, *J* = 7.1, $CH_2$), 1.18 (3 H, t, *J* = 7.1, $CH_2$). $\delta_C$ (100 MHz DMSO-$d_6$) 194.2 (SeCS), 169.0 (CO), 135.3 (Ar*C*=C), 132.7 (ArC) 132.5 (ArC), 132.2 (ArC), 130.9 (ArC=*C*) 129.1 (ArC) 128.9 (ArC), 127.3 (CBr), 41.2 ($NCH_2$), 12.3 ($CH_3$). Mass spectrum: HRMS (ASAP+) found 371.8925, $C_{12}H_{10}BrNOSSe$ requires 371.8937. υ $cm^{-1}$: 2980.0, 2933.7, 1685.8, 1653.0, 1647.2, 1593.2, 1541.1, 1382.9, 1371.4, 1338.5, 1305.8, 1276.9, 1234.4, 1203.6, 1120.6, 1072.4, 923.9, 875.7, 833.3, 798.5, 771.5, 732.9, 717.5, 667.4, 623.0, 574.8, 553.6, 545.9.

### 6.10.7. (Z)-3-ethyl-5-(4-fluorobenzylidene)-2-thioxo-1,3-selenazolidin-4-one (**22**)

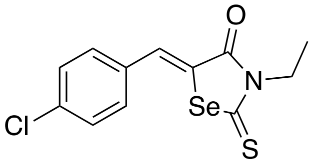


Orange/brown solid, 65%

$\delta_H$ (400 MHz, $CDCl_3$): 8.14 (1 H, s, CH=CSe), 7.41 (2 H, m, ArCH), 7.21 (2 H, t, *J* = 8.4, ArCH), 4.26 (2 H, q, *J* = 7.0, CH2), 1.30 (3 H, t, *J* = 7.0, CH3). $\delta_C$ (100 MHz, $CDCl_3$): 193.0 (CS), 169.6 (CO), 163.7 (d, *J* = 254.7 ArCF), 135.8 (C*H*=CSe), 132.10 (d, *J* = 8.6, ArCH), 130.5 (d, *J* = 2.8, ArC), 123.1 (CSe), 116.6 (d, *J* = 22.4, ArCH), 41.07 ($CH_2$), 12.1 ($CH_3$). $\delta_F$ (400 MHz, $CDCl_3$): -111.0. Mass spectrum: HRMS

(ASAP$^+$) found 311.9725, $C_{12}H_{10}ClNOSSe$ requires 311.9738. υ cm$^{-1}$: 2980.0, 2889.4, 1683.7, 1595.1, 1508.3, 1429.3, 1371.4, 1313.5, 1288.5, 1228.7, 1159.2, 1122.6, 1093.6, 921.9, 854.5, 831.3, 802.4, 767.7, 609.5, 547.8, 522.7

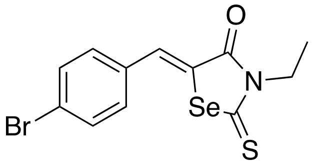
### 6.10.8. (Z)-5-(4-chlorobenzylidene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**23**)



Orange solid, 23 %

δ$_H$ (500 MHz, $CDCl_3$): 8.02 (1 H, s, CHC=CSe), 7.38 (2 H, d, *J* = 7.0, ArCH), 7.26 (2 H, d, *J* = 7.0, ArCH), 4.16 (2 H, q, *J* = 7.1, $CH_2$), 1.21 (3 H, t, *J* = 7.1, $CH_3$). δ$_C$ (100 MHz, $CDCl_3$): 192.8 (CS), 169.6 (CO), 136.7 (*C*H=CSe), 135.6 (ArC), 132.6 (ArC), 131.6 (ArC), 131.1 (CSe), 129.8 (ArCH), 129.7 (ArCH), 43.0 ($CH_2$), 12.0 ($CH_3$). Mass spectrum: HRMS (ASAP$^+$) found 327.9434, $C_{12}H_{11}ClNOS^{76}Se$ requires 327.9442. υ cm$^{-1}$: 2978.1, 2929.9, 1685.8, 1654.9, 1647.2, 1595.1, 1560.4, 1541.1, 1487.1, 1429.3, 1363.7, 1338.6, 1305.8, 1278.8, 1236.4, 1124.5, 893.0, 813.9, 788.9, 734.9, 700.2, 646.2, 605.7, 511.1

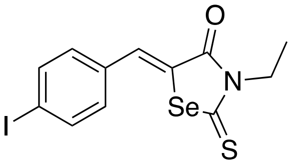
### 6.10.9. (Z)-5-(4-bromobenzylidene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**24**)



Orange solid, 30%

δ$_H$ (400 MHz, CDCl3): 8.10 (1 H, s, CH=CSe), 7.64 (2 H, d, *J* = 8.4, ArCH), 7.27 (2 H, d, *J* = 8.5, ArCH), 4.26 (2 H, q, *J* = 7.0, $CH_2$), 1.30 (3 H, t, *J* = 7.1, $CH_3$). δ$_C$ (100 MHz, $CDCl_3$): 192.8 (CS), 169.6 (CO), 135.7 (CH=CSe), 133.0 (ArC), 132.6 (ArC), 131.92 (2 C, ArCH), 131.7 (2 C, ArCH), 131.2 (ArCBr), 41.1 ($CH_2$), 12.1 ($CH_3$). Mass spectrum: HRMS (ASAP$^+$) found 371.8929, $C_{12}H_{10}BrNOSSe$ requires 371.8937. υ cm$^{-1}$: 2978.1, 2927.9, 1685.8, 1595.1, 1577.8, 1483.3, 1363.7, 1336.7, 1305.8, 1236.4, 1126.4, 1097.5, 1070.5, 1001.1, 894.9, 812.0, 802.4, 786.9, 734.9, 696.3, 605.7, 509.2.

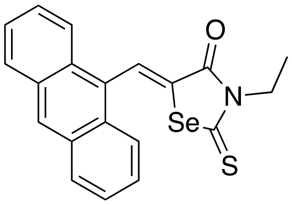
### 6.10.10. Z)-3-ethyl-5-(4-iodobenzylidene)-2-thioxo-1,3-selenazolidin-4-one (**25**)



Orange solid, 82%

$\delta_H$ (400 MHz, $CDCl_3$): 8.16 (1 H, s, CH=CSe), 7.92 (2 H, dt, *J* = 8.2, *J* = 2.1, ArCH), 7.33 (2 H, dt, *J* = 8.2, *J* = 2.1, ArCH), 7.49 (2 H, m, ArCH), 4.10 (2 H, q, *J* = 7.1, $CH_2$), 1.17 (3 H, t, *J* = 7.1, $CH_2$), $\delta_C$ (100 MHz, $CDCl_3$): 192.8 (CS), 169.6 (CO), 138.6 (ArC), 135.8 (CH=CSe), 133.5 (ArCH), 131.2 (CSe), 124.3 (ArCH), 97.3 (CI), 41.1 ($CH_2$), 12.1 ($CH_3$). Mass spectrum: HRMS (ASAP$^+$) found 423.8762, $C_{12}H_{10}BrNOSSe$ requires 422.8693. $\upsilon$ cm$^{-1}$: 3049.5, 2976.2, 2929.9, 1685.8, 1647.2, 1591.3, 1558.5, 1381.0, 13675, 1338.6, 1306.7, 1274.9, 1228.7, 1203.6, 1116.8, 1064.7, 989.5, 925.8, 902.7, 829.4, 794.7, 767.7, 731.0, 669.2, 621.1, 574.8, 542.0

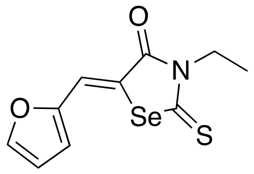
### 6.10.11. (Z)-5-(anthracen-9-ylmethylene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**26**)



Orange solid, 71 %

$\delta_H$ (400 MHz, $CDCl_3$): 8.46 (1 H, s, C*H*C=CSe), 8.20 (1 H, s, ArCH) (1 H, 8.12-7.81 (4 H, m, ArCH), 7.66-7.35 (4 H, m, ArCH), 4.22 (2 H, q, *J* = 7.1, $CH_2$), 1.29 (3 H, t, *J* = 7.1, $CH_3$). $\delta_C$ (100 MHz, $CDCl_3$): 194.1 (CS), 168.4 (CO), 135.8 (*C*H=CSe), 133.0 (CSe), 131.2 (ArC), 129.5 (ArC), 129.2 (ArCH), 128.3 (ArCH), 127.0 (ArCH), 125.7 (ArCH), 125.0 (ArC), 123.5 (ArCH), 41.0 ($CH_2$), 12.3 ($CH_3$). Mass Spectrum: HRMS (ES$^+$) found 394.0136, $C_{20}H_{15}NOSSe$ requires 394.0145. $\upsilon$ cm$^{-1}$: 3047.5, 3024.4, 2997.4, 1705.1, 1676.1, 1622.1, 1604.8, 1382.9, 1371.4, 1346.3, 1307.7, 1282.7, 1236.4, 1157.3, 1121.6, 1049.3, 1020.3, 999.1, 958.6, 881.5, 839.0, 802.4, 783.0, 758.1, 723.3, 694.4, 638.4, 705.7, 582.5, 553.6, 542.0, 526.6, 509.2.

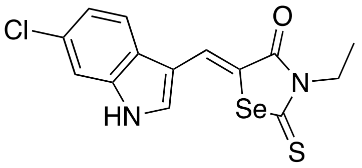
### 6.10.12. (Z)-3-ethyl-5-(furan-2-ylmethylene)-2-thioxo-1,3-selenazolidin-4-one (**27**)



Brown solid, 90 %

$\delta_H$ (400 MHz, CDCl$_3$): 7.87 (1 H, s, CHC=CSe), 7.72 (1 H, s, ArC), 6.86 (1 H, t, *J* = 2.2, ArC), 6.60 (1 H, d, *J* = 2.1, ArC), 4.23 (2 H, q, *J* = 7.2, CH$_2$), 1.28 (3 H, t, *J* = 7.2, CH$_3$). $\delta_C$ (100 MHz, CDCl$_3$): 195.6 (CS), 169.4 (CO), 150.4 (OCAr), 146.8 (*C*H=CSe), 122.1 (ArCH), 121.2 (CSe), 118.2 (ArCH), 113.4 (ArCH). Mass spectrum: HRMS (EI$^+$) found 286.9515, $C_{10}H_9NO_2S^{80}Se$ requires 286.9519. $\upsilon$ cm$^{-1}$: 3049.5, 2976.2, 2931.8, 2358.9, 2322.3, 1685.8, 1593.2, 1558.5, 1541.1, 1508.3, 1489.1, 1473.6, 1456.3, 1431.7, 1367.5, 1338.6, 1311.6, 1230.6, 1205.5, 1116.8, 1068.0, 925.8, 881.5, 765.7, 731.0, 667.4, 623.0, 574.8, 543.0.

### 6.10.13. (Z)-5-((6-chloro-1H-indol-3-yl)methylene)-3-ethyl-2-thioxo-1,3-selenazolidin-4-one (**28**)

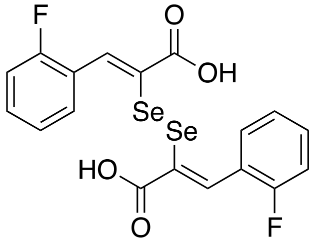


Orange solid, 60%

$\delta_H$ (400 MHz, DMSO-d$_6$): 12.43 (1 H, s, NH), 8.48 (1 H, s, CH=Cse), 8.09 (1 H, d, *J* = 1.7, ArCH), 7.77 (1 H, s, ArCH), 7.51 (1 H, d, *J* = 8.6, ArCH), 7.26 (1 H, dd, *J* = 8.6, *J* = 1.7, ArCH), 4.11 (2 H, q, *J* = 7.2, CH$_2$), 1.18 (3 H, t, *J* = 7.2, CH$_3$). $\delta_C$ (400 MHz, DMSO-d$_6$): 193.4 (CS), 169.0 (CO), 135.4 (CSe), 131.1 (ArCH), 130.2 (*C*H=CSe), 128.6 (ArC), 126.5 (ArC) 123.9 (ArCH) 118.7 (ArCH), 116.1 (ArC), 114.5 (ArCH), 112.0 (ArC) ,40.4 (CH$_2$), 12.6 (CH$_3$). HRMS (ES$^+$) found 365.9482, $C_{14}H_{11}ClN_2OS^{76}Se$ requires 365.9472. $\upsilon$ cm$^{-1}$: 3219.2, 3049.5, 2976.2, 2929/9, 2360.9, 2322.3, 1683.9, 1648.4, 1591.3, 1558.5, 1541.1, 1508.3, 1471.7, 1456.3, 1400.3, 1367.5, 1338.6, 1309.1, 1226.7, 1203.6, 1116.8, 1067.0, 929.7, 890.4, 767.7, 731.0, 667.4, 621.1, 574.8,

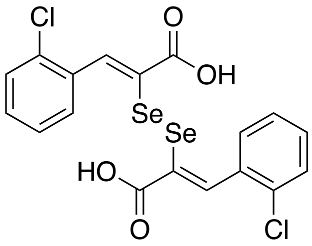
## 6.11. General Method for Hydrolysis and Oxidation Reaction

### 6.11.1. (2Z,2'Z)-2,2'-diselanediylbis(3-(2-fluorophenyl)acrylic acid) (**29**)



(Z)-3-ethyl-5-(2-fluorobenzylidene)-2-thioxo-1,3-selenazolidin-4-one (**16**, 21 mg, 0.066 mmol) was dissolved in 15% NaOH/methanol (2:1) and heated to 95 °C for 90 minutes. The mixture was cooled on ice and acidified to pH 2 with HCl, the resultant yellow solid was collected by filtration. The solid was dissovled in 15% NaOH, washed with dichloromethane (3x 10 mL), acidifed to pH 2 with HCl and extracted with ethyl acetate (3x 5 mL). The combined organic fractions were washed with cold deionised water, brine, dried with $MgSO_4$ and solvent removed *in vacuo*. The product was isolated as yellow crystals (11 mg, 0.023 mmol, 29%). $\delta_H$ (400 MHz, MeOH-$d_4$): 8.14 (2 H, s, CH=CSe), 7.70 (2 H, d, *J* = 7.8, ArCH), 7.32 (2 H, m, ArCH), 7.08 (4 H, m, ArCH) $\delta_C$ (100 MHz MeOH-d4): 167.33 (CO), 160.39 (d, *J* = 249.6, CF) 138.34 (d, *J* = 4.3, ArCH=C), 131.41 (d, *J* = 8.8, ArCH), 130.07 (d, *J* = 1.7, ArCH), 123.58 (CSe), 123.18 (d, *J* = 12.7, ArCH), 115.00 (d, *J* = 22.1, ArC), 114.83 (d, *J* = 22.9, ArCH). $\delta_F$ (400 MHz, MeOH-$d_4$): -111.02. Mass spectrum: HRMS (AP$^+$) found 472.9011, [$C_{18}H_{12}Cl_2O_4Se_2$]- [$H_2O$] requires 472.9007. $\upsilon$ cm$^{-1}$: 3298.3, 2924.1, 2850.8, 1683.9, 1508.3, 1419.6, 1398.4, 1163.1, 1068.6, 1008.8, 800.5, 545.9.

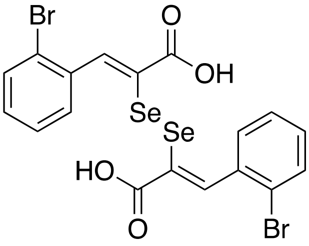
### 6.11.2. (2Z,2'Z)-2,2'-diselanediylbis(3-(2-chlorophenyl)acrylic acid) (**30**)



Yellow solid, 31%

$\delta_H$ (400 MHz, DMSO-$d_6$): 7.94 (2 H, s, CH=CSe), 7.42 (2 H, d, *J* = 7.4, ArCH), 7.32 (2 H, d, *J* = 7.2, ArCH), 7.24 (4 H, m ArCH). $\delta_C$ (100 MHz DMSO-$d_6$) 167.19 (CO), 142.75 (Ar*C*H=C), 134.05 (CCl) 133.85 (ArC), 130.96 (ArCH), 130.31 (ArCH), 129.04 (ArCH), 128.43 (CSe), 126.30 (ArCH).Mass Spectrum: HRMS (AP$^+$) found 498.8459, requires 498.8484 [$C_{18}H_{11}Cl_2O_3Se_2$]- [$H_2O$]. υ cm$^{-1}$: 3383.1, 2914.4, 2848.9, 1670.4, 1560.4, 1467.8, 1419.6, 1199.72, 1130.3, 1020.3, 950.91, 833.3, 798.5, 721.4, 518.9

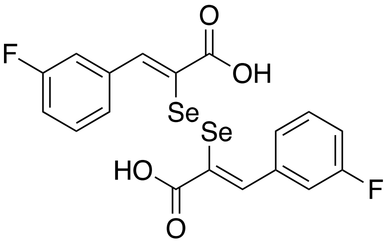
### 6.11.3. (2Z,2'Z)-2,2'-diselanediylbis(3-(2-bromophenyl)acrylic acid) (**31**)



Yellow solid, 32%

$\delta_H$ (400 MHz, DMSO-$d_6$): 7.84 (2 H, s, CH=CSe), 7.66 (2 H, d, *J* = 7.9, ArCH), 7.45 (2 H, d, *J* = 7.9, ArCH), 7.42 (2 H, t, *J* = 7.9, ArCH), 7.32 (2 H, t, *J* = 7.9, ArCH). $\delta_C$ (100 MHz $CDCl_3$) 167.16 (CO), 145.01 (Ar*C*=C), 135.99 (ArC), 135.77 (ArC), 132.27 (ArCH), 130.96 (ArCH), 130.39 (ArCH), 126.92 (ArCH), 123.60 (CSe). HRMS (ES-) found 602.7399, $C_{18}H_{12}{}^{79}Br_2O_4\ {}^{77}Se_2$ requires 602.7422. υ cm$^{-1}$: 3327.2, 3049.5, 2976.2, 2929.9, 1685.7, 1593.2, 1558.5, 1541.1, 1508.3, 1473.6, 1404.2, 1367.5, 1336.7, 1311.6, 1276.9, 1226.7, 1203.6, 1116.8, 1064.7, 931.6, 765.7, 729.1, 666.5, 623.0, 574.8.

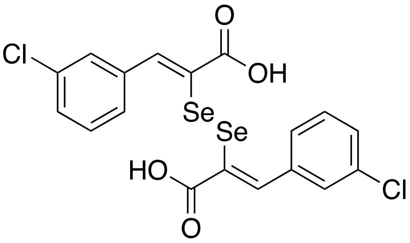
### 6.11.4. (2Z,2'Z)-2,2'-diselanediylbis(3-(3-fluorophenyl)acrylic acid) (**32**)



Yellow solid, 41%

$\delta_H$ (500 MHz, DMSO-d$_6$): 8.07 (2 H, s, CH=CSe), 7.32 - 7.03 (8 H, m, ArCH). $\delta_C$ (100 MHz, DMSO-d$_6$): 167.23 (CO), 162.05 (d, *J* = 247.8, CF), 143.71 (Ar*C*H=C), 137.61 (d, *J* = 7.8, ArCH), 130.66 (d, *J* = 8.8, ArC), 127.70 (CSe), 126.98 (d, *J* = 2.3, ArCH), 117.02 (d, *J* = 22.4, ArCH), 116.82 (d, *J* = 21.5, ArCH). $\delta_F$ (500 MHz, DMSO-d$_6$): -122.77. Mass spectrum: HRMS (AP$^+$) found 489.9034, [$C_{18}H_{12}F_2O_4Se_2$] requires 488.8962. υ cm$^{-1}$: 3039.85, 2922.2, 1674.2, 1575.8, 1481.3, 1446.6, 1275.0, 1228.7, 1170.8, 1143.8, 1004.9, 952.8, 925.8, 873.8, 790.8, 752.2, 711.7, 677.0, 636.5, 590.2, 520.8. Purity >95% by LC/MS.

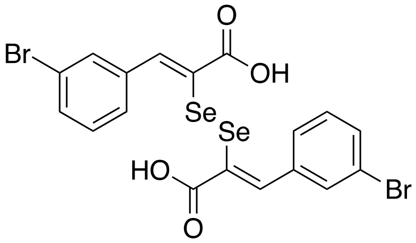
### 6.11.5. (2Z,2'Z)-2,2'-diselanediylbis(3-(3-chlorophenyl)acrylic acid) (**33**)



Orange solid, 23 %

$\delta_H$ (400 MHz, DMSO-d$_6$): 8.18 (2 H, s, CH=CSe), 7.48 (2 H, d, *J* = 7.8, ArCH), 7.32 (2 H, d, *J* = 7.8, ArCH), 7.28-7.16 (4 H, m, ArCH). $\delta_C$ (100 MHz, DMSO-d$_6$): 167.7 (COOH), 144.8 (*C*H=CSe), 133.5 (CCl), 130.5 (ArCH), 129.6 (ArCH), 128.3 (ArCH), 127.5 (ArCH), 127.3 (ArCH), 119.4 (CSe). Mass spectrum: HRMS (AP$^+$) found 498.8459, [$C_{18}H_{11}Cl_2O_3Se_2$]- [$H_2O$] requires 498.8484. υ cm$^{-1}$: 3346.5, 2924.1, 2850.8, 1654.9, 1589.3, 1560.4, 1375.3, 1220.9, 1151.5, 1045.4, 1022.3, 995.3, 783.1, 761.9, 682.8

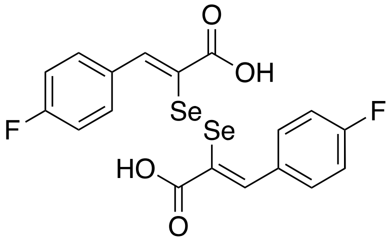
### 6.11.6. (2Z,2'Z)-2,2'-diselanediylbis(3-(3-bromophenyl)acrylic acid) (**34**)



Orange solid, 27 %

$\delta_H$ (400 MHz, acetone-$d_6$): 7.97 (2 H, s, CH=CSe), 7.70 (2 H, s, ArCH), 7.62 (2 H, d, *J* = 7.7, ArCH), 7.36-7.25 (2 H, m, ArCH), 7.13 (2 H, d, *J* = 7.7, ArCH). $\delta_C$ (100 MHz, acetone-d8): 166.2 (COOH), 145.6 (CH=CSe), 138.2 (ArC), 132.1 (ArCH), 131.2 (ArCH), 131.0 (ArCH), 130.0 (ArCH), 123.9 (CBr), 122.6 (CSe). Mass spectrum: HRMS (ES-) found 608.7349, $C_{18}H_{12}Br_2O_4Se_2$ requires 608.7355. υ $cm^{-1}$: 3402.4, 2918.3, 1678.1, 1637.6, 1585.5, 1485.2, 1396.5, 1259.5, 1224.8, 1103.4, 1068.6, 1006.8, 926.8, 877.6, 848.7, 800.5, 758.0, 696.3, 578.6, 545.9.

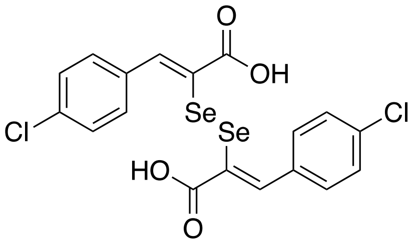
### 6.11.7. (2Z,2'Z)-2,2'-diselanediylbis(3-(4-fluorophenyl)acrylic acid) (**35**)



Yellow solid, 48 %

$\delta_H$ (500 MHz, DMSO-$d_6$): 7.91 (2 H, s, CH=CSe), 7.82 (4 H, dd, *J* = 7.6 *J* = 3.6, ArCH), 7.37 (4 H, d, *J* = 7.6, ArCH). $\delta_C$ (100 MHz, DMSO-$d_6$): 166.9 (CO), 165.2 (d, *J* = 249.7, ArCF), 132.5 (ArCH, d, *J* = 8.5), 131.2 (CH=CSe), 128.0 (CSe), 128.6 (ArC), 116.1 (ArCH, d, *J* = 22.4), 115.1 (ArCH). $\delta_F$ (500 MHz, DMSO-$d_6$): -107.08. Mass spectrum: HRMS (EI$^+$) found 483.9107, $C_{18}H_{12}F_2O_4Se_2$ requires 483.9094. υ $cm^{-1}$: 3082.3, 2885.5, 1670.4, 1598.9, 1506.4, 1421.5, 1290.4, 1224.8, 1157.3, 1130.3, 852.5, 844.8, 767.7, 609.5, 547.8. Purity >95% by LC/MS.

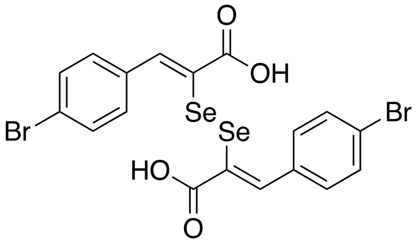
### 6.11.8. (2Z,2'Z)-2,2'-diselanediylbis(3-(4-chlorophenyl)acrylic acid) (**36**)



Orange solid, 44 %

$\delta_H$ (400 MHz, DMSO-d$_6$): 7.72 (2 H, s, CH=CSe), 7.58 (4 H, d, *J* = 8.7, ArCH), 7.44 (4 H, d, *J* = 8.7, ArCH). $\delta_C$ (100 MHz, DMSO-d$_6$): 167.3 (CO), 143.9 (*C*H=CSe), 134.1 (CCl), 133.9 (ArC), 132.91 (ArCH), 128.8 (ArCH), 126.9 (CSe). Mass spectrum: HRMS (EI$^-$) found 521.8410, $C_{18}H_{12}Cl_2O_4Se_2$ requires 521.8443. υ cm$^{-1}$: 3442.9, 2960.7, 2916.4, 2848.9, 1680.0, 1591.3, 1489.1, 1400.3, 1257.6, 1087.9, 1010.7, 792.7, 731.0, 701.0, 582.5, 547.8, 511.1. Purity >95% by LC/MS.

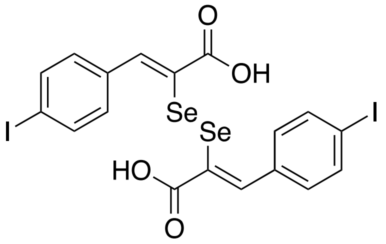
### 6.11.9. (2Z,2'Z)-2,2'-diselanediylbis(3-(4-bromophenyl)acrylic acid) (**37**)



Orange solid, 30 %

$\delta_H$ (400 MHz, MeOD-d$_4$): 7.94 (2 H, s, CH=CSe), 7.61 (4 H, d, *J* = 8.3, ArCH), (4 H, d, *J* = 8.3, ArCH). $\delta_C$ (100 MHz, MeOD-d$_4$): 167.4 (CO), 138.0 (*C*H=CSe), 134.2 (ArC), 131.4 (ArCH), 130.0 (ArCH), 129.7 (CBr), 127.3 (CSe). Mass spectrum: HRMS (ES$^-$) found 602.7393, $C_{18}H_{11}O_4{}^{77}Se_2{}^{79}Br_2$ requires 602.7422. υ cm$^{-1}$: 2958.8, 2922.2, 2850.8, 1676.1, 1653.0, 1585.5, 1541.1, 1485.2, 1456.3, 1417.7, 1396.5, 1319.3, 1257.6, 1176.6, 1070.5, 1006.8, 848.7, 798.5, 756.1, 698.2, 607.6, 545.9, 505.4. Purity >95% by LC/MS.

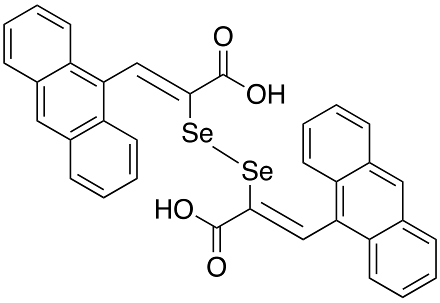
### 6.11.10. (2Z,2'Z)-2,2'-diselanediylbis(3-(4-iodophenyl)acrylic acid) (**38**)



Brown solid, 41%

$\delta_H$ (500 MHz, DMSO-d$_6$): 7.81 (2 H, s, CH=CSe), 7.76 (4 H, d, *J* = 7.9, ArCH), 7.37 (4 H, d, *J* = 7.9, ArCH). $\delta_C$ (100 MHz, DMSO-d$_6$): 167.3 (CO), 144.2 (*C*H=CSe), 139.7 (ArCH), 137.5 (ArCH), 134.7 (ArC), 132.5 (CSe), 94.8 (CI). Mass spectrum: HRMS (ES$^-$) found 698.7162 requires $C_{18}H_{11}{}^{127}I_2O_4{}^{77}Se_2$ 698.7145. υ cm$^{-1}$: 3385.1, 3057.2, 2976.2, 2929.9, 1683.9, 1662.6, 1558.5, 1541.1, 1508.6, 1371.0, 1340.1, 1228.7, 1202.6, 1117.8, 1022.3, 994.7, 823.6, 765.7, 729.1, 667.4, 623.0, 574.8, 542.7. Purity >95% by LC/MS.

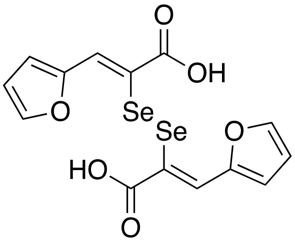
### 1.1.1. (2Z,2'Z)-2,2'-diselanediylbis(3-(anthracen-9-yl)acrylic acid) (**39**)



Brown solid, 22%

$\delta_H$ (500 MHz, DMSO-d$_6$): 8.59 (2 H, s, CH=CSe), 8.18-7.81 (6 H, m, ArC), 7.60-7.29 (12 H, m, ArC). $\delta_C$ (100 MHz, DMSO-d$_6$): 166.7 (CO), 135.0 (CH=CSe), 131.3 (ArCH), 129.2 (ArCH), 129.1 (ArC), 128.4 (ArCH), 126.7 (ArCH), 126.0 (ArCH), 125.9 (ArC), 125.7 (ArC), 125.6 (CSe). Mass spectrum: HRMS (ES$^+$) found 649.0017, $C_{34}H_{23}O_4{}^{77}Se_2$ requires 648.9994. υ cm$^{-1}$: 3327.2, 3049.5, 2920.2, 2850.8, 1757.2, 1670.4, 1587.4, 1558.5, 1541.1, 1508.3, 1458.2, 1417.7, 1257.6, 1229.1, 1089.8, 1026.5, 1001.9, 933.6. 892.8, 844.8, 729.2, 684.7, 590.2, 513.1.

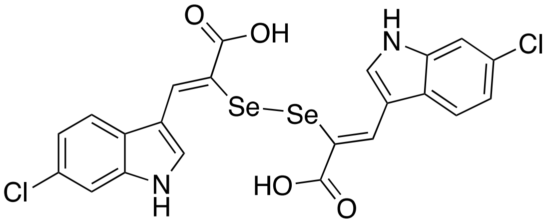
### 6.11.11. (2Z,2'Z)-2,2'-diselanediylbis(3-(furan-2-yl)acrylic acid) (**40**)



Brown solid, 48%

$\delta_H$ (400 MHz, DMSO-d$_6$): 7.90 (2 H, d, *J* = 1.3, ArCH), 7.7 (2 H, s, CH=CSe), 7.15 (2 H, d, *J* = 3.6, ArCH), 6.67 (2 H, dd, *J* = 3.6, *J* = 1.3 ArCH). $\delta_C$ (100 MHz, DMSO-d$_6$) 167.3 (COOH), 150.4 (ArC), 146.6 (CH=CSe), 131.8 (ArCH), 121.4 (CSe), 118.6 (ArCH), 113.3 (ArCH). Mass spectrum: HRMS (ES$^-$) found 456.8713, $C_{14}H_{10}O_6Na^{80}Se_2$ requires 456.8706. $\upsilon$ cm$^{-1}$: 3115.0. 2918.3, 2792.9, 1670.4, 1585.5, 1527.6, 1460.1, 1411.9, 1384.9, 1259.5, 1211.3, 1172.7, 1147.6, 1087.9, 1022.3, 939.3, 918.1, 881.5, 788.9, 759.9, 748.4, 684.7, 651.9, 630.7, 588.3. Purity >95% by LC/MS.

### 6.11.12. (2Z,2'Z)-2,2'-diselanediylbis(3-(6-chloro-1H-indol-3-yl)acrylic acid) (**41**)



Brown solid, 38%

$\delta_H$ (400 MHz, DMSO-d$_6$): 12.10 (1 H, s, NH), 8.02 (1 H, s, C*H*=CSe), 7.91 (1 H, s, 7.42 (1 H, d, *J* = 8.6, ArCH), 7.18-7.07 (2 H, m, ArCH). $\delta_C$ (100 MHz, DMSO-d$_6$): 173.9 (CO), 134.5 (CCl), 129.2 (ArC), 129.2 (*C*H=CSe), 124.6 (ArC), 122.1 (ArCH), 122.0 (CSe), 117.8 (ArCH), 113.8 (ArCH), 112.0 (ArCH) 111.7 (ArC). $\upsilon$ cm$^{-1}$: 3288.6 3057.2, 2976.2, 2931.8, 1683.9, 1653.0, 1558.5, 1541.1, 1508.3, 1456.3, 1419.6, 1338.6, 1228.7, 1122.9, 1099.7, 937.4, 891.1, 796.6, 728.7, 669.2, 594.6, 507.9.

# References

1. Miller, D. J., Adams, S. E., Hallett, M. B. & Allemann, R. K. Calpain-1 inhibitors for selective treatment of rheumatoid arthritis: what is the future? *Future Med. Chem.* **5,** 2057–2074 (2013).
2. Mahajan, T. D. & Mikuls, T. R. Recent advances in the treatment of rheumatoid arthritis. *Curr. Opin. Rheumatol.* **30,** 231–237 (2018).
3. Pelletier, J. P., Martel-Pelletier, J. & Abramson, S. B. Osteoarthritis, an inflammatory disease: potential implication for the selection of new therapeutic targets. *Arthritis Rheum.* **44,** 1237–47 (2001).
4. Arend, W. P. The pathophysiology and treatment of rheumatoid arthritis. *Off. J. Am. Coll. Rheumatol.* **40,** 595–597 (1997).
5. Otero, M. & Goldring, M. B. Review Cells of the synovium in rheumatoid arthritis Chondrocytes. **13,** 1–13 (2007).
6. Burrage, P. S., Mix, K. S. & Brinckerhoff, C. E. Matrix Metalloproteinases: Role in Arthritis. *Front. Biosci.* **11,** 529–543 (2006).
7. Yuan, G. M. D. *et al.* Characterization of cells from pannus-like tissue over articular cartilage of advanced osteoarthritis. **12,** 38–45 (2004).
8. Lefèvre, S. *et al.* Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. *Nat. Med.* **15,** 1414–1422 (2009).
9. Fossati, G., Bucknall, R. C. & Edwards, S. W. Insoluble and soluble immune complexes activate neutrophils by distinct activation mechanisms: changes in functional responses induced by priming with cytokines. *Ann. Rheum. Dis.* **61,** 13–9 (2002).
10. T., T. Interleukin-6 inhibition in inflammatory diseases: Results achieved and tasks to accomplish. *J. Scleroderma Relat. Disord.* **2,** S20–S28 (2017).
11. Hallett, M. B. & Dewitt, S. Ironing out the wrinkles of neutrophil phagocytosis. *Trends Cell Biol.* **17,** 209–214 (2007).
12. Springer, T. A. Traffic signals for lymphocyte recirculation and leukocyte emigration: The multistep paradigm. *Cell* **76,** 301–314 (1994).
13. Mcdonald, B. & Kubes, P. Chemokines : Sirens of Neutrophil Recruitment — but Is It Just One Song ? *Immunity* **33,** 148–149 (2010).
14. Dewitt, S., Francis, R. J. & Hallett, M. B. Ca2+ and calpain control membrane expansion during the rapid cell spreading of neutrophils. **126,** 4627–4635 (2013).

15. Ananthakrishnan, R. & Ehrlicher, A. The Forces Behind Cell Movement. **3,** 303–317 (2007).

16. Irisa, M. A Brownian Ratchet Model of Actin Polymerization Motor by using Extended Scaled Particle Theory. *AIP Conf. Proc.* **708,** 294–297 (2004).

17. Goll, D. E., Thompson, V. F., Li, H., Wei, W. E. I. & Cong, J. The Calpain System. **1990,** 731–801 (2003).

18. Strobl, S. *et al.* The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 588–592 (2000).

19. Sorimachi, H., Hata, S. & Ono, Y. Calpain chronicle - enzyme family under multidisciplinary characterization. *Proc. Japan Acad. Ser. B* **87,** 287–327 (2011).

20. Goll, D. E., Thompson, V. F., Li, H., Wei, W. & Cong, J. The calpain system. *Physiol. Rev.* **83,** 731–801 (2003).

21. Sato, K. & Kawashima, S. Calpain function in the modulation of signal transduction molecules. *Biol. Chem.* **382,** 743–751 (2001).

22. Ji, J., Su, L. & Liu, Z. Critical role of calpain in inflammation (Review). *Biomed. Reports* **5,** 647–652 (2016).

23. Ono, Y. & Sorimachi, H. Calpains - An elaborate proteolytic system. *Biochim. Biophys. Acta - Proteins Proteomics* **1824,** 224–236 (2012).

24. Hosfield, C. M., Elce, J. S., Davies, P. L. & Jia, Z. Crystal structure of calpain reveals the structural basis for Ca(2+)-dependent protease activity and a novel mode of enzyme activation. *EMBO J.* **18,** 6880–6889 (1999).

25. Fersht, A. *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding. Lavoisier.Fr* **13409,** (1999).

26. Campbell, R. L. & Davies, P. L. Structure – function relationships in calpains 1. **351,** 335–351 (2012).

27. Form, H. *et al.* Molecular Cloning of the cDNA for the Large Subunit of the. **27,** 8122–8128 (1988).

28. Sorimachi, H., Hata, S. & Ono, Y. Impact of genetic insights into calpain biology. *J. Biochem.* **150,** 23–37 (2011).

29. Dutt, P. *et al.* m-Calpain is required for preimplantation embryonic development in mice. *BMC Dev. Biol.* **6,** 1–11 (2006).

30. Pal, G. P., De Veyra, T., Elce, J. S. & Jia, Z. Crystal Structure of a μ-like Calpain Reveals a Partially Activated Conformation with Low Ca2+ Requirement. *Structure* **11,** 1521–1526 (2003).

31. Badugu, R., Garcia, M., Bondada, V., Joshi, A. & Geddes, J. W. N Terminus of Calpain 1 Is a Mitochondrial Targeting Sequence *. **283,** 3409–3417 (2008).
32. Zimmerman, U.-J. P. & Schlaepfer, W. W. Two-stage autolysis of the catalytic subunit initiates activation of calpain I. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1078,** 192–198 (1991).
33. Congs, J., Gollq, D. E., Peterson, A. M. & Kapprell, H. The Role of Autolysis in Activity of the Ca2 + -dependent Proteinases. **264,** 10096–10103 (1989).
34. Li, H., Thompson, V. F. & Goll, D. E. Effects of autolysis on properties of μ- and m-calpain. *Biochim. Biophys. Acta - Mol. Cell Res.* **1691,** 91–103 (2004).
35. Hanna, R. a, Campbell, R. L. & Davies, P. L. Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature* **456,** 409–412 (2008).
36. Moldoveanu, T. *et al.* A Ca2+ Switch Aligns the Active Site of Calpain. *Cell* **108,** 649–660 (2002).
37. Todd, B. *et al.* A structural model for the inhibition of calpain by calpastatin: Crystal structures of the native domain VI of calpain and its complexes with calpastatin peptide and a small molecule inhibitor. *J. Mol. Biol.* **328,** 131–146 (2003).
38. Arthur, J. S. C., Gauthier, S. & Elce, J. S. Active site residues in m-calpain: identification by site-directed mutagenesis. **368,** 397–400 (1995).
39. Low, K. E., Partha, S. K., Davies, P. L. & Campbell, R. L. Allosteric inhibitors of calpains : Reevaluating inhibition by PD150606 and LSEAL. *Biochim. Biophys. Acta* **1840,** 3367–3373 (2014).
40. Siklos, M., BenAissa, M. & Thatcher, G. R. J. Cysteine proteases as therapeutic targets: does selectivity matter? A systematic review of calpain and cathepsin inhibitors. *Acta Pharm. Sin. B* **5,** 506–519 (2015).
41. Tompa, P., Emori, Y., Sorimachi, H., Suzuki, K. & Friedrich, P. Domain III of Calpain Is a Ca 2 ϩ -Regulated Phospholipid-Binding Domain. **1339,** 1333–1339 (2001).
42. Shao, H. *et al.* Spatial localization of m-calpain to the plasma membrane by phosphoinositide biphosphate binding during epidermal growth factor receptor-mediated activation. *Mol. Cell. Biol.* **26,** 5481–5496 (2006).
43. Maki, M., Maemoto, Y., Osako, Y. & Shibata, H. Evolutionary and physical linkage between calpains and penta-EF-hand Ca 2+-binding proteins. *FEBS J.* **279,** 1414–1421 (2012).
44. Bhattacharya, S., Bunick, C. G. & Chazin, W. J. Target selectivity in EF-hand

calcium binding proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **1742,** 69–79 (2004).

45. Lewit-Bentley, A. & Rety, S. EF-hand calcium-binding proteins. *Curr. Opin. Struct. Biol.* **10,** 637–643 (2000).
46. Lewit-Bentley, A, R. S. EF-hand calcium-binding proteins. **10,** 637–643 (2000).
47. Moldoveanu, T., Hosfield, C. M., Lim, D., Elce, J. S. & Jia, Z. Switch Aligns the Active Site of Calpain between calpain and its endogenous inhibitors like cal. *Cell* **108,** 649–660 (2002).
48. Adams, S. E. *et al.* The structural basis of differential inhibition of human calpain by indole and phenyl α-mercaptoacrylic acids. *J. Struct. Biol.* **187,** 236–241 (2014).
49. Campbell, R. L. & Davies, P. L. Structure–function relationships in calpains. *Biochem. J.* **447,** 335–351 (2012).
50. Pal, G. P., Elce, J. S. & Jia, Z. Dissociation and Aggregation of Calpain in the Presence of Calcium. *J. Biol. Chem.* **276,** 47233–47238 (2001).
51. Ravulapalli, R., Campbell, R. L., Gauthier, S. Y., Dhe-Paganon, S. & Davies, P. L. Distinguishing between calpain heterodimerization and homodimerization. *FEBS J.* **276,** 973–982 (2009).
52. Dutt, P., Spriggs, C. N., Davies, P. L., Jia, Z. C. & Elce, J. S. Origins of the difference in Ca2+ requirement for activation of mu- and m-calpain. *Biochem. J.* **367,** 263–269 (2002).
53. Richard, I. *et al.* Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* **81,** 27–40 (1995).
54. Hata, S., Ueno, M., Kitamura, F. & Sorimachi, H. Efficient expression and purification of recombinant human m-calpain using an Escherichia coli expression system at low temperature. *J. Biochem.* **151,** 417–422 (2012).
55. Leloup, L. *et al.* M-calpain activation is regulated by its membrane localization and by its binding to phosphatidylinositol 4,5-bisphosphate. *J. Biol. Chem.* **285,** 33549–33566 (2010).
56. Cuerrier, D., Moldoveanu, T., Inoue, J., Davies, P. L. & Campbell, R. L. Calpain inhibition by ??-ketoamide and cyclic hemiacetal inhibitors revealed by X-ray crystallography. *Biochemistry* **45,** 7446–7452 (2006).
57. Bagur, R. & Hajnooczky, G. Intracellular Ca 2 + Sensing: Its Role in Calcium Homeostasis and Signaling. *Mol. Cell* **66,** 780–788 (2017).
58. Moldoveanu, T., Jia, Z. & Davies, P. L. Calpain Activation by Cooperative Ca 2 ⁒ Binding at Two Non-EF-hand Sites *. **279,** 6106–6114 (2004).

59. Ono, Y., Saido, T. C. & Sorimachi, H. Calpain research for drug discovery: challenges and potential. *Nat. Rev. Drug Discov.* **12,** 854–876 (2016).
60. Santella, L., Kyozuka, K., De Riso, L. & Carafoli, E. Calcium, protease action, and the regulation of the cell cycle. *Cell Calcium* **23,** 123–130 (1998).
61. Davies, E. V. & Hallett, M. B. High micromolar Ca2+ beneath the plasma membrane in stimulated neutrophils. *Biochem. Biophys. Res. Commun.* **248,** 679–683 (1998).
62. Franco, S., Perrin, B. & Huttenlocher, A. Isoform specific function of calpain 2 in regulating membrane protrusion. *Exp. Cell Res.* **299,** 179–187 (2004).
63. Adams, S. E. *et al.* Conformationally restricted calpain inhibitors. *Chem. Sci.* **6,** 6865–6871 (2015).
64. Nishino. Acyl-coA-binding protein is a potent m-calpain activator. *J. Biol. Chem.* **275,** 82–86 (2000).
65. Fernández-Montalván, A. *et al.* μ-Calpain binds to lipid bilayers via the exposed hydrophobic surface of its Ca2+-activated conformation. *Biol. Chem.* **387,** 617–627 (2006).
66. Melloni, E., Michetti, M., Salamino, F., Minafra, R. & Pontremoli, S. Modulation of the Calpain Autoproteolysis by Calpastatin and Phospholipids in an active form throughout an autoproteolytic process, triggered by the binding of calcium present available on the nature of the signal induced by binding of Ca 2 / , it is wel. *Biochem. Biophys. Res. Commun.* **197,** 193–197 (1996).
67. Elce, J. S., Hegadorn, C. & Arthur, J. S. C. Autolysis, Ca2+ requirement, and heterodimer stability in m-calpain. *J. Biol. Chem.* **272,** 11268–11275 (1997).
68. Sorimachi, H. & Ono, Y. Regulation and physiological roles of the calpain system in muscular disorders. *Cardiovasc. Res.* **96,** 11–22 (2012).
69. Shinkai-Ouchi, F. *et al.* Predictions of Cleavability of Calpain Proteolysis by Quantitative Structure-Activity Relationship Analysis Using Newly Determined Cleavage Sites and Catalytic Efficiencies of an Oligopeptide Array. *Mol. Cell. Proteomics* **15,** 1262–1280 (2016).
70. Cuerrier, D., Moldoveanu, T. & Davies, P. L. Determination of peptide substrate specificity for μ-calpain by a peptide library-based approach: The importance of primed side interactions. *J. Biol. Chem.* **280,** 40632–40641 (2005).
71. Tompa, P. *et al.* On the sequential determinants of calpain cleavage. *J. Biol. Chem.* **279,** 20775–20785 (2004).
72. Alías, M., Ayuso-Tejedor, S., Fernández-Recio, J., Cativiela, C. & Sancho, J.

Helix propensities of conformationally restricted amino acids. Non-natural substitutes for helix breaking proline and helix forming alanine. *Org. Biomol. Chem.* **8,** 788–792 (2010).

73. Momeni, H. R. Role of calpain in apoptosis. *Cell J.* **13,** 65–72 (2011).

74. Kerr, J. F., Winterford, C. M. & Harmon, B. V. Apoptosis. Its significance in cancer and cancer therapy. *Cancer* **73,** 2013–2026 (1994).

75. Momeni, H. R., Azadi, S. & Kanje, M. Calpain activation and apoptosis in motor neurons of cultured adult mouse spinal cord. *Funct. Neurol.* **22,** 105–110 (2007).

76. Santella, L. & Carafoli, E. Calcium signaling in the cell nucleus. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **11,** 1091–1109 (1997).

77. Das, M., Subbayya Ithychanda, S., Qin, J. & Plow, E. F. Mechanisms of talin-dependent integrin signaling and crosstalk. *Biochim. Biophys. Acta - Biomembr.* **1838,** 579–588 (2014).

78. Nagano, M., Hoshino, D., Koshikawa, N., Akizawa, T. & Seiki, M. Turnover of focal adhesions and cancer cell migration. *Int. J. Cell Biol.* **2012,** (2012).

79. Robles, E., Huttenlocher, A. & Gomez, T. M. Filopodial calcium transients regulate growth cone motility and guidance through local activation of calpain. *Neuron* **38,** 597–609 (2003).

80. Saraiva, N. *et al.* HGAAP promotes cell adhesion and migration via the stimulation of store-operated Ca2+ entry and calpain 2. *J. Cell Biol.* **202,** 699–713 (2013).

81. Schollmeyer, J. E. Calpain II involvement in mitosis. *Science (80-. ).* **240,** 911–913 (1988).

82. Jánossy, J. *et al.* Calpain as a multi-site regulator of cell cycle. *Biochem. Pharmacol.* **67,** 1513–1521 (2004).

83. Mellgren, R. L. Evidence for participation of a calpain-like cysteine protease in cell cycle progression through late G1phase. *Biochem. Biophys. Res. Commun.* **236,** 555–558 (1997).

84. Leloup, L. & Wells, A. Calpains as potential anti-cancer targets. *Expert Opin. Ther. Targets* **15,** 309–323 (2011).

85. Martin, T. A., Ye, L., Sanders, A. J., Lane, J. & Jiang, W. G. Cancer Invasion and Metastasis : Molecular and Cellular Perspective Cancer Invasion and Metastasis : The Role of Cell Adhesion Molecules. *Madame Curie Biosci. Database* (2013).

86. Mamoune, A., Luo, J.-H., Lauffenburger, D. A. & Wells, A. Calpain-2 as a

target for limiting prostate cancer invasion. *Cancer Res.* **63,** 4632–4640 (2003).

87. Ferreira, A. Calpain Dysregulation in Alzheimer's Disease. *ISRN Biochem.* **2012,** 1–12 (2012).
88. Tu, S. Oligomeric A β-induced synaptic dysfunction in Alzheimer' s disease. *Mol. Neurodegener.* **9,** 1–12 (2014).
89. Saito, K., Elce, J. S., Hamos, J. E. & Nixon, R. A. Widespread activation of calcium-activated neutral proteinase (calpain) in the brain in Alzheimer disease: a potential molecular basis for neuronal degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 2628–2632 (1993).
90. Kling, A. *et al.* Discovery of Novel and Highly Selective Inhibitors of Calpain for the Treatment of Alzheimer's Disease: 2-(3-Phenyl-1H-pyrazol-1-yl)-nicotinamides. *J. Med. Chem.* **60,** 7123–7138 (2017).
91. Liang, B., Duan, B. Y., Zhou, X. P., Gong, J. X. & Luo, Z. G. Calpain activation promotes BACE1 expression, amyloid precursor protein processing, and amyloid plaque formation in a transgenic mouse model of alzheimer disease. *J. Biol. Chem.* **285,** 27737–27744 (2010).
92. Bano, D. & Nicotera, P. Ca2+ signals and neuronal death in brain ischemia. *Stroke* **38,** 674–676 (2007).
93. Luo, T. *et al.* PD150606 protects against ischemia/reperfusion injury by preventing μ-calpain-induced mitochondrial apoptosis. *Arch. Biochem. Biophys.* **586,** 1–9 (2015).
94. Yamashima, T. Ca2+-dependent proteases in ischemic neuronal death. A conserved 'calpain-cathepsin cascade' from nematodes to primates. *Cell Calcium* **36,** 285–293 (2004).
95. Wendt, A., Thompson, V. F. & Goll, D. E. Interaction of calpastatin with calpain: A review. *Biol. Chem.* **385,** 465–472 (2004).
96. Tompa, P., Mucsi, Z., Orosz, G. & Friedrich, P. Calpastatin subdomains A and C are activators of calpain. *J. Biol. Chem.* **277,** 9022–9026 (2002).
97. Neuhof, C. Calpain system and its involvement in myocardial ischemia and reperfusion injury. *World J. Cardiol.* **6,** 638–652 (2014).
98. Hanna, R. A., Garcia-Diaz, B. E. & Davies, P. L. Calpastatin simultaneously binds four calpains with different kinetic constants. *FEBS Lett.* **581,** 2894–2898 (2007).
99. Betts, R., Weinsheimer, S., Blouse, G. E. & Anagli, J. Structural determinants of the calpain inhibitory activity of calpastatin peptide B27-WT. *J. Biol. Chem.*

**278,** 7800–7809 (2003).

100. Anagli, J. *et al.* A novel calpastatin-based inhibitor improves postischemic neurological recovery. *Biochem. Biophys. Res. Commun.* **385,** 94–99 (2009).
101. Low, K. E. *et al.* Rational Design of Calpain Inhibitors Based on Calpastatin Peptidomimetics. *J. Med. Chem.* **59,** 5403–5415 (2016).
102. Marsault, E. & Peterson, M. L. Macrocycles are great cycles: Applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem.* **54,** 1961–2004 (2011).
103. Kalash, L. *et al.* Structure-based design of allosteric calpain-1 inhibitors populating a novel bioactivity space. *Eur. J. Med. Chem.* **157,** 1264–1275 (2018).
104. Moldoveanu, T., Campbell, R. L., Cuerrier, D. & Davies, P. L. Crystal structures of calpain-E64 and -leupeptin inhibitor complexes reveal mobile loops gating the active site. *J. Mol. Biol.* **343,** 1313–1326 (2004).
105. Saito, K. I. & Nixon, R. A. Specificity of calcium-activated neutral proteinase (CANP) inhibitors for human μCANP and mCANP. *Neurochem. Res.* **18,** 231–233 (1993).
106. Breen, C. J., Raverdeau, M. & Voorheis, H. P. Development of a quantitative fluorescence-based ligand-binding assay. *Nat. Publ. Gr.* 1–9 (2016). at <http://dx.doi.org/10.1038/srep25769>
107. KURAMOCHI, H., NAKATA, H. & ISHII, S. Mechanism of Association of a Specific Aldehyde Inhibitor, Leupeptin, with Bovine Trypsin. *J. Biochem.* **86,** 1403–1410 (1979).
108. Atta-ur-Rahman. in *Bioactive Natural Products (Part N)* (ed. Atta-ur-Rahman) **13,** 180–187 (Elsevier, 2016).
109. Turk, V. *et al.* Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochim. Biophys. Acta - Proteins Proteomics* **1824,** 68–88 (2012).
110. Donkor, I. O. An updated patent review of calpain inhibitors (2012 - 2014). *Expert Opin. Ther. Pat.* **25,** 17–31 (2015).
111. Pencillide, a nonpeptide calpain inhibitor, produced by Penicillium sp. F60760. *J. Microbiol. Biotechnol.* **8,** 188–190, April (1998).
112. Graybill, T. L. *et al.* Inhibition of human erythrocyte calpain I by novel quinolinecarboxamides. *Bioorganic Med. Chem. Lett.* **5,** 387–392 (1995).
113. Pencillide, a nonpeptide calpain inhibitor, produced by Penicillium sp. F60760. *J. Microbiol. Biotechnol.* **8,** 188–190, April (1998).

114. Montero, A., Mann, E., Chana, A. & Herradon, B. Peptide-biphenyl hybrids as calpain inhibitors. *Chem. Biodivers.* **1,** 442–457 (2004).

115. Pal, G. P., Elce, J. S. & Jia, Z. Dissociation and Aggregation of Calpain in the Presence of Calcium. *J. Biol. Chem.* **276,** 47233–47238 (2001).

116. Trinchese, F. *et al.* Inhibition of calpains improves memory and synaptic transmission in a mouse model of Alzheimer disease. *J. Clin. Invest.* **118,** 2796–2807 (2008).

117. Wang, K. K. *et al.* An alpha-mercaptoacrylic acid derivative is a selective nonpeptide cell-permeable calpain inhibitor and is neuroprotective. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6687–92 (1996).

118. Van Den Bosch, L. *et al.* An α-mercaptoacrylic acid derivative (PD150606) inhibits selective motor neuron death via inhibition of kainate-induced Ca2+ influx and not via calpain inhibition. *Neuropharmacology* **42,** 706–713 (2002).

119. Gilli, P. *et al.* Enthalpy-Entropy Compensation in Drug-Receptor Binding. *J. Phys. Chem.* **98,** 1515–1518 (1994).

120. Ali, M. A. M., Stepanko, A., Fan, X., Holt, A. & Schulz, R. Calpain inhibitors exhibit matrix metalloproteinase-2 inhibitory activity. *Biochem. Biophys. Res. Commun.* **423,** 1–5 (2012).

121. Hu, J., Van den Steen, P. E., Sang, Q.-X. A. & Opdenakker, G. Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases. *Nat. Rev. Drug Discov.* **6,** 480–498 (2007).

122. Ali, M. A. M., Stepanko, A., Fan, X., Holt, A. & Schulz, R. Calpain inhibitors exhibit matrix metalloproteinase-2 inhibitory activity. *Biochem. Biophys. Res. Commun.* **423,** 1–5 (2012).

123. Arai, K. *et al.* Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue. *Angew. Chemie - Int. Ed.* **56,** 5522–5526 (2017).

124. Beld, J., Woycechowsky, K. J., Hilvert, D. & Hci, F. Selenoglutathione : Efficient Oxidative Protein Folding by a Diselenide †. 5382–5390 (2007).

125. Back, T. G. & Codding, P. W. Studies of the dihedral angle of a crowded diselenide by X-ray crystallography and ultraviolet spectroscopy. *Can. J. Chem.* **61,** 2749–2752 (1983).

126. Houk, J. & Whitesides, G. M. Structure-Reactivity Relations for Thiol-Disulfide. **II,** 6825–6836 (1987).

127. Van Wart, H. E., Lewis, A., Scheraga, H. A. & Saeva, F. D. Disulfide Bond Dihedral Angles from Raman Spectroscopy. *Proc. Natl. Acad. Sci.* **70,** 2619–2623 (1973).

128. Adams, S. E. *et al.* Conformationally restricted calpain inhibitors. *Chem. Sci.* **6,** 6865–6871 (2015).

129. Adams, S. E., Parr, C., Miller, D. J., Allemann, R. K. & Hallett, M. B. Potent inhibition of Ca2+-dependent activation of calpain-1 by novel mercaptoacrylates. *Med. chem Commun* **3,** 566–570 (2012).

130. Tomai, T. & Mai, L. P. Rhodanine as a Privileged Scaffold in Drug Discovery. *Curr. Med. Chem.* **16,** 1596–1629 (2009).

131. Li, J. J. in *Name Reactions: A Collection of Detailed Mechanisms and Synthetic Applications Fifth Edition* 615–616 (Springer International Publishing, 2014). doi:10.1007/978-3-319-03979-4_279

132. Smith, M. B. & March, J. *March's advanced organic chemistry*. *Wiley* (2007).

133. Raja, T. K., Ananthapadmanabhan, S., Gopalakrishnan, R. & Vimala, K. M. Synthesis of some styryl dyes from 3-ethyl-2-thio-4-selenazolidone. *Current Science* **57,** 795–796 (1988).

134. Sneddon, J., Masuram, S. & Richert, J. C. Gas chromatography-mass spectrometry-basic principles, instrumentation and selected applications for detection of organic compounds. *Anal. Lett.* **40,** 1003–1012 (2007).

135. Jones, G. in *Organic Reactions* (John Wiley & Sons, Inc., 2004). doi:10.1002/0471264180.or015.02

136. Giles, N. M. *et al.* Metal and Redox Modulation of Cysteine Protein Function. **10,** 677–693 (2003).

137. Dickinson, D. A. & Forman, H. J. Cellular glutathione and thiols metabolism. *Biochem. Pharmacol.* **64,** 1019–1026 (2002).

138. Owen, J. B. & Butterfield, D. A. in *Protein Misfolding and Cellular Stress in Disease and Aging: Concepts and Protocols* (eds. Bross, P. & Gregersen, N.) 269–277 (Humana Press, 2010). doi:10.1007/978-1-60761-756-3_18

139. Kosower, N. S. & Kosower, E. M. The glutathione status of cells. *Int. Rev. Cytol.* **54,** 109–160 (1978).

140. Lukesh, J. C., Palte, M. J. & Raines, R. T. A Potent, Versatile Disulfide-Reducing Agent from Aspartic Acid. 4057–4059 (2012).

141. Cline, D. J. *et al.* New Water-Soluble Phosphines as Reductants of Peptide and Protein Disulfide Bonds : Reactivity and Membrane Permeability. **43,** 15195–15203 (2004).

142. Kumar, P., Chiku, T., Carvan, M. J. & Sem, D. S. Fluorescence-based detection of thiols in vitro and in vivo using dithiol probes. **352,** 265–273 (2006).

143. Burns, J. A., Butler, J. C., Moran, J. & Whitesides, G. M. Selective Reduction of Disulfides by Tris(2-carboxyethy1)phosphine. **02138,** 2648–2650 (1991).

144. Pullela, P. ., Chiku, T., Carvan III, M. & Sem, D. S. Fluorescence-based detection of thiols in vitro and in vivo using dithiol probes. *Anal. Biochem.* **352,** 265–273 (2006).

145. Mirković, B., Sosič, I., Gobec, S. & Kos, J. Redox-Based Inactivation of Cysteine Cathepsins by Compounds Containing the 4-Aminophenol Moiety. *PLoS One* **6,** 27197 (2011).

146. Gunther, W. H. H. The Reduction of Diselenides with Dithiothreitol. *Methods Selenium Chem.* **32,** 3931–3933 (1967).

147. Steinmann, D., Nauser, T. & Koppenol, W. H. Selenium and Sulfur in Exchange Reactions: A Comparative Study. 6696–6699 (2010).

148. Dubey, R., Lee, H., Nam, D. H. & Lim, D. Mild generation of selenolate nucleophiles by thiol reduction of diselenides: Convenient syntheses of selenyl-substituted aryl aldehydes. *Tetrahedron Lett.* **52,** 6839–6842 (2011).

149. Ruszczy, A. *et al.* Investigation of biotransformation of selenium in plants using spectrometric methods. *Spectrochim. Acta Part B* **130,** 7–16 (2017).

150. Shimodaira, S., Asano, Y., Arai, K. & Iwaoka, M. Selenoglutathione Diselenide: Unique Redox Reactions in the GPx- Like Catalytic Cycle and Repairing of Disul fi de Bonds in Scrambled Protein. 5644–5653 (2017). doi:10.1021/acs.biochem.7b00751

151. Bachrach, S. M., Demoin, D. W., Luk, M. & Jr, J. V. M. Nucleophilic Attack at Selenium in Diselenides and Selenosulfides . A Computational Study. 4040–4046 (2004). doi:10.1021/jp037972o

152. Dubey, R., Lee, H., Nam, D. H. & Lim, D. Mild generation of selenolate nucleophiles by thiol reduction of diselenides: Convenient syntheses of selenyl-substituted aryl aldehydes. *Tetrahedron Lett.* **52,** 6839–6842 (2011).

153. Zhang, L. *et al.* Use of Recombinant Peptide as Substrate in FRET Based Protease Assays. *Blood* **106,** 3941–3941 (2005).

154. Dewitt, S. & Hallett, M. B. Cytosolic free Ca2+ changes and calpain activation are required for β integrin-accelerated phagocytosis by human neutrophils. *J. Cell Biol.* **159,** 181–189 (2002).

155. Lock, J. T., Parker, I. & Smith, I. F. A comparison of fluorescent Ca2+ indicators for imaging local Ca2+ signals in cultured cells. *Cell Calcium* **58,** 638–648 (2015).

156. Franco, S. J. Regulating cell migration: calpains make the cut. *J. Cell Sci.* **118,**

3829–3838 (2005).

157. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli : advances and challenges. **5,** 1–17 (2014).

158. Jia, B. & Jeon, C. O. High-throughput recombinant protein expression in Escherichia coli : current status and future perspectives. (2016).

159. Adams, S. E. PhD thesis. *PhD thesis* (2013). at <http://orca.cf.ac.uk/53271/1/2013Adamsphd.pdf>

160. Finally, A. Expression Using the T7 RNA Polymerase / Promoter System. *Curr. Protoc. Mol. Biol.* **16.2,** 1–11 (2000).

161. Sweet, C. R. in *E. coli Plasmid Vectors: Methods and Applications* (eds. Casali, N. & Preston, A.) 277–288 (Humana Press, 2003). doi:10.1385/1-59259-409-3:277

162. Elce, J. S., Hegadorn, C., Gauthier, S., Vince, J. W. & Davies, P. L. Recombinant calpain II: Improved expression systems and production of a C105A active-site mutant for crystallography. *Protein Eng. Des. Sel.* **8,** 843–848 (1995).

163. Manual, I. BL21-CodonPlus Competent Cells Instruction Manual. *Agil. Technol.* 1–16 (2015).

164. Tabka, M. G. *et al.* pET System Manual. *pET Syst. Man.* **78,** 1–16 (2012).

165. GE Healthcare. Size exclusion chromatography: Principles and Methods. *GE Heal. Handbooks* 139 (2012).

166. Hong, P., Koza, S., Bouvier, E. S. P. & Corporation, W. A review size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. *J. Liq. Chromatogr.* 2923–2950 (2012). doi:10.1080/10826076.2012.743724

167. Benvenuti, M. & Mangani, S. Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. **2,** 1633–1651 (2007).

168. GE Healthcare. Superdex 75 10/300 GL and Superdex 200 10/300 GL Instructions. *GE Heal. Handbooks* 1–2 (2002).

169. Kelly, S. M. & Price, N. C. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Curr. Protein Pept. Sci.* **1,** 349–384 (2000).

170. Hill, A. F., Barnham, K. J. & Bottomley, S. P. *Protein Folding, Misfolding, and Diseases*. (2011).

171. Stojanovic, N., Murphy, L. D. & Wagner, B. D. Fluorescence-based comparative binding studies of the supramolecular host properties of PAMAM dendrimers using anilinonaphthalene sulfonates: Unusual host-dependent

fluorescence titration behavior. *Sensors* **10,** 4053–4070 (2010).

172. Shvadchak, V. V, Kucherak, O., Kseniia, A., Dziuba, D. & Yushchenko, D. A. Biochimica et Biophysica Acta Environmentally sensitive probes for monitoring protein-membrane interactions at nanomolar concentrations. **1859,** 852–859 (2017).
173. Liu, Y. *et al.* Use of a Fluorescence Plate Reader for Measuring Kinetic Parameters with Inner Filter Effect Correction. *Anal. Biochem.* **335,** 331–335 (1999).
174. Lampe, J. N. & Atkins, W. M. Time-Resolved Fluorescence Studies of Heterotropic Ligand Binding to Cytochrome. *Biochemistry* **45,** 12204–12215 (2006).
175. Venien-bryan, C., Li, Z., Vuillard, L. & Boutin, J. A. Cryo-electron microscopy and X-ray crystallography : complementary approaches to structural biology and drug discovery. **73,** (2017).
176. Shoemaker, S. C. & Ando, N. X-rays in the Cryo-Electron Microscopy Era: Structural Biology's Dynamic Future. *Biochemistry* **57,** 277–285 (2018).
177. Haupteman, H. A. The phase problem of X-ray crystallography. *Reports Prog. Phys.* **54,** 1427 (1991).
178. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43,** 186–190 (2010).
179. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
180. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D* **53,** 240–255 (1997).
181. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D* **67,** 355–367 (2011).
182. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–2132 (2004).
183. Wilkinson, K. D. Quantitative Analysis of Protein – Protein Interactions. *Methods Mol. Biol.* **261,** 15–31 (2004).
184. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).
185. Lebedev, A. A. *et al.* JLigand: A graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68,** 431–440 (2012).
186. Krissinel, E. Stock-based detection of protein oligomeric states in jsPISA.

*Nucleic Acids Res.* **43,** W314–W319 (2015).

187. Sumbul, F., Acuner-Ozbabacan, S. E. & Haliloglu, T. Allosteric Dynamic Control of Binding. *Biophys. J.* **109,** 1190–1201 (2015).

188. Liljas, L., Lundahl, P. & Hjertén, S. Selective solubilization with tween 20 of proteins from water-extracted human erythrocyte membranes analysis by gel electrophoresis in dodecylsulfate and in tween 20. *Biochim. Biophys. Acta - Biomembr.* **352,** 327–337 (1974).

189. Ye, Q., Campbell, R. L. & Davies, P. L. Structures of human calpain-3 protease core with and without bound inhibitor reveal mechanisms of calpain activation. *J. Biol. Chem.* **293,** 4056–4070 (2018).

190. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden Gate Shuffling : A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS One* **4,** 1–9 (2009).

191. Kosobokova, E. N., Skrypnik, K. A. & Kosorukov, V. S. Overview of Fusion Tags for Recombinant Proteins. *Biokhimiya* **81,** 299–314 (2016).

192. Kapust, R. B. *et al.* Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **14,** 993–1000 (2001).

193. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling , prediction and analysis. *Nat. Protoc.* **10,** 845–858 (2015).

194. Lamb, A. L., Kappock, T. J. & Silvaggi, N. R. You are lost without a map: Navigating the sea of protein structures. *Biochim. Biophys. Acta* **1854,** 258–268 (2015).

195. Liebschner, D. *et al.* Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. Sect. D, Struct. Biol.* **73,** 148–157 (2017).

196. Liu, K. & Kokubo, H. Exploring the Stability of Ligand Binding Modes to Proteins by Molecular Dynamics Simulations: A Cross-docking Study. *J. Chem. Inf. Model.* **57,** 2514–2522 (2017).

197. Jasial, S. & Hu, Y. How Frequently Are Pan-Assay Interference Compounds Active? Large-Scale Analysis of Screening Data Reveals Diverse Activity Profiles, Low Global Hit Frequency, and Many Consistently Inactive Compounds. *J. Med. Chem.* **60,** 3879–3886 (2017).

198. Lagorce, D., Sperandio, O., Galons, H., Miteva, M. A. & Villoutreix, B. O. FAF-Drugs2 : Free ADME / tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics* **9,** 1–9 (2008).

199. Pollastri, M. P. Overview on the Rule of Five. *Curr. Protoc. Pharmacol.* **49,** 9.12.1-9.12.8 (2010).
200. Doak, B. C., Giordanetto, F. & Kihlberg, J. Oral Druggable Space beyond the Rule of 5 : Insights from Drugs and Clinical Candidates. *Chem. Biol.* **21,** 1115–1142 (2014).
201. Pascolutti, M. & Quinn, R. J. Natural products as lead structures: chemical transformations to create lead-like libraries. *Drug Discov. Today* **19,** 215–221 (2014).
202. Kalash, L. *et al.* Structure-based design of allosteric calpain-1 inhibitors populating a novel bioactivity space. *Eur. J. Med. Chem.* **157,** 1264–1275 (2018).
203. Busse, W. W. *et al.* Safety and efficacy of the prostaglandin D2 receptor antagonist AMG 853 in asthmatic patients. *J. Allergy Clin. Immunol.* **131,** 339–345 (2013).
204. Liu, J. *et al.* Supporting Information Discovery of AMG 853 , A CRTH2 and DP Dual Antagonist. *ACS Med. Chem. Lett.* **2,** 1–12 (2011).
205. Autmizguine, J. *et al.* Population Pharmacokinetics of Trimethoprim-Sulfamethoxazole in Infants and Children. *Antimicrob. Agents Chemother.* **62,** e01813-17 (2018).
206. Wright, J. M., Musini, V. M. & Gill, R. First-line drugs for hypertension. *Cochrane Database Syst. Rev.* (2018). doi:10.1002/14651858.CD001841.pub3
207. Ilardi, E. A., Vitaku, E. & Njardarson, J. T. Data-Mining for Sulfur and Fluorine: An Evaluation of Pharmaceuticals To Reveal Opportunities for Drug Design and Discovery. *J. Med. Chem.* **57,** 2832–2842 (2014).
208. Davis, T. L. *et al.* The Crystal Structures of Human Calpains 1 and 9 Imply Diverse Mechanisms of Action and Auto-inhibition. *J. Mol. Biol.* **366,** 216–229 (2007).
209. Hata, S., Kitamura, F. & Sorimachi, H. Efficient expression and purification of recombinant human μ-calpain using an Escherichia coli expression system. *Genes to Cells* **18,** 753–763 (2013).
210. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. Sect. D* **66,** 22–25 (2010).
211. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 1002–1011 (2006).

212. Cowtan, K. Fitting molecular fragments into electron density. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64,** 83–89 (2007).
213. Vangone, A., Spinelli, R., Scarano, V., Cavallo, L. & Oliva, R. COCOMAPS : a web application to analyze and visualize contacts at the interface of biomolecular complexes. **27,** 2915–2916 (2018).
214. Brennan, R. G. & Link, T. M. Hfq structure, function and ligand binding. *Curr. Opin. Microbiol.* **10,** 125–133 (2007).
215. Fortas, E. *et al.* New insight into the structure and function of Hfq Bioscience Reports. *Biosci. Rep.* **35,** 1–9 (2015).
216. Schulz, E. C. Structure of an Escherichia coli Hfq:RNA complex resolution at 0.97 A. **F70,** 1492–1497 (2014).
217. Wilusz, C. J. & Wilusz, J. Lsm proteins and Hfq Life at the 3 ’ end. **10,** 592–601 (2013).
218. Bukowska, A., Lendeckel, U., Bode-b, S. M. & Goette, A. Physiologic and Pathophysiologic Role of Calpain : Implications for the Occurrence of Atrial Fibrillation General Properties of the Calpain Family. **30,** 115–127 (2012).
219. Danial, M. & Postma, A. Disulfide conjugation chemistry: a mixed blessing for therapeutic drug delivery? *Ther. Deliv.* **8,** 359–362 (2017).
220. Hallett, M. B., Pettit, E. J. & Davies, E. V. Capacitative Ca2+ influx and a diffusible influx factor. *The Biochemical journal* **314 ( Pt 3,** 1054–1056 (1996).