Structural & Kinetic Studies of CaMK1D Interactions

by

Steffane McLennan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biochemistry
University of Alberta

ABSTRACT

Protein kinases are critical players in cell signaling processes and their activity leads to the phosphorylation of substrates, often regulating enzymatic activity. 518 protein kinases are encoded in the human genome, however, there are many features within the catalytic domain of these proteins that have been evolutionarily conserved. These similarities pose problems for the drug discovery of kinase targets as off-target effects are a threat with substrate mimetic inhibitors. Therefore, understanding the structure and regulation of these cell signaling proteins is key for the discovery of therapeutics that regulate enzymatic activity through novel binding space. This thesis focuses on CaMK1D, an understudied member of the calcium/calmodulin dependent protein kinase (CaMK) family.

CaMK1D is a putative driver of Triple Negative Breast Cancer (TNBC), a particularly aggressive breast cancer subtype with no targeted treatment option. Overexpression of CaMK1D in TNBC tumours leads to increased cell proliferation and migration, moreover, it is a validated TNBC target as the knockdown of CaMK1D reverses tumourigenesis. The work in this thesis shows that calmodulin (CaM) binds CaMK1D in a two-part mechanism wherein CaM has a final resting place on the C-terminal lobe of CaMK1D as shown with structural data from small angle x-ray scattering. Phosphorylation plays an important role in the activation of CaMK family members, and this work shows that the phosphorylation of the CaMK1D activation loop occurs post CaM binding to sustain activity. Lastly, the development of therapeutics to inhibit CaMK1D utilizing novel binding space is the overall goal of the group working on this target and this thesis explores Protein Thermal Shift assays as an inexpensive, rapid screening method for the initial steps in drug discovery.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Amino Acids

Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic acid
Cys, C	Cysteine
Gln, Q	Glutamine
Glu, E	Glutamic acid
Gly, G	Glycine
His, H	Histidine
Ile, I	Isoleucine
Lys, K	Lysine
Leu, L	Leucine
Met, M	Methionine
Phe, F	Phenylalanine
Pro, P	Proline
Ser, S	Serine
Thr, T	Threon[2,3]ine
Trp, W	Tryptophan

Tyr, Y	Tyrosine
Val, V	Valine

Protein Families

CaMK	Kinase family consisting majorly of Ca^{2+}/CaM regulated protein kinases.

Proteins, Domains, Motifs, & Genes

AID	Auto-inhibitory domain
A-loop	Activation Loop
CaM	Calmodulin
CaMK	Ca^{2+}/CaM dependent protein kinase
CaMK1A	Ca^{2+}/CaM dependent protein kinase α
CaMK1B	Ca^{2+}/CaM dependent protein kinase β
CaMK1D	Ca^{2+}/CaM dependent protein kinase δ
CaMK1G	Ca^{2+}/CaM dependent protein kinase γ
CaMKK	Ca^{2+}/CaM dependent protein kinase kinase
CBD	CaM binding domain
C-lobe	C-terminal lobe
CREB	cAMP response binding protein 1
N-lobe	N-terminal lobe

Metabolites and Reagents

ADP	Adensoine 5'diphosphate
AMP	Adenosine 5'-monophosphate
AMPPCP	β,γ-methyleneadenosine 5'-triphosphate
ATP	Adenosine 5'-triphosphate
DMSO	Dimethyl sulfoxide
DTT	1,4-Dithiothreitol
GSK-3 XIII	Glycogen synthase kinase 3 XIII inhibitor (5-methyl-1H-pyrazol-3-yl)-(2phenylquinazolin-4-yl)amine
^2H_2O	Deuterium oxide
HEPES	4-(2-hydroxethyl)piperazine-1-ethanesulfonic acid
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
LB	Luria broth
M9	M9 Minimal Media
SOC	Super Optimized Broth + Glucose
TB	Terrific Broth
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
Tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol

Methods

HSQC   Heteronuclear single quantum coherence

OD_{600} Optical Density at 600nm

PTS   Protein thermal shift

rtPCR   Real-time polymerase chain reaction

SAXS   Small angle x-ray scattering

SDS PAGE   Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Chapter One

Introduction

CHAPTER ONE – INTRODUCTION

1.1 An Introduction to Protein Phosphorylation

Reversible phosphorylation has emerged as the most common post-translational modification which acts as a major regulator of enzymatic activity. The regulation of protein phosphorylation is carried out by three main players: a kinase, a phosphatase, and partners such as calmodulin acting on a substrate protein. The act of protein phosphorylation is carried out by protein kinases and can be described as the movement of the gamma phosphate group of adenosine triphosphate (ATP) to a hydroxyl group of an amino acid residue (shown in Fig. 1.1). Phosphorylation occurs on exposed Ser, Thr, and less commonly, Tyr residues of the substrate protein. The reverse process, or the removal of a phosphate group, is performed by protein phosphatases which act with less specificity than protein kinases. Currently, 518 human protein kinases have been identified that specifically act on diverse substrates through unique molecular architecture of both the protein binding surfaces and active sites[1]. In contrast, approximately 200 protein phosphatases are encoded by the human genome and act with a higher level of promiscuity[2].

1.1.1 Then and now: the road to understanding protein phosphorylation

The importance of protein phosphorylation became apparent in 1955 with the discovery of two distinct forms of glycogen regulatory protein, protein phosphorylase: active and inactive, or respectively, *a* and *b*[3]. It was subsequently determined that the covalent attachment of a phosphate was necessary for the conversion from inactive to the active state[4] and that a multi-step, phosphorylation cascade initiated by Protein Kinase A (PKA) resulted in the phosphorylation and consequent activation of phosphorylase[5]. This work on phosphorylase built the foundation for understanding the reciprocity of phosphorylation-dephosphorylation in enzymatic regulation[6].

It was initially thought that this phosphoryl transfer phenomenon was unique to the regulation of

glycogenolysis, however, it was soon realized that phosphorylation is a ubiquitous covalent

modification altering enzymatic activity[7]. Protein phosphorylation cascades are now understood

to be sophisticated networks of signaling macromolecules that ultimately lead to the regulation of
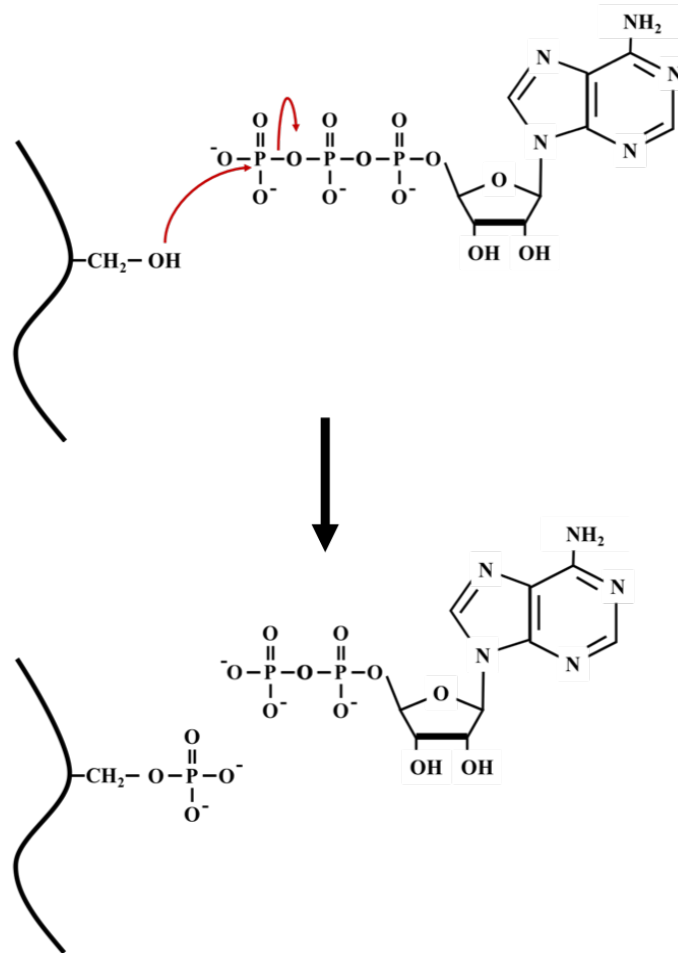
Figure 1.1 A general overview of the phosphorylation mechanism
A lone pair of electrons from a hydroxyl containing residue in the protein (Ser is shown but Thr
or Tyr could be substituted) initiates a nucleophilic attack on the gamma-phosphate of an ATP
molecule. The phosphate is released from the ATP molecule, breaking the phosphoanhydride bond
and yielding pSer and an ADP molecule.

a wide range of cellular functions such as metabolism, gene transcription, apoptosis, and cellular proliferation. The complexity and potential for amplification of these signalling cascades is a factor that renders them vulnerable; a multitude of problems can arise when the activity of protein kinases or phosphatases is altered by either a genetic mutation or an abnormal covalent modification[8,9].

1.1.2 Conserved protein kinase domain features

The first protein kinase structure to be solved was that of cyclic adenosine monophosphate-dependent protein kinase in 1991 (Human PKA)[10], a serine/threonine specific kinase, revealing a bi-lobed structure with a deep cleft separating the two sub-domains (Figure 1.2). The smaller N-domain consists of mainly beta-sheets and contains a beta-sheet necessary for ATP binding. The larger C-domain is majorly alpha-helical and interacts with peptide substrate. A small loop connects the two lobes, appropriately called the hinge and is the point where the kinase articulates to open and close around substrate molecules.

This structure revealed the orientation of conserved amino acid sequences throughout the protein[11]. Firstly, nucleotide binding proteins contain a conserved G-X-G-X-X-G motif that, in PKA and other kinases, is spatially near the catalytic site on the p-loop and aids in the binding of ATP (PKA: G50-G55). The p-loop is not only important for ATP binding but its structure also plays a role in substrate specificity, as shown in the example of the kinase MAP2K2[12]. An invariant Lys residue (PKA: K72) aids in the stabilization of the α and β phosphates of ATP. Furthermore, this Lys forms a salt bridge with a conserved Glu (PKA: E194) located on the αC helix. The last conserved motif that aids in ATP binding is the conserved DFG residues in the C-lobe (PKA: D148-G186): the
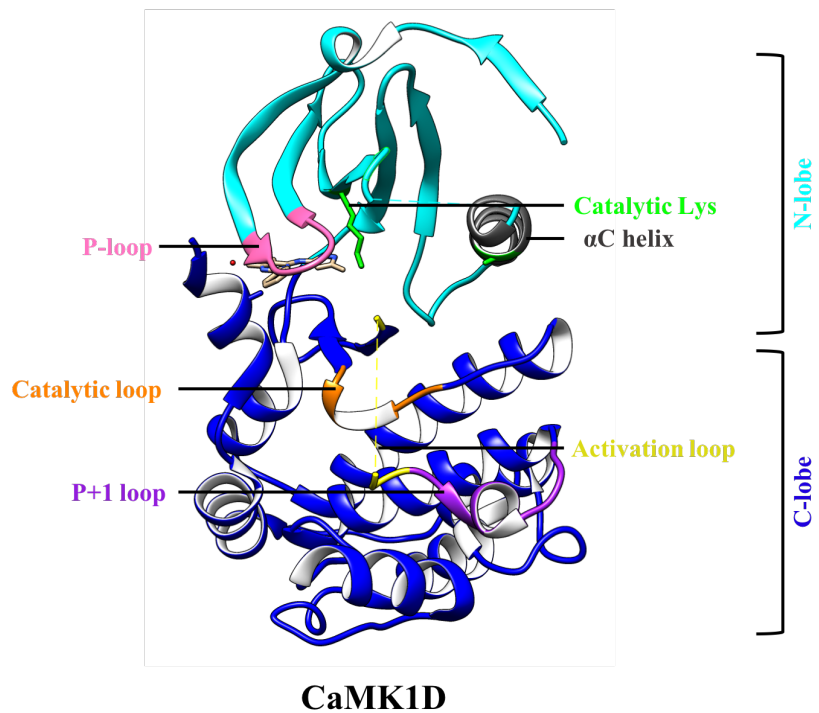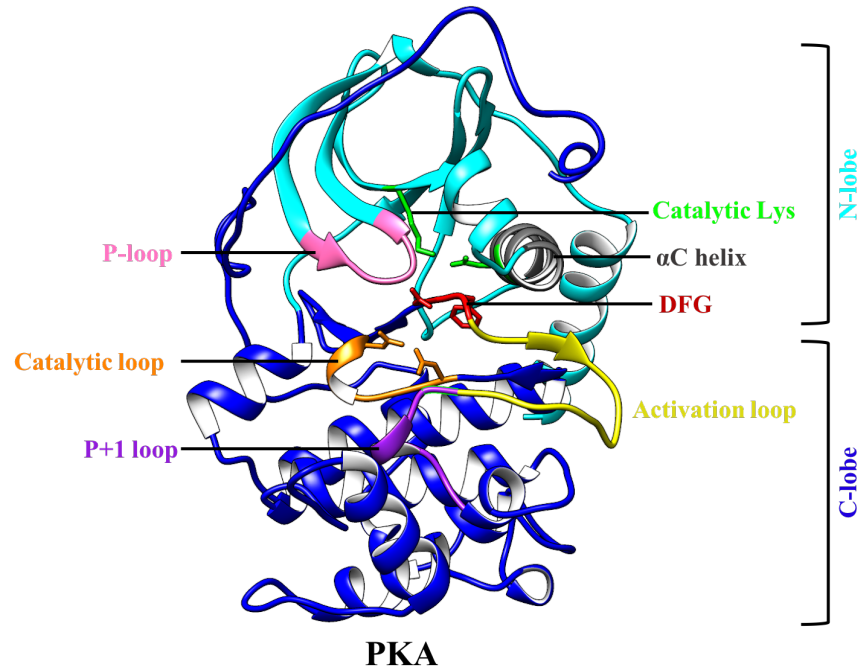
Figure 1.2 Crystal structure of PKA and CaMK1D with conserved kinase regions
The N-lobe (teal) consists of the nucleotide binding G-X-G-X-X-G motif in the p-loop (pink), and the catalytic Lys and Gln (green) which is located the αC helix (grey). The C-lobe (dark blue) contains the catalytic loop (orange), the Mg^{2+} coordinating DFG motif (red, not observed in CaMK1D), the activation loop (yellow), and the p+1 loop (purple) (PDB IDs: 2CPK & 2JC6).

negative charge on the Asp R-group coordinates the Mg^{2+} cofactor that is concurrently coordinated to and stabilizes the negative charges of the phosphates in the ATP molecule.

The catalytic site of protein kinases is evolutionarily conserved in a catalytic loop motif with a consensus sequence of D-X-X-X-X-N (PKA: D166-N171)[11]. The catalytic Asp residue acts a general base to deprotonate a hydroxyl group on the substrate peptide; a lone pair of electrons on this newly formed substrate oxyanion then performs a nucleophilic attack on the γ-phosphate of a substrate ATP molecule (Fig. 1.1). The activation loop (PKA: F187-G200), though known to be relatively disordered, is a site of phosphorylation that mediates the kinase catalytic activity. Lastly, the p+1 loop located immediately downstream of the activation loop lies against the surface of the kinase and confers substrate specificity along with the p-loop.

The structural features described above is more specific to serine/threonine kinases, however, protein-tyrosine kinases share similar sequence and 3-dimensional structure. They exhibit a bi-lobed structure much like serine/threonine kinases, but their specificity is changed by a different catalytic loop motif. Serine/threonine kinases have a well-conserved catalytic loop consensus sequence of D-L-K-P-E-N whereas Tyrosine kinases will display a catalytic loop composed of D-L-R-A-A-N which lacks the negative charge of the glutamate residue[11]. This confers specificity to the tyrosine substrate. However, it has been noted that under certain conditions, it is possible for serine/threonine kinases to phosphorylate substrate tyrosine residues[13].

1.1.3 Targeting protein kinases in the development of therapeutics

The prominence of protein phosphorylation in cellular function makes kinases attractive targets for the development of new drug molecules for the treatment of disease. In cancer cells, specifically breast and colorectal tumours, it was previously found that kinases show the highest rate of mutation[14] which is thought to contribute to the disease state. Kinases often act in cascades where sequential, serial phosphorylation of multiple kinases acts to amplify signals and creates diversity in the pathways. Identification of specific nodes in signaling pathways where the disease signal may originate or become amplified gives insight into where the inhibition of a specific kinase, or its activation, may minimize the disease signal. A textbook example of how kinases can be used as therapeutic targets is the development of imatinib, a potent inhibitor of Bcr-Abl tyrosine kinase that is overexpressed in Chronic Myeloid Leukemia[15,16]. The Bcr-Abl tyrosine kinase is structurally similar to the serine/threonine kinase target discussed in this thesis, demonstrating that specific kinase targeting is viable method of cancer treatment.

In the past, kinases were classified as "difficult to target" enzymes, largely due to the evolutionarily conserved active site architecture[17] that limits the selectivity and specificity of potential drug molecules. ATP mimetics are potent kinase inhibitors and useful for *in vitro* assays but poor therapeutics as they inhibit a wide variety of kinases and other nucleotide binding proteins causing toxicity in cells. The advancement of techniques used for both preliminary screening and obtaining specific molecular contact points between a drug molecule and its target has made the drug discovery pathway much more streamlined in recent years. Techniques such as virtual ligand screening[18] are now being used to discover novel binding pockets on target proteins to give insight into which drug fragment molecules may be good candidates for initial experimental screening.

Structural biology techniques such as Nuclear Magnetic Resonance (NMR) and X-Ray Crystallography (XRC) can pinpoint the molecular contacts between a target protein and its ligand, ensuring that novel binding space is utilized. This atomic resolution data enables scientists to hone in on which molecular contacts need further optimization and aids in the refinement of the final drug molecules. This thesis focuses on a novel target within the CaMK superfamily implicated in Triple Negative Breast Cancer.

1.2 Breast Cancer Biology

Cancer is a disease characterized by the loss of cell cycle control. Cancerous cells are characterized by the 6 hallmarks of cancer: sustaining proliferative signaling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis, and resisting cell death[19]. In breast cancer, the cells most commonly affected are the epithelial cells that line breast milk ducts which present as ductal carcinomas. According to a European epidemiologic study in 2012, breast cancer still remains the most common cancer subtype, identifying an incidence of 464,000 new cases in 2012 alone[20]. Great strides have been made in breast cancer treatments in the form of hormone and small molecule therapies, however, aggressive and advanced cases still lack targeted therapeutic options.

1.2.1 Breast Cancer Classification

The relative abundance of Human Epidermal Growth Factor Receptor 2 (HER2), the Estrogen Receptor (ER), the Progesterone Receptor (PR), and arguably the nuclear proliferative marker, Ki-67[21], can be used as markers to categorize breast cancers and to devise personalized treatment plans. The basic categories of breast cancer subtypes are summarized in Table 1.1[22]. Breast cancer

Table 1.1 Breast Cancer Subtypes

INTRINSIC SUBTYPE	MARKERS PRESENT	GRADE	RELATIVE OUTCOME	PREVALENCE
LUMINAL A	[ER+][PR+] HER2- KI67-	1-2	Good	23.7%
LUMINAL B	[ER+][PR+] HER2-KI67+	2-3	Intermediate	38.8%
	[ER+][PR+] HER2+ KI67+	2-3	Poor	14%
HER2 OVEREXPRESSION	[ER-][PR-] HER2+	2-3	Poor	11.2%
BASAL-LIKE	[ER-][PR-] HER2- Basal-marker +	3	Poor	12.3%
NORMAL-LIKE	[ER+][PR+] HER2-KI67-	1-3	Intermediate	7.8%

Table modified from Dai *et al*[22].

is plagued by a high degree of heterogeneity, and each subtype listed in Table 1.1 can be further subdivided with numerous other molecular markers. However, for general treatment and classification purposes the table above is useful.

Hormone therapies such as Tamoxifen, an anti-estrogen drug, have proven effective in early stage Luminal A (ER+, PR+, HER2-, Ki67-), Luminal B (ER+, PR+, HER2+/-, Ki67+), and Normal-like (ER+, PR+, HER2-, Ki67-) subtypes as they display upregulated expression of the ER and PR hormone receptors. Luminal A cases generally have the better outcomes than Luminal B due to the absence of Ki67. Ki67 is a nuclear marker that denotes cellular proliferation as it is only produced when a cell is actively dividing. Luminal A tumour cells proliferate at a slower rate, and this slower progression gives physicians a longer time frame in which detection can be made for a lower grade tumour and treatments subsequently initiated. The normal-like subtype is similar to a

Luminal A profile, however, the expression levels of the hormone receptors and HER2 resemble that of normal, healthy breast tissue which complicates diagnosis and treatment options.

The HER2 overexpression subtype displays a hormone receptor negative profile, limiting specific treatment options to drug molecules such as Herceptin that targets HER2 directly. However, the outcomes for this subtype remain poor. The HER2 overexpression classification is made using a gene expression array which groups tumours that overexpress other genes in the HER2 amplicon, such as GRB7[23,24] and PGAP3[23], as HER2 overexpression. This leaves room for error where tumours that appear to overexpress HER2 in fact do not, and this misclassification could lead to ineffective treatment with Herceptin.

Basal-like breast cancers are negative for all common markers but show increased levels of basal markers such as keratins[24,25]. Basal-like tumours, which make up the largest group of the Triple Negative Breast Cancer (TNBC) tumours, are more aggressive[26] and target a younger, pre-menopausal population[27]. Currently the treatment options for TNBC tumours are general surgery, chemotherapy and palliative care in advanced cases as no targeted treatment options have successfully passed clinical trials to this point.

1.2.2 Tumour Progression

Ductal carcinomas, if localized, are generally treatable and yield positive outcomes, however, it is the transition from a contained to a metastatic/invasive state that creates tumours which are difficult to treat. Breast cancer tumour progression is defined in 4 states: epithelial hyperproliferation followed by *in situ*, invasive, and then metastatic carcinomas (Fig. 1.3). It has

been realized that the microenvironment surrounding the breast tumour and epigenetic factors are important in its progression to a metastatic carcinoma. Under normal physiological conditions, the breast duct epithelium is a well-defined, uniform layer of luminal and myoepithelial cells with an intact basement membrane. Genetic and epigenetic alternations[28] of the luminal epithelial cells cause the cells to assume a phenotype characterized by the 6 hallmarks of cancer described by Hanahan *et al*[19]. This leads to the development of a Ductal Carcinoma *in situ* (DCIS) where epigenetic alterations cause a change in the myoepithelial cell phenotype to overexpress the gene for chemokine *CXCL14*[29]. *CXCL14* acts as a paracrine factor by binding to receptors on nearby luminal epithelial cells increasing their proliferative, migratory, and invasive properties. Furthermore, the initial degradation of the basement membrane is observed at this stage as the myoepithelial cells upregulate the expression of proteins such as matrix metalloproteinases (MMPs) that act to degrade the extracellular matrix[29]. It is also at this point that the microenvironment surrounding the tumour dramatically; the number of lymphocytes, stromal fibroblasts, myofibroblasts, and endothelial cells increase. The increased presence of endothelial cells promotes angiogenesis. Cancer associated fibroblasts (CAFs), much like myoepithelial cells, act as paracrine factors through the secretion of factors including the chemokine *CXCL12*. *CXCL12* activates the G-protein coupled receptor (GPCR), C-X-C chemokine receptor type 4 (CXCR4) on luminal epithelial cells which controls downstream signaling involved in increasing intracellular Ca^{2+} levels and the activation of mitogen activated protein kinase (MAPK) which impacts cell proliferation, survival, and tumourigenesis. The transition from a DCIS to an invasive carcinoma is characterized by the disappearance of structured myoepithelial layer and basement membrane[30] which, at least in part, is due to the upregulation of MMPs in myoepithelial cells. The

complete disappearance of this basement membrane allows tumour cells to invade surrounding

tissues and eventually leads to distant metastasis.

1.2.3 Calcium Signaling in Breast Cancer

Ca^{2+} is a critical ionic signal in a variety of cellular processes including cell cycle control,

apoptosis, gene transcription, and motility. Cytosolic Ca^{2+} is maintained at a low concentration of

~10^{-7} M in quiescent cells by Ca^{2+} ATPases such as Sarco/Endoplastmic Reticulum Ca^{2+} ATPase

(SERCA) which pumps Ca^{2+} against its concentration gradient with the aid of ATP hydrolysis to

store Ca^{2+} in either the sarco- or endoplasmic reticulum. Even in non-excitable cells such as breast

tissue, Ca^{2+} concentrations in micro domains near channels transporting extracellular or
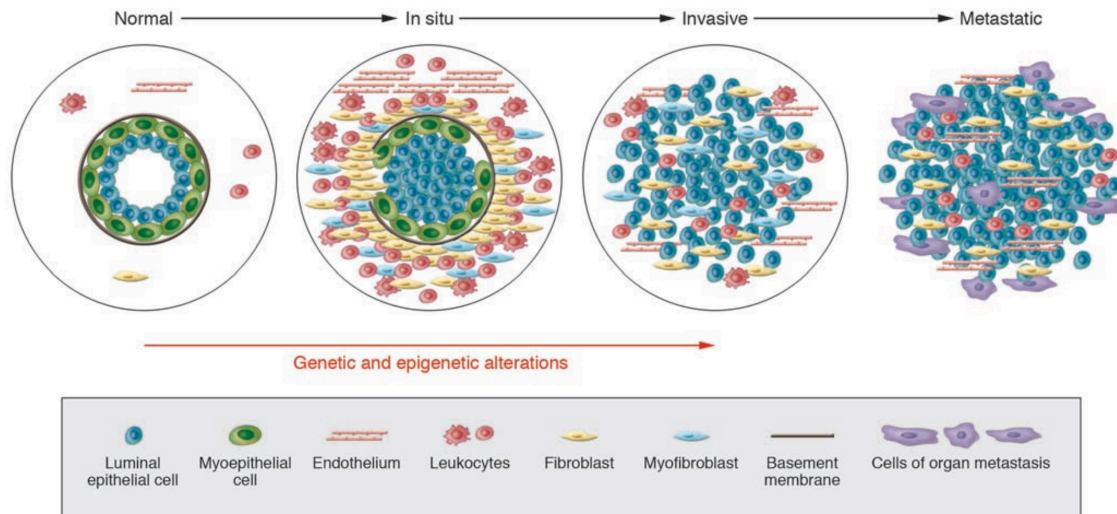
Figure 1.3 Schematic of Breast Cancer Progression
The breast cancer tumour progression is characterized by 4 stages: epithelial hyperproliferation
followed by *in situ,* invasive, and metastatic carcinomas. *In situ* carcinomas are characterized by
the decrease in myoepithelial cells and partial degradation of the basement membrane. There is an
increase in endothelial, lymphocytes, stromal fibroblasts, and myofibroblasts in the tumour
microenvironment at this stage. The transition to an invasive carcinoma shows a complete
degradation of the basement membrane which then allows the tumour cells to invade surround
tissue and metastasize. Figure from Polyak (2007)[31].

intracellularly stored Ca^{2+} can increase to $\sim 10^{-3}$ or $\sim 10^{-5}$ M, respectively, upon stimulation by various signals. One such signal that is the most common Ca^{2+} regulatory in non-excitable cells, inositol triphosphate (IP_3), which binds to and activates Ca^{2+} channels known as inositol triphosphate receptors (IP_3Rs), inducing a calcium influx[32]. Specific oscillations of these Ca^{2+} concentration spikes, which have are also described as waves, have been shown to provide signals that are necessary for cellular differentiation and proliferation[33–35]. Downstream effectors in Ca^{2+} signaling include calmodulin (CaM), Ca^{2+}/calmodulin-dependent protein kinases (CaMKs), calpain, and more specific to immune function are nuclear factor of activated T-cells (NFAT) and nuclear factor kappa-light chain-enhancer of activated B-cells (NF-kB).

Calcium is intimately linked to breast tissue as calcium is a major component of milk, and its signalling and transport has been shown to be altered in breast cancer cells. As previously discussed, the tumour microenvironment in breast cancer is critical for its progression to a metastatic state. The GPCR, CXCR4, which is stimulated by the *CXCL12* that is overexpressed by CAFs in DCISs has a direct impact on the calcium signaling pathway within cancer cells[36]. Specifically, upon its activation it activates Phospholipase C (PLC) which converts phosphatidylinositol 4,5-bisphosphate (PIP_2) into IP_3 to activate the signaling mechanism described in the previous paragraph. This effectively increases the cytoplasmic calcium concentration. There is also some evidence to suggest that calcium signaling is altered from the downregulation of certain Ca^{2+} channels on the cell surface, and subsequent compensatory upregulation of other calcium channels[37]. Ultimately, it is clear that calcium signaling controls important regulatory processes within breast cells and these processes are altered in the cancerous phenotype.

1.3 The Ca²⁺/Calmodulin-dependent Protein Kinase Family

The Ca^{2+}/Calmodulin-dependent protein kinase (CaMK) family of kinases is synchronized with calcium signaling through the ubiquitous calcium sensor protein, calmodulin (CaM) which, once loaded with calcium, binds to and activates members of the CaMK family. As illustrated in Figure 1.3, kinases in the CaMK family display similar domain organization. With the exception of CaMKIII, the catalytic kinase domain is the most N-terminal domain. Downstream of the catalytic
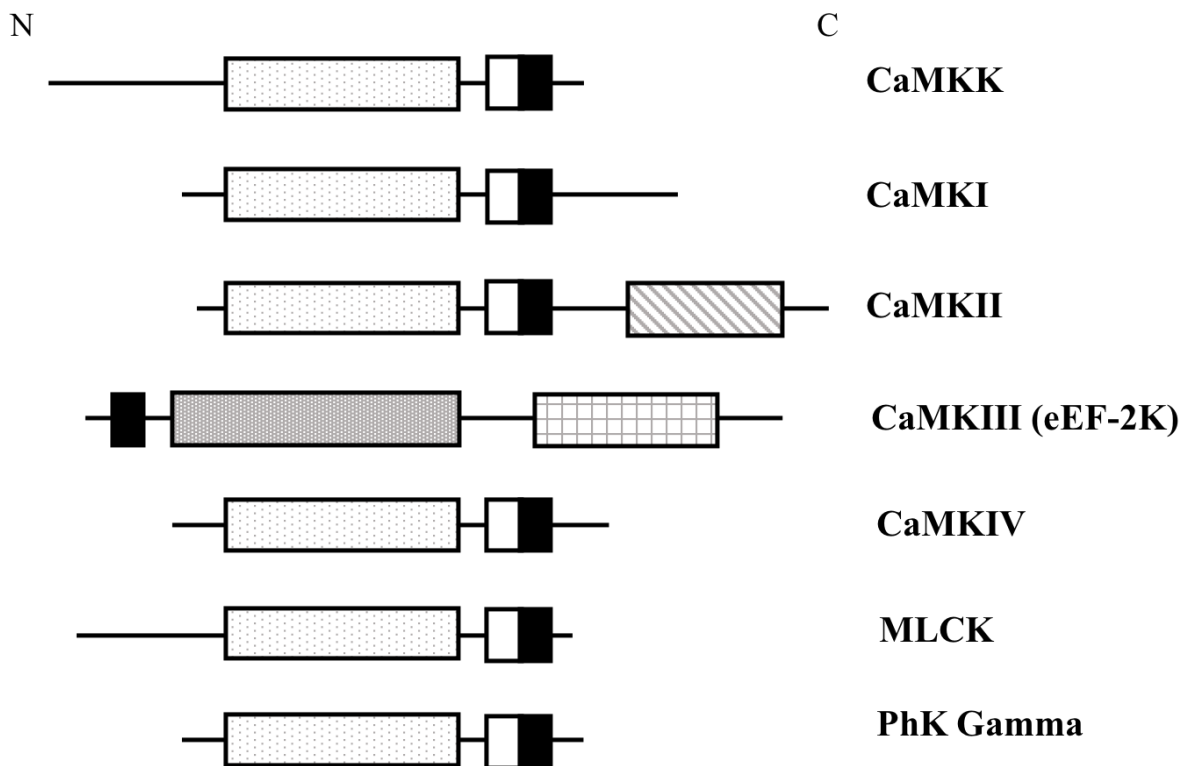
Figure 1.3 General domain organization of Ca²⁺/Calmodulin-dependent protein kinases
The conserved domains are shown by black dots (catalytic domains), white (auto-inhibitory domain, AID), and black (calmodulin binding domain, CBD). CaMKII contains an additional association domain used for oligomerization shown by grey/white stripes. CaMKIII has a unique catalytic domain (grey) with a C-terminal eEF-2K recognition domain (white boxes). Adapted from Swulius *et al*[38].

domain are the overlapping auto-inhibitory domain (AID) and calmodulin binding domain (CBD). Without CaM, the AID lays along the active site surface, leaving the kinase auto-inhibited, however, when CaM binds the CBD it dislodges the AID from the active site of the kinase rendering it active and able to phosphorylate Ser and Thr residues on substrate proteins[39,40].

CaMKs have been shown to undergo phosphorylation as a secondary method of enzymatic regulation and this has been well characterized in CaMKII isoforms. CaMKII forms a dodecamer via its association domain where other CaMKs act as monomers, however, there are conserved phosphorylation sites throughout the CaMK family which provides rationale to conclude that the CaMKII model is relevant to other CaMKs. Phosphorylation of T286/287 on CaMKII, which is located in the AID, is an activating phosphorylation site that acts as a "CaM lock" and ensures that the kinase remains in its active form. In contrast, the phosphorylation of T305/306/307 in the CBD of CaMKII blocks CaM binding which consequently does not allow CaM to dislodge the AID from the surface of the kinase to expose the active site . CaMKII and CaMKIV exhibit auto-phosphorylation after Ca^{2+}/CaM binding, which permits autonomous activation and inhibition. Other isoforms are phosphorylated by the upstream Ca^{2+}/CaM bound CaMKK.

1.3.1 CaMK1D in Disease

Ca^{2+}/Calmodulin-dependent protein kinase 1 δ (CaMK1D) is a 37kDa Ser/Thr kinase of the CaMK superfamily that has been identified as a putative driver of Triple Negative Breast Cancer (TNBC)[42,43]. TNBC is a particularly aggressive breast cancer subtype for which there are currently no targeted treatment options. TNBC tumours lack the cellular marks such as Estrogen Receptor (ER), Progesterone Receptor (PR) and Human Epidermal Growth Factor Receptor 2 (HER2) that

are targeted by proven hormone and small molecule therapies. For this reason, the discovery of a viable, druggable target for targeted treatment options for TNBC patients is of utmost importance. CaMK1D is overexpressed in 80% of TNBC tumors[26] and cells that display this overexpression show increased cell proliferation and migration[42]. Though an increase in the abundance of a protein does not necessarily lead to a disease state, it has been shown that TNBC tumour cells are susceptible to the knockdown of CaMK1D by siRNA[44]. Furthermore, preliminary studies of CaMK1D inhibition via small molecules show promising results, which further validate this kinase as a TNBC target suitable for drug discovery. The link between CaMK1D and cancer is becoming a rather well-documented story, and a recent publication indicates that CaMK1D knockdown by endogenous microRNAs may play a role in promoting angiogenesis in late stage lung cancers[45].

The CaMK1D-cancer link is the focus of this thesis work, however, CaMK1D was first identified as a molecular player in Type II Diabetes (T2D). Genetic variants in the CaMK1D gene, *CAMK1D*, found in both European and Asian populations, display increased susceptibility to T2D[46,47]. On a molecular level, CaMK1D inhibition via siRNA has been shown to decrease the expression of Phosphoenolpyruvate Carboxykinase 1 (PCK1) in hepatocytes[44]. PCK1 is a key regulator of gluconeogenesis, converting oxaloacetate in the Kreb's Cycle into the glycolysis intermediate, phosphoenolpyruvate, to be utilized in the formation of glucose for storage as glycogen. With this knowledge, it is likely that the dysregulation of CaMK1D expression due to mutations in the *CAMK1D* gene has a direct impact on the ability of hepatocytes to regulate glucose levels and glycogen formation, contributing to the T2D disease state.

Finally, there are links between Late Onset Alzheimer's Disease (LOAD) and CaMK1D. Risk-factors have been identified on chromosome 10, including the CaMK1D locus (10p13)[48]. Furthermore, a Genome-Wide Association Study (GWAS) of >200,000 individuals revealed Single Nucleotide Polymorphisms on chromosome 10 located near the CaMK1D locus were associated with clinically diagnosed Alzheimer's Disease (AD)[49]. On the protein level, it has been shown via Western Blot analysis of brain tissue of patients with AD that CaMK1D is differentially cleaved in these cells[50]. It has been postulated that calpain, a Ca^{2+} dependent protease, may cleave CaMK1D and cause aberrant activity, altering the phosphorylation states of its substrates. CaMKs have been shown to phosphorylate Tau protein, which is the protein which misfolds and causes neurofibrillary tangles in AD[51]. The alteration of the Tau phosphorylation state would have implications on fibrillation kinetics, which is positively correlated to the development of the disease state[52].

1.3.2 CaMK1D: A Molecular Level View

CaMK1D is one of four CaMKI isoforms (α, β, γ, δ) (Fig. 1.4) which are expressed from four distinct genes. All four of these isoforms are relatively similar in sequence and domain organization, with the exception of CaMK1G which contains a CAAX motif in its C-terminal to allow for targeting and interactions with the plasma membrane (Fig. 1.4). The splice variant of CaMK1D, known as CaMKI-like kinase (cKLiK), which is almost exclusively expressed in polymorphonuclear leukocytes (PMNs) and plays a role in neutrophilic differentiation of CD34(+) stem cells and chemo-attractant induced responses of human granulocytes[53,54]. In contrast to cKLiK, expression of CaMK1D from the aptly named *CaMK1D* gene is widespread: rtPCR experiments indicate expression in tissues including the liver, spleen, thymus, testes, ovaries, and

colon[55], however, expression is highest in the brain. CaMK1D exhibits many conserved kinase regions that are seen in PKA (Fig. 1.2).

As illustrated in Figures 1.3-1.5, CaMKI isoforms follow the general domain organization of CaMKs: a catalytic kinase domain closest to the N-terminal, followed by the regulatory domain consisting of the overlapping AID (residues 279-313) and CBD (residues 303-322). Like other CaMKs, CaMK1D activity is regulated by both CaM and phosphorylation. Ishikawa *et al* showed that CaMKK1 & CAMKK2 phosphorylate T180 to produce sustained activation, but only after Ca^{2+}/CaM is bound. This suggests that T180, which is found the CaMK1D activation loop (A-loop, residues 164-184), is not accessible to CaMKK until the AID is dislodged from the kinase
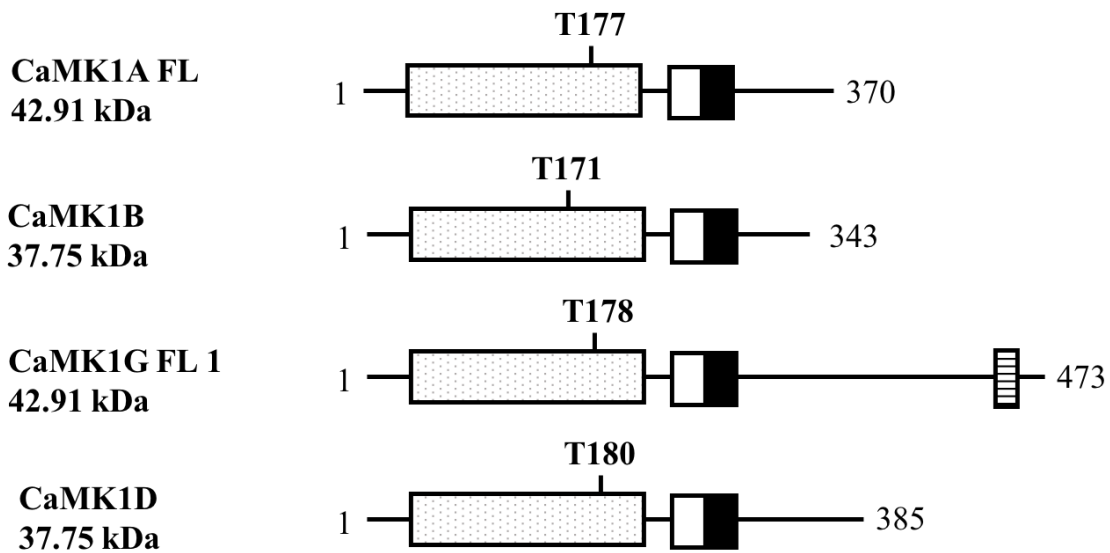
Figure 1.4 Overview of CaMK1 Isoforms
The four CaMK1 isoforms follow the general domain organization of CaMKs with the catalytic kinase domain (black polka dots) containing a phosphorylation site that contributes to sustained activity *in vivo*. The overlapping AID and CBD are shown in white and black, respectively. CaMK1G also contains a CAAX motif at its C-terminal (horizontal stripes) which plays a role in membrane targeting. Adapted from Swulius *et al*[56].

surface. It has not been established as of yet whether CaMK1D exhibits autonomous activation after Ca^{2+}/CaM binding through auto-phosphorylation at sites other than T180 (ie. a site homologous to T286/287 in CaMKII variants), but the C-terminal region of CaMK1D (residues 326-362, containing the CBD) is described as containing many potential phosphorylation sites by Uniprot, as it is Ser/Thr rich.

The biological role and signaling pathways of CaMK1 isoforms have not been studied extensively. However, in primary hippocampal neurons, CaM bound CaMK1D has been shown to translocate to the nucleus upon KCl depolarization and results in the phosphorylation of transcription factor CREB, activating pCREB-dependent gene transcription. A pCREB response and subsequent expression of the intermediate filament protein vimentin indicates that CaMK1D is likely a component of a signaling cascade that leads to cell proliferation and cytoskeletal remodeling[57]. More recently, CaMK1D was found to undergo "cross-talk" with eukaryotic initiation factor 5A (eIF5A) in trophoblast cells[58]: knockdown of eIF5A via siRNA decreased CaMK1D expression in HTR-8 cells and decreased their proliferative, migratory, and invasive properties. Furthermore, CaMK1s have been implicated in the cell cycle wherein it was shown that CaMK1 deficient cell lines display cell cycle arrest, where CaMKII deficient cell lines did not[59], suggesting that this class of CaMK plays a unique, pivotal role cell cycle progression. This supports the notion that CaMK1D is involved in cancer: a disease characterized by the loss of cell cycle control.

A

Catalytic Domain   Regulatory Domain

T180

1   385

Auto-Inhibitory Domain   CaM Binding Domain

B

CaM Binding Domain

Auto-Inhibitory Domain

Activation Loop
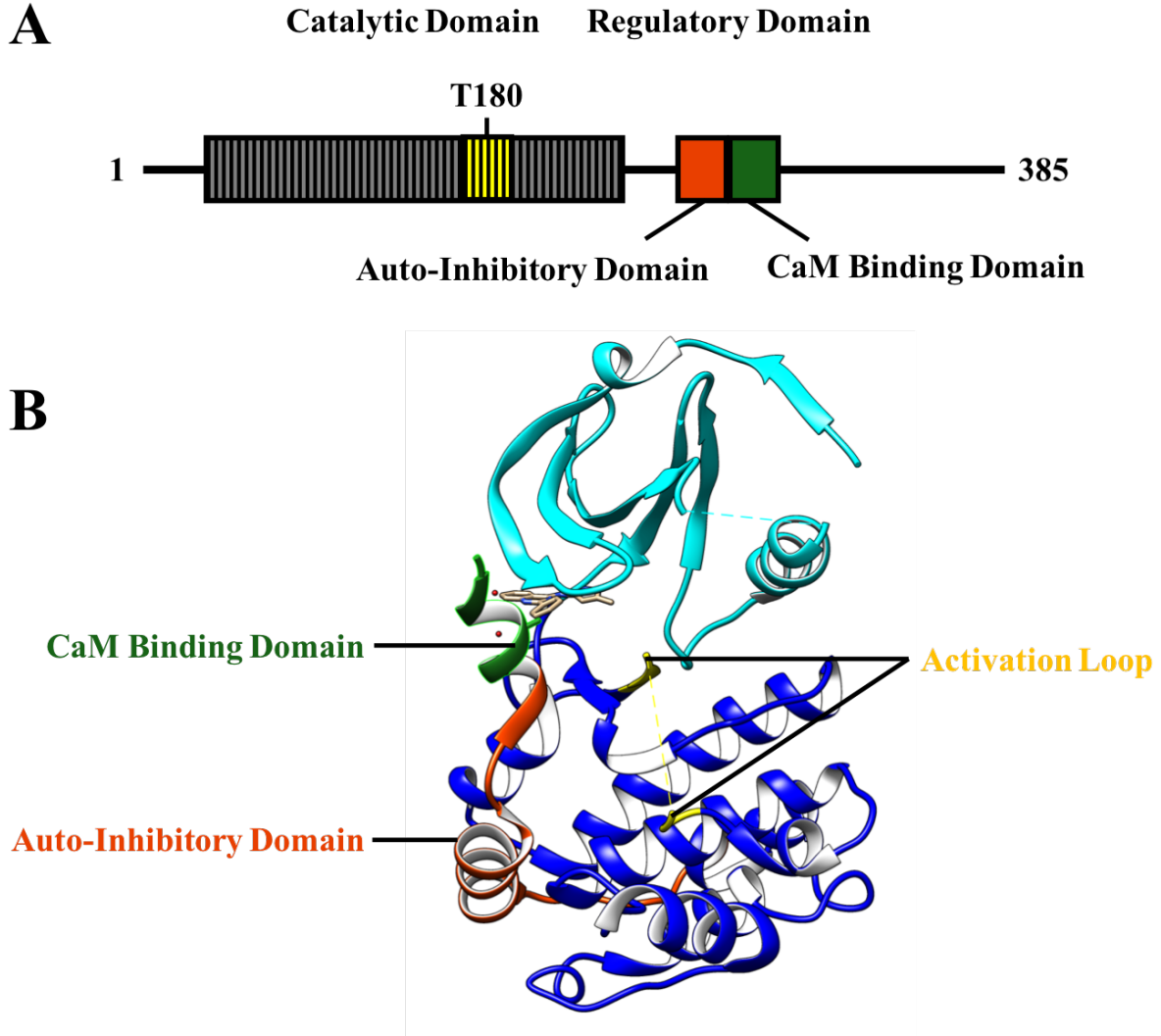
Figure 1.5 Key regulatory regions of CaMK1D

A) CaMK1D consists of two main domains: the catalytic kinase domain (vertical stripes) and the regulatory domain (solid red & green). The catalytic domain contains the kinase active site as well as the activation loop (yellow) which is phosphorylated by CaMKK after Ca^{2+}/CaM is bound. The regulatory domain consists of the auto-inhibitory domain (red) and the CaM binding domain (green). B) Crystal structure of CaMK1D (PDB ID: 2JC6)

1.3.3 Calmodulin: the master calcium sensor

Calmodulin (CaM) is a 17kDa calcium sensor protein that plays an essential role in Ca^{2+} signalling by acting as an activator of many signalling cascades, including those driven by CaMKs. Mutations that interfere with the calcium binding ability of this critical calcium sensor have been found in infants that present with recurrent cardiac arrest[60]. Interestingly, elevated blood CaM levels may act as a potential biomarker for Alzheimer's Disease [61].
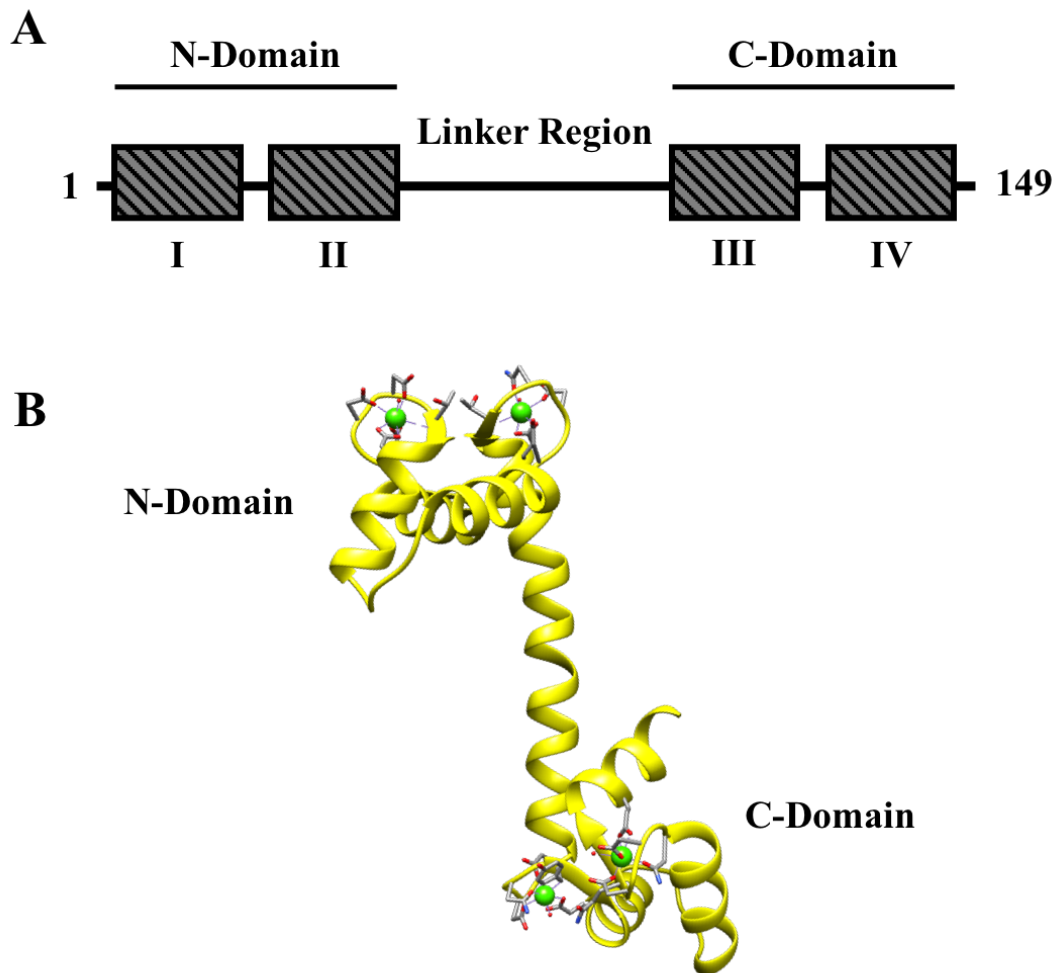
Figure 1.6 Domain architecture and structure of CaM
A) Calmodulin consists of a 149 amino acid chain containing two globular domains: the N-domain and the C-domain. Each globular domain is composed of 2 EF hand motifs (numbers I-IV, grey/black stripes) which bind cytosolic calcium. The domains are connected by the linker region which is flexible to facilitate substrate binding. B) Crystal structure (PDB ID: 3CLN) of calcium (green) bound CaM (yellow).

CaM structure is characterized by two globular domains, the N-domain and the C-domain, which are connected via a flexible linker, resembling a dumbbell (Figure 1.6). Each globular domain is composed of two helix-loop-helix motifs, commonly known as EF hands, each binding one calcium ion. It has been argued that the central linker adopts a helical conformation upon calcium binding, as shown in the crystal structure [62] (Figure 1.6B). However, NMR relaxation studies show that the linker domain remains flexible even upon calcium binding [63] which facilitates substrate binding as CaM wraps around its substrate. Calmodulin binds substrates through hydrophobic interactions with its globular domains and the IQ motif of CaM's ligand proteins (consensus sequence: IQXXXR(GX)XXX, where X represents any amino acid).

1.4 Biophysical Techniques

Biophysical techniques unveil details of macromolecular structure at an atomic level, giving insight into molecular function as structure often dictates function. The biophysical techniques used in this thesis work include Nuclear Magnetic Resonance (NMR) and Small Angle X-Ray Scattering (SAXS) to uncover molecular details of the Ca^{2+}-CaM/CaMK1D interaction.

1.4.1 Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) is a powerful structural technique that utilizes the nuclear spin of spin active atoms, including ^{15}N, ^{13}C and ^{1}H which display spin quantum numbers (I) of 1/2 which enables their detection in NMR experiments. Specific chemical shifts of these nuclei allow for the deduction of information about the solution structure, conformational changes, ligand binding and other dynamic properties of proteins uniformly isotopically labelled with the latter NMR active isotopes[64]. Once assignments are completed, Chemical Shift Perturbations (CSPs)

from the expected random coil value is used as an indicator of secondary protein structure as chemical shift is directly dependent upon the local chemical environment in which the nuclei resides. Furthermore, the Nuclear Overhauser Effect (NOEs), which utilizes the transfer of nuclear spin from one nuclei to another in close proximity by cross-relaxation[66], can give information about tertiary protein structure. Amino acids which are not close in primary structure may be in close proximity in the folded structure, therefore, NOEs give critical information in regard to how the secondary structures identified with CSPs fold in three-dimensional space.

For the purpose of this thesis work, NMR was specifically utilized to determine protein-protein interaction characteristics using Heteronuclear Single Quantum Coherence (HSQC) experiments. HSQC experiments are two-dimensional NMR experiments where chemical shifts are resolved on the ^{1}H and ^{15}N dimensions. For protein NMR, useful information is gathered from the correlation between the backbone amide nitrogen and its respective proton, resulting in one peak for each amino acid, with the exception of Pro as it does not contain an amide proton. In what is known as the fingerprint region, the region containing the amide peaks, mirror peaks from the side chains of Asn and Gln residues will appear but are easily distinguishable from the backbone amide peaks.

Even without backbone assignments, chemical shift perturbations in HSQC experiments can be used to determine characteristics of protein-protein interactions based on the type of exchange they display: slow, intermediate, or fast[67]. Since the introduction of a ligand will change the local chemical environment of the amino acids in the binding interface between the two binding partners, the chemical shift of these amino acids will change accordingly. As depicted in Figure 1.7, an interaction characterized by fast exchange kinetics will display a chemical shift that

increases with increasing ligand concentration until it reaches a final saturation point. Often, it is possible to calculate a relative binding affinity for binding partners such as these by plotting the relative change in chemical shift and identifying the concentration of ligand where 50 % of the binding sites are saturated. Slow exchange kinetics presents itself not as a gradual change in chemical shift, but as the disappearance of one peak representing the unbound state and the appearance of a second peak representing the bound state. Slow exchange kinetics generally characterizes a strong, non-transient interaction where the population of the half-bound state seen in fast exchange kinetics is too low to resolve as the labeled protein almost exclusively exists in either the bound, or unbound form. Interaction kinetics are more difficult to deduce from a slow exchange interaction; however, it is possible to calculate the change in area under either the bound
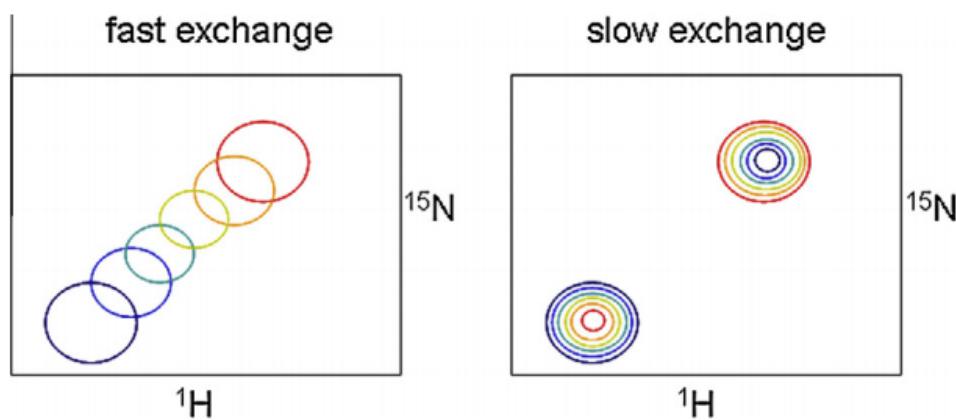
Figure 1.7 Exchange kinetics as observed by NMR
A schematic of peak shape and shift in NMR HSQC titrations of increasing ligand concentration. Fast exchange (left) is identified by a gradual peak shift from the unbound (purple) to a final saturated state (red). Fast exchange generally describes a transient interaction between partners as there is a high population of intermediary states throughout the binding interaction before saturation. Slow exchange (right) is identified by a gradual peak disappearance in the unbound form (purple) and the reappearance of a peak with a different chemical shift that represents the bound state (red). Slow exchange generally describes a tight binding interaction where the molecules of interest are almost exclusively in either a bound or unbound state, with minimal time in a transient, intermediate state. Figure from Williamson (2013)[67].

or unbound state peaks to calculate a rough estimate of affinity. An interaction characterized by intermediate exchange may display very small peaks representing the intermediately bound state, but the peak shape will be much narrower due to the decrease in population size as compared to fast exchange binding partners.

The greatest limitation of protein NMR is that only proteins of up to about 30 kDa can be easily resolved. This is both due to the sheer number of peaks that will appear on the HSQC spectrum, where proteins with >300 amino acids will show a high level of spectral crowding and overlap, but also due to the faster relaxation properties of the molecules with increasing size. Specific labeling methods, which are discussed in detail in section 4.1.2, are being employed to combat this problem. Crystallography, another high resolution biophysical method for structure determination, does not face the problem of size, however, structures solved by crystallography are static. In cellular conditions, a protein is in solution and therefore crystal structures are not always useful for visualizing the moving parts associated with protein function. NMR can give insight into dynamics, which makes it particularly useful even if a high-resolution crystal structure has already been solved, such as with the case of CaMK1D in this thesis work.

1.4.2 Small Angle X-Ray Scattering

Small Angle X-Ray Scattering (SAXS) is a relatively low resolution (10-20 Å range) structural technique that utilizes elastic scattering of x-rays to back calculate a molecular envelope of a macromolecule of interest[68]. An isotropic, monodispersed solution of the molecule of interest is subjected to a monochromatic x-ray beam and the scattering of x-rays due to the macromolecule is subtracted from background scattering to result in scattering solely from the molecule of interest.

The one-dimensional scatter data is most commonly analyzed with the ATSAS suite or with ScÅtter software (download available at bioisis.net). This analysis is further discussed in section 2.7.

The biggest limitation that SAXS faces is sample aggregation. Aggregates will display a different scatter pattern than the monodispersed form convoluting experimental data, specifically with the high angle scatter, or Guinier region, which contains the most useful, structural information. Therefore, SAXS samples must stable and non-aggregating for the duration of the experiment which is generally a couple of hours. However, it is now becoming more commonplace to couple SAXS to size exclusion chromatography wherein a size exclusion column separates macromolecules of varying sizes immediately prior to data collection. This lessens the impact of sample aggregation on data quality, granted that the population of native protein is still high enough to obtain quality data.

Despite that techniques such as NMR and X-Ray Crystallography provide high resolution structural information, sample preparation and data analysis are often the bottleneck in obtaining this data. SAXS samples can be run and data analyzed on the same day, making SAXS a valuable structural tool until higher resolution data can be analyzed. Furthermore, SAXS is a highly complementary technique wherein high-resolution crystallography data can be fitted into the experimental SAXS envelope. This is particularly useful to determine regions of flexibility as SAXS is a solution technique and the envelope will represent an average of conformations in solution. Additional volume in a SAXS model that does not appear in a static crystal structure may indicate a region of flexibility in the molecule of interest. Finally, as SAXS does not have any size

limitations, it can be used to determine how multi-protein complexes interact in three-dimensional space. The example of the nidogen-1 and laminin proteins is a textbook example of how SAXS data was utilized in conjunction with previously elucidated crystal structures to identify interaction surfaces between two proteins partners[69].

1.5 Thesis Objectives and Hypotheses

As an alleged driver of TNBC, CaMK1D is a potential target for drug discovery; previous data shows that TNBC tumourigenesis is reversed by CaMK1D knockdown[44], validating its legitimacy as a target. Discovery of drug molecules for CaMK1D is particularly challenging with the conserved CaMK architecture and therefore, understanding the structure of CaMK1D and its specific interactions with other molecules such as CaM, ATP, substrate, and small molecule inhibitors is critical. Preliminary investigation on small molecule inhibitors of CaMk1D has been initiated and relies on optimizing a rapid screening method for small molecules, which is one area of focus for this thesis.

A crystal structure of CaMK1D in complex with a small molecule kinase inhibitor was previously released in the protein data bank (2JC6), however, the kinase crystallized as a dimer which is not consistent with the current understanding that CaMK1D acts as a monomer. Furthermore, crystallization studies of CaMK1D in complex with its activator molecule, CaM, have not been successful thus far, emphasizing the need for additional structural investigation.

This leads to the 2 objectives of this thesis:

1. Determine if Protein Thermal Shift, NMR, and MST assays can be utilized for small molecule ligand screening, specifically in cross validation and mechanistic analysis.

2. Further the understanding of CaMK1D enzymatic regulation through uncovering the CaM/CaMK1D structure and the impact of phosphorylation events on complex formation and regulation of enzymatic activity.

My hypotheses are as follows:

1. Binding assays including Protein Thermal Shift and NMR can be used for identifying small molecule ligands of CaMK1D if the shifts in melting temperature (Tm) are reproducible and can be corroborated with other techniques such as NMR.

2. CaM binds to the C-terminal CBD of CaMK1D to activate its enzymatic activity.

3. T180 phosphorylation will not change complex formation of CaM/CaMK1D as T180 is supposedly sterically hindered from CaMKK phosphorylation until CaM binds to remove the AID from the kinase surface.

Chapter Two

Materials & Methods

CHAPTER TWO – MATERIALS & METHODS

All materials, reagents, and equipment purchased or used were from BioRad, Maybridge Chemicals, ThermoFischer Scientific, Sigma-Aldrich, Agilent, NanoTemper, ForteBio, TA Instruments, Bruker, and Invitrogen.

2.1 cDNA Clones

All cDNA clones used in this thesis work are summarized in Table 2.1. All constructs were expressed in a pNIC28-Bsa4 vector which results in the expression of proteins flanked by an N-terminal His_6 tag used for purification and fluorescent labeling purposes. TEV cleavage sites were located on the C-terminal end of the His_6 affinity tag to allow for final production of affinity tag-free protein when necessary. Mutagenesis for this project was completed by other Overduin Lab members.

Table 2.1. List of cDNA clones.

cDNA Clone	Optimized Expression Cell Line	Antibiotic Resistance	Expression Vector	Affinity Tag	TEV Cleavage Site (Y/N)
CaMK1D 1-333 WT	Rosetta 2 (DE3)	Kanamycin	pNIC-Bsa4	His_6	Y
CaMK1D 1-333 S179E/T180E	Rosetta 2 (DE3)	Kanamycin	pNIC-Bsa4	His_6	Y
CaMK1D 10-329 WT	Rosetta 2 (DE3)	Kanamycin	pNIC-Bsa4	His_6	Y
Calmodulin	T7 Express	Kanamycin	pNIC-Bsa4	His_6	Y

2.2 Transformations

cDNA stock concentrations were in the range of 100-200 ng/µL. 1-2 µL of stock cDNA was added to 50 µL of competent cells, specified in Table 2.1, and incubated on ice for 30 mins, followed by a 42 °C heat shock for 30 s. The cells were returned to incubate on ice for 2 mins. 100 µL of SOC broth was added and incubated at 37 °C for 1 hour. 100 µL of the cell suspension was spread on LB agar supplemented with 1 mgmL^{-1} Kanamycin and incubated for 16-18 hours at 37 °C. Plates were stored at 4 °C for up to 2 weeks.

2.3 Protein Expression and Purification

2.3.1 Protein Expression in LB

Initial expression tests were completed by inoculating 2 mL LB (supplemented with 1 mg/mL Kanamycin) with fresh transformants and incubating at 37 °C with shaking (150-200 rpm) overnight. 100 µL of this overnight culture was used to inoculate 2 mL LB supplemented with 1 mgml^{-1} Kanamycin. The culture was incubated with shaking at 150-200 rpm at 37 °C for 2 hours. Protein expression was induced with 0.5 mM IPTG and expression took place over 1 hour at 37 °C. Expression levels were visualized with SDS-PAGE.

Once an optimal cell line was identified, overnight cultures were set up by inoculating 25 mL LB, supplemented with 1 mgml^{-1} Kanamycin, with fresh transformants of the optimal cell line. These cultures were incubated overnight, shaking at 150-200 rpm on a benchtop shaker at 37 °C for 16 hours. 10 mL of overnight culture was used to inoculate 1 L of LB or TB Broth, which was subsequently incubated at 37 °C with shaking (150-200 rpm). The OD$_{600}$ was monitored using a

spectrophotometer until 0.6-0.8 for LB and 2-4 for TB was reached. IPTG was added at this point to a final concentration of 0.5 mM before dropping the temperature to 18 °C and incubating for another 16 hours with shaking at 150-200 rpm. Cell cultures were harvested by centrifugation at 10,000 x g for 15 minutes and either used immediately or stored at -20 °C.

2.3.2 Protein Expression in M9 Minimal Media for Protein Labelling

For proteins to be used in NMR experiments, the expression method was adapted to uniformly label proteins with ^{15}N and/or ^{13}C using ^{15}N enriched Ammonium Chloride and ^{13}C enriched Glucose, respectively. The protocol outlined in section 2.3.1 was followed until the 1L LB or TB Broth reached the above mentioned OD_{600} prior to IPTG induction. The cultures were pelleted by centrifuging at 10,000 x g for 15 mins. The supernatant media was decanted, and the cells were resuspended in 1 L M9 Minimal Media Base at pH 7.3 combined with Nutrient Mix (Summarized in Tables 2.2, 2.3 & 2.4). 0.5mgmL^{-1} Kanamycin was added to the M9 Media. The resuspended cell culture was allowed to adjust to the new, low nutrient media by shaking (150-200rpm) at 18 °C for 30 mins. IPTG was added to a final concentration of 0.5 mM and the cultures were incubated at 18°C for 16 hours. Cell cultures were harvested by centrifuging at 10,000 x g for 15 minutes and either used immediately or stored at -20 °C.

Table 2.2 M9 Minimal Media Base Recipe (For 1L M9 Media)

Chemical Name	Amount
Sodium phosphate, dibasic	42 mM
Potassium phosphate, monobasic	22 mM
Sodium Chloride	9 mM
ddH$_2$O	Up to 0.95 L

Table 2.3 M9 Minimal Media Nutrient Mix Recipe (For 1L M9 Media)

Chemical Name	Amount	Stock Concentration
Ammonium Chloride*	2 g	-
Glucose*	1 g	-
Magnesium sulfate	4 mL	1 M
Calcium Chloride, dihydrate	2 mL	50 mM
Thiamine	2 mL	20 mgml^{-1}
Iron (III) Chloride	800 μl	3 mM
Metal Mix (See Table 2.3)	500 μl	-
ddH$_2$O	Up to 50 mL	-

^{15}N Ammonium chloride and/or ^{13}C glucose were used depending on what type of labeling was desired

Table 2.4 M9 Minimal Media Metal Mix Recipe (500mL)

Chemical Name	Amount
Zinc sulfate, heptahydrate	0.575 g
Manganese (II) chloride, tetrahydrate	0.099 g
Boric Acid	0.145 g
Copper (II) sulfate, pentahydrate	0.087 g
ddH$_2$O	Up to 500 mL

2.3.3 Protein expression in modified M9 media for specific amino acid labelling

To increase NMR spectral quality by decreasing the number of peaks on the HSQC, addition of 1 g/L unlabeled amino acids (Ser, Gly) were added to the normal M9 Minimal Media recipe. The protocol in 2.3.2 was followed in the same manner following the addition of the unlabeled amino acid(s).

2.3.4 Primary Ni^{2+} NTA-affinity chromatography

Half of an EDTA-free protease inhibitor tablet (Sigma or Fischer) was dissolved in 25 mL binding buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole). Cell pellets, either fresh or frozen, were homogenized on ice with the protease inhibitor containing buffer and were lysed using sonication (5, 45 s cycles – 1 s on, 1 s off – amplitude 70-80 %). Soluble and insoluble fractions were separated via centrifugation at 70,000 x g for 45 minutes. The soluble supernatant fraction was filtered through a 0.45 μm syringe filter and purified via Ni^{2+} NTA-affinity chromatography on an AKTA Pure system (GE Healthcare) at 4°C. The filtrate was passed over a 5 mL His HP column (GE Healthcare) previously equilibrated with 3 column volumes of binding buffer at a flow rate of 1 mL/min. After the filtrate had been loaded onto the column at 0.5 mL/min, it was subsequently washed with 7 column volumes (35mL) of binding buffer at a flow rate of 1 mL/min. The elution of bound protein was performed with the up-flow setting in elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 250 mM imidazole) at 0.5 mL/min. 1 mL aliquots were collected until the A_{280} detecting aromatic residues in the protein reached baseline levels, indicating all protein had been eluted. Generally, elution was complete within 10 fractions (10 mL). Protein purity was assessed using SDS-PAGE and fractions containing the purest protein were pooled.

2.3.5 TEV Protease Cleavage and Dialysis

Protein purified from Ni^{2+} NTA-affinity chromatography was placed in a dialysis bag with 1 ml of 4.6 mgmL^{-1} His$_6$-TEV Protease to cleave the His$_6$ affinity tag of the purified protein. Cleavage took place over 16 hours in dialysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT) at 4 °C with constant stirring.

2.3.6 Reverse Ni^{2+} NTA-affinity chromatography

Cleaved protein was run over 2-3 mL Ni^{2+}-NTA resin on ice and the flow through was collected. Protein content in the resin flow through and wash steps was monitored using 5 drops of 1X Bradford Reagent and 10 μL of the eluent; a blue colour indicated that protein was present in the sample. Washes of 1 column volume were completed with dialysis buffer until no protein was present. Successful TEV cleavage was determined by the presence of protein in the flow through fractions and confirmed by an increase in band migration with SDS-PAGE. Protein was stored at 4 °C or flash frozen in liquid nitrogen and stored at -80 °C in 50 % glycerol.

2.4 Protein Thermal Shift for Buffer Optimization & Fragment Binding

Protein Thermal Shift (PTS) can be used to screen for optimal buffer conditions for subsequent assays by measuring the thermal stability, or the ability of a molecule to resist unfolding when subjected to thermal stress under various conditions[70,71]. The ThermoFischer QuantStudio 3 and the BioRad CFX-Connect rtPCR systems were utilized for these experiments.
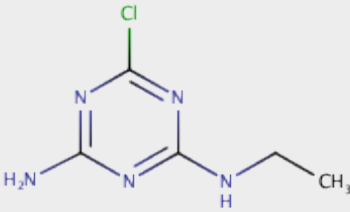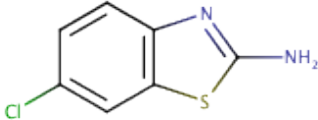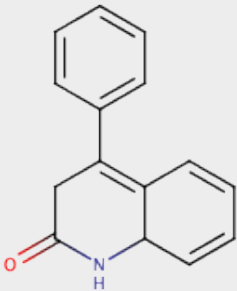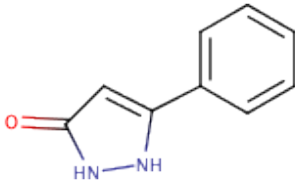
2.4.1 Buffer Optimization

A final protein concentration of 0.08 mgmL^{-1} was used in each assay with either 1X ROX or SYPRO Orange fluorescent dye in various buffer and salt conditions with a final volume of 25 μL. The 96-well plate was subjected to a temperature gradient from 25 to 96 °C and the fluorescence output was measured upon protein unfolding. The T_m, or the inflection point of the melting curve, was calculated with GraphPad Prism software and describes the thermal stability of the system; a higher T_m indicates a more thermally stable system, and theoretically correlates with increased sample longevity for other assay conditions.

2.4.2 Drug Fragment Screening

Fragment binding assays were completed in a similar set up to the buffer screens mentioned above. Buffer conditions for all fragment binding experiments were 50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP. Fragment concentrations varied from 0.1 to 200 μM and each well contained a final volume of 25 μl. The fragments used are summarized in Table 2.5 and were chosen based on initial screens performed by collaborator, Marc Lenoir. The fragments outlined in Table 2.5 are 4 of the 10 compounds chosen for further investigation. Fragment 2 was not soluble in ddH₂O, therefore, it was dissolved in DMSO prior to data collection. DMSO was normalized to 0.5% in all wells containing Fragment 2 and accompanying blanks.

Table 2.5 Summary of Fragments Utilized in Protein Thermal Shift Experiments

Compound	Structure
Deisopropylatrazine (Fragment 1)	
6-chloro-1,3-benzothiazol-2-amine (Fragment 2)	
4-phenyl-1,2-dehydroquinazolin-2-one (Fragment 3)	
5-phenyl-2,4-dehydro-3H-pyrazol-3-one (Fragment 4)	

2.5 ForteBio Octet RED96 Dip & Read Assay

The Octet RED96 system (ForteBio) was utilized to measure the binding affinity between CaM and CaMK1D. In these experiments, S-Tag CaM[72] (CaM + S-Tag fusion peptide with the sequence: KETAAAKFERQHMES) was provided by Dr. Larry Fliegel's laboratory as in-lab CaM purification and expression had not been optimized at the time this equipment was available. 200 nM WT CaMK1D with the His_6 affinity tag in 50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 1 mM $CaCl_2$ was loaded onto 6 Ni-NTA sensor chips on the Octet RED96 system and subsequently dipped into S-Tag CaM (concentrations of 31.25, 62.5, 125, 250, 500, and 1000 nM) in the same buffer, simultaneously detecting the binding interaction. This experiment was also completed with the CaMK1D 10-329 construct. Data was subtracted and analyzed using the built-in Octet® software.

2.6 Microscale Thermophoresis

Microscale Thermophoresis (MST) utilizes the concept that molecules of differing charge, size, and hydrodynamic radius will migrate through a thermal gradient at different speeds. These differences translate to a measurable difference in migration speed between the bound and unbound states of a molecule of interest[73]. Migration speeds are measured in different modes including fluorescence, which was used in this work, but also by A_{280} granted that the ligand of interest does not display inherent fluorescence. MST was used to measure the binding affinity of CaMK1D WT and S179E/T180E constructs with their activator molecule, CaM.

CaM-His_6 was labelled with reconstituted RED-tris-NTA dye (NanoTemper) which coordinates to the His_6 affinity tag of the purified protein. Each 200 μL labelling reaction contained 200 nM CaM in 50 mM HEPES pH 7.5, 10 mM NaCl, 1 mM $CaCl_2$, and 2 μL reconstituted dye. This

37

solution was incubated at room temperature for 30 mins followed by a 10 min centrifugation at 4 °C and 15,000 x g. 100 nM of labelled CaM was used in a 1:1 serial dilution of CaMK1D (WT and S179E/T180E) in the same buffer and which ranged in concentration from 6 nM to 200 µM, resulting in a final concentration of 50 nM CaM. All data was collected on a Monolith NT.115 (NanoTemper) and data analysis was completed using MO.Affinity Analysis software provided by NanoTemper.

2.7 Size Exclusion Chromatography coupled Small Angle X-Ray Scattering

Small Angle X-Ray Scattering (SAXS) was used to elucidate a low-resolution structure of the CaM/CaMK1D complex in an isotropic solution. Purified protein samples were concentrated to 9 mgmL^{-1} in Amicon Centrifugal Filters (0.5 mL, 10,000 MWCO for CaMK1D and 3,000 MWCO for CaM) and buffer exchanged into 50 mM HEPES pH 7.5, 10 mM NaCl, 5 mM CaCl$_2$. Samples for the CaMK1D/CaM complex contained 3 mgmL^{-1} CaM and 6 mgmL^{-1} CaMK1D for a roughly equimolar sample whereas single protein samples were each 9 mgmL^{-1}. 100 µL samples were flash frozen in liquid nitrogen sent to Diamond Light Source, Oxfordshire, UK for data collection. Each sample was passed through an appropriate SEC column (Superdex 200 or Shodex KW402.5 column) equilibrated with the above-mentioned buffer and subsequently subjected to the synchrotron X-ray source.

Data subtraction for SEC-SAXS was completed with the ScÅtter software (download available at bioisis.net) and the ATSAS suite[74] was utilized for the rest of the data analysis. The sample containing areas for subtraction were determined by plotting the Radius of Gyration (R_g) on the signal plot and only including points where the R_g values were similar (quality samples tended to have >20 data points with similar R_g values for averaging). The PRIMUS/qt program[75] was used

next to merge the subtracted data and generated a plot of $\ln(I)$ vs. q^2 to visualize the linearity of the Guinier region and its accompanying residuals plot to ensure the sample was monodispersed and non-aggregating[76,77]. The Kratky plot was used to determine the relative state of "unfoldedness" of the molecules, where the plot of q^2 x l(q) vs q forms a Gaussian peak when the macromolecule of interest is globular or folded and plateaus for unfolded molecules[78].

The Guinier analysis exclusively uses the data points from the upper portion of the curve to determine R_g, however, an indirect Fourier Transform was used to convert the reciprocal-space information of $\ln(I(q))$ vs. (q) into real space distance distribution function $(p(r))$ to utilize the whole data set. The p(r) distribution gives information on R_g and the maximum particle dimensions (D_{max}) of the experimental data and the shape of the function gives insight to the macromolecular shape of the sample[79,80]. *ab initio* bead modeling was then completed using DAMMIN[81] and DAMMIF[82] with 15 iterations completed on slow mode and averaged with the DAMAVER package[83] to give a final envelope structure. Manual fitting was completed in Chimera[84].

2.8 Nuclear Magnetic Resonance

All double and triple resonance experiments were performed at 25 °C (298K) on a Bruker 800 MHz NMR spectrometer with a 5 mm HCN CryoProbe. Optimized buffer conditions were 20 mM Tris pH 8.0, 150 mM NaCl, 0.5 mM TCEP for CaMK1D and 10 mM $CaCl_2$ was added to samples where CaM was present to load it with Ca^{2+}. Acquisition and analysis of NMR data was completed using TopSpin software.

Chapter Three

Representative Results

CHAPTER THREE – REPRESENTATIVE RESULTS

3.1 Protein Production

3.1.1 CaMK1D Purification & Buffer Optimization

CaMK1D was purified to >95 % purity, measured via densitometry, as shown in Figure 3.1. The chromatogram of the elution from the first IMAC column is shown on the left of Figure 3.1 and fractions in the peak area were pooled (shown on the SDS-PAGE gel as "IMAC 1 Pooled Fractions"). The TEV cleavage of these initial pooled fractions was successful due to the presence of CaMK1D in the flow through fractions of the second IMAC column (labeled "Reverse IMAC FT") and the noticeable increase in band migration which correlates to the decrease in CaMK1D size upon His_6 cleavage.
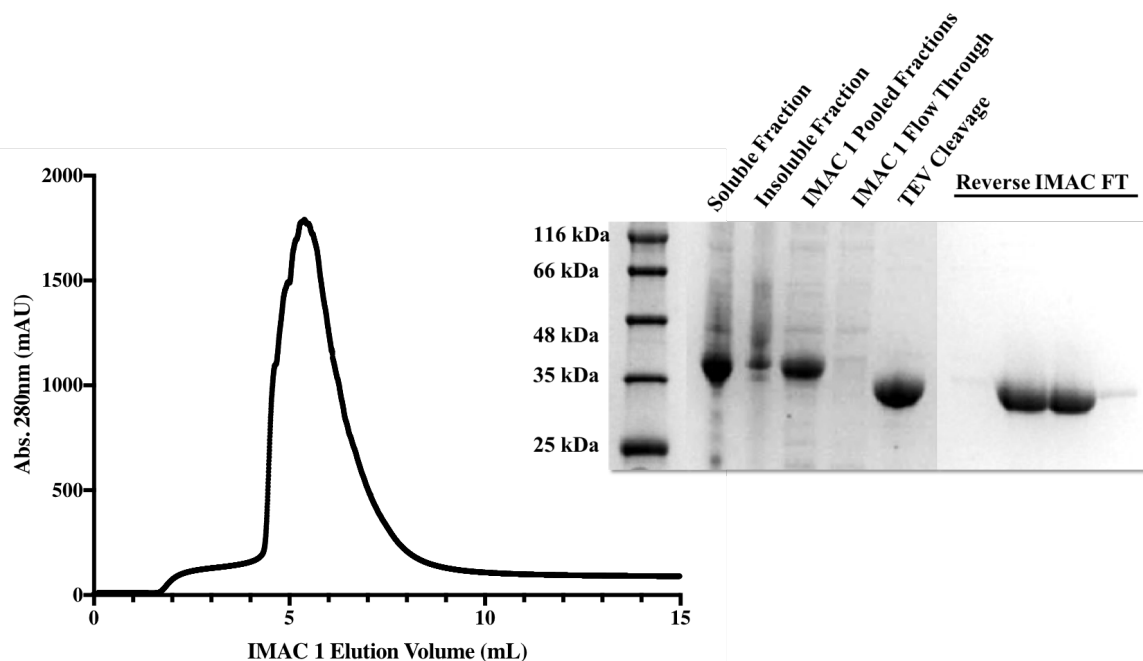
Figure 3.1 IMAC Chromatogram & SDS-PAGE of CaMK1D Purification
CaMK1D was purified to >95 % purity through an initial IMAC column (chromatogram shown on left) where the soluble protein fraction was loaded onto a 5 mL HisHP column (GE Healthcare). The protein containing fractions from the elution curve (IMAC 1 Pooled Fractions on SDS-PAGE) were subjected to TEV protease overnight. The Flow Through was collected from a second IMAC column. An any-kDa stain-free BioRad pre-cast SDS-PAGE gel (right) was run at 200 V for 25 minutes. Visualization was completed with A_{280}.

CaMK1D stability proved to be a problem early as it has a high propensity for precipitation. Structural techniques such as NMR require a highly concentrated sample that is non-aggregating over a long period of time and, therefore, PTS was used to optimize the buffer conditions for future experiments. PTS assays performed tested buffer conditions of with pH ranging from 6.0-8.0 and different salts such as NaCl, KCl, NH_4Cl, $MgCl_2$, and $(NH_4)_2SO_4$ in concentrations ranging from 0-300 mM. Of the buffer conditions tested, 25 mM Tris pH 8.0, 150 mM & 300 mM NaCl and 25 mM HEPES pH 7.5, 150 mM & 300 mM NaCl were deemed the optimal buffer conditions as they demonstrated the highest melting temperatures (Fig. 3.2). For purification
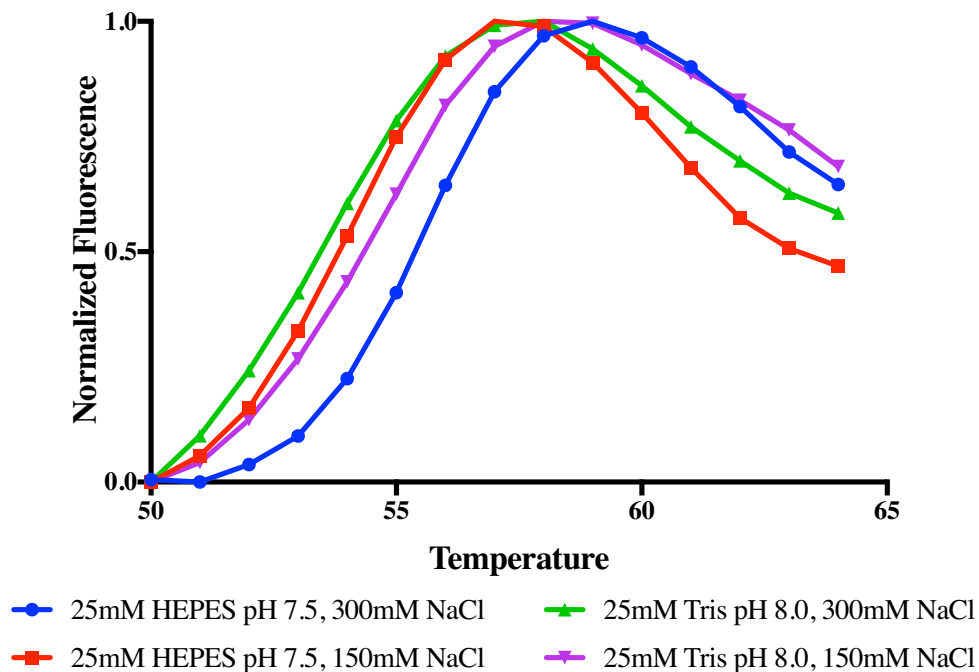
Figure 3.2 Protein Thermal Shift Assay Buffer Optimization for CaMK1D
PTS Buffer Screens were run on a ThermoFischer Quantstudio3 with a final concentration of 0.08 mgml[-1] WT CaMK1D 1-333 and 1X SYPRO Orange Fluorescent Dye. 25 mM HEPES pH 7.5 with 300 mM NaCl showed the highest melting temperature. Samples were prepared as described in section 2.4.

processes a HEPES buffer containing high salt was utilized as mentioned in section 2.3, however, increasing salt concentrations in NMR experiments leads to a sharp decrease in signal intensity. Therefore, a Tris buffer condition at pH 8.0 was implemented in these experiments and showed increased sample longevity at concentrations of up to 1 mM CaMK1D.

3.1.2 CaM Expression & Purification

CaM expression was determined to be the cleanest in the T7 Express *E. coli* cell line (Fig. 3.3) due to the absence of any strong bands other than that correlating to the desired CaM, which runs at ~16kDa. CaM was found to remain in the soluble fraction upon cell lysis in the 50 mM HEPES, pH 7.5, 500 mM NaCl, 20 mM imidazole, 0.5 mM TCEP buffer condition also utilized for CaMK1D purification. CaM expressed in T7 Express was purified to >95 % purity, measured with densitometry, with a single IMAC column (Fig. 3.4) utilizing the His$_6$ affinity tag. Data is not shown for additional steps of TEV cleavage and a second, reverse IMAC column where flow through fractions were collected to obtain cleaved CaM were utilized when an affinity tag free protein
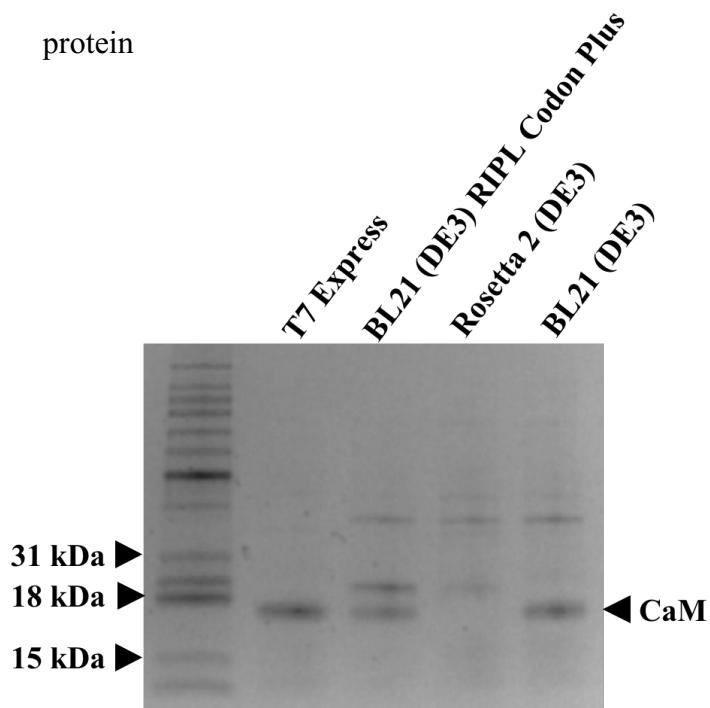
Figure 3.3 CaM Expression in *E. coli* cell lines CaM expression was induced with 0.5 mM IPTG and expressed for 1 hour at 37 °C. T7 Express showed the cleanest expression and was used for optimal CaM expression in subsequent experiments. An any-kDa BioRad pre-cast polyacrylamide SDS-PAGE gel was run at 200 V for 25 minutes and stained with Coomassie Blue dye for visualization.
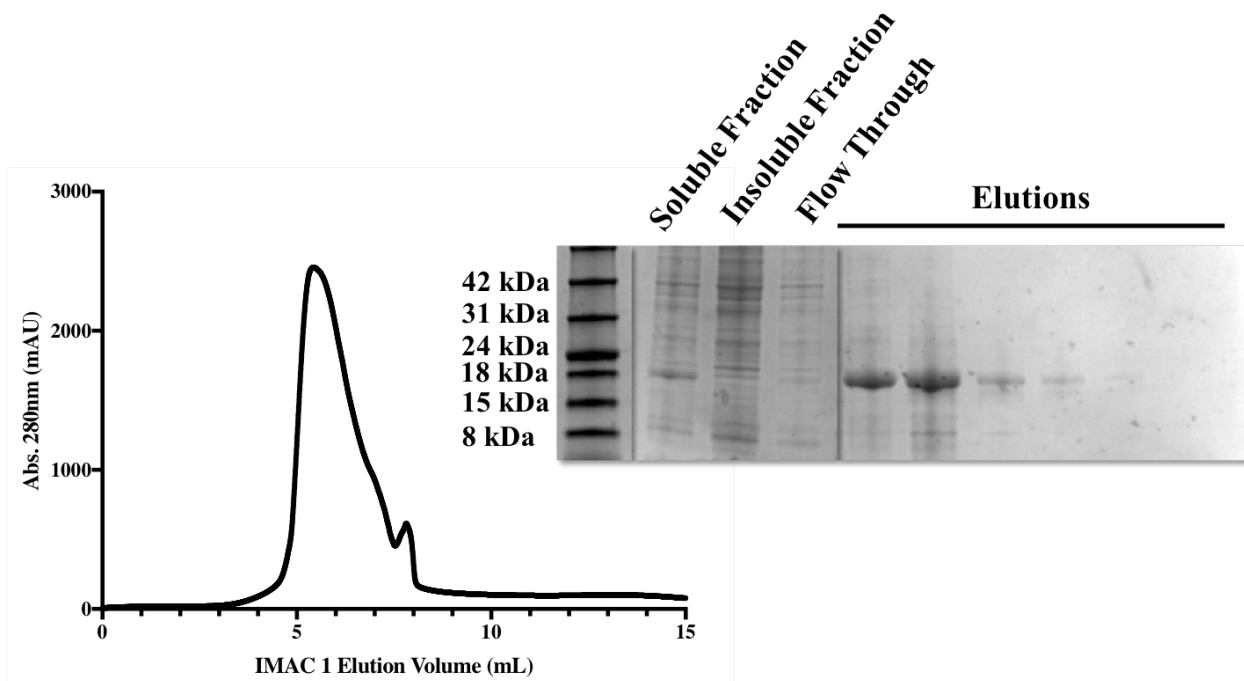
Figure 3.4 IMAC Purification & SDS-PAGE of CaM Purification
CaM was purified to >95 % purity through an initial IMAC column (chromatogram shown on left) where the soluble protein fraction was loaded onto a 5 mL HisHP column (GE Healthcare). An any-kDa pre-cast BioRad polyacrylamide SDS-PAGE gel (right) was run at 200 V for 25 minutes and stained with Coomassie Blue dye for visualization.

3.2 Drug Fragment Screening

Protein Thermal Shift was utilized as an initial screening method to determine which small molecules would bind to CaMK1D by measuring the change in the melting temperature (T_m). T_m is used as a measure of protein thermal stability, or how well the protein resists unfolding under thermal stress. Theoretically, thermal stability of a system should change upon ligand binding which results in a shift of the curve to either the left, suggesting a destabilizing effect, or to the right, a stabilizing effect. It was shown that fragments 1 and 2 (Table 2.6) bound to wildtype CaMK1D as a shift in the curves was observed, however, fragment 1 showed a concentration dependent decrease in T_m whereas fragment 2 did not (Figure 3.5).
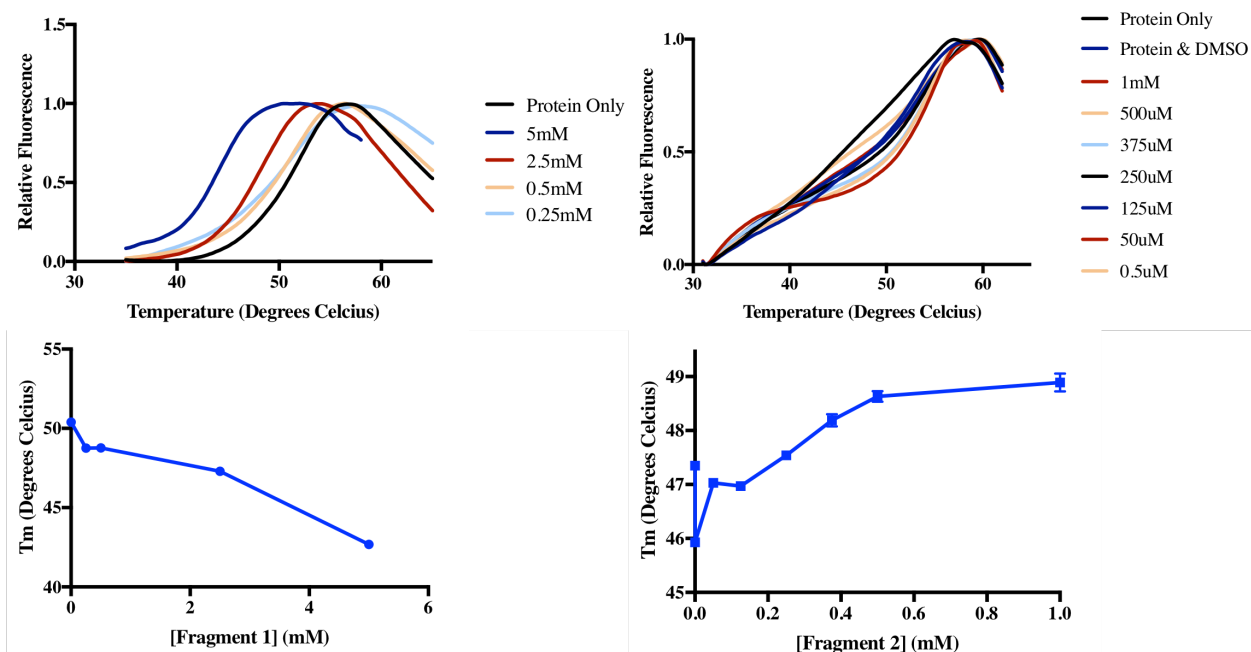
Figure 3.5 Protein Thermal Shift Fragment Screens
PTS fragment screening was performed on a BioRad CF-x rtPCR with Fragment one (Deisopropylatrazine) and Fragment 2 (6-chloro-1,3-benzothiazol-2-amine) at various concentrations in 50mM HEPES pH 7.5 buffer containing 150mM NaCl. [WT CaMK1D 1-333] = 0.08mgml^{-1} (n=4).

3.3 Structure and Kinetics of CaMK1D Natural Ligand Interactions

3.3.1 Small Angle X-Ray Scattering Solution Structures

Small Angle X-Ray Scattering provided solution structures of CaMK1D (WT and S179E/T180E), CaM, and their complex (Figs. 3.6-3.9). Initially, sample quality was based on the linearity of the Guinier, or upper region of the 1D scatter curve for all samples. The residuals plot for each sample was linear and did not show a "smile" which is observed when a sample exhibits any level of aggregation. All samples revealed a Gaussian-like peak in the Kratky analysis, suggesting that the samples were folded which was expected as both CaMK1D and CaM are globular proteins each containing one and two hydrophobic cores, respectively. The hydrodynamic parameters for each of the four samples are summarized in Table 3.1. SAXS analysis revealed R_g values of 2.4 nm, 2.5

nm, 2.2 nm, and 2.6 nm for WT CaMK1D, S179E/T180E CaMK1D, CaM, and the CaM/CaMK1D complex, respectively. D_{max} values calculated with the p(r) distribution for these samples were 8.5 nm, 8.7 nm, 7.2 nm, and 10.2 nm, respectively.

The DAMMIN envelope calculations all had Chi values of ~1.1, denoting a high-quality estimate of sample shape. The DAMMIN envelopes of the CaMK1D samples have a larger volume than that of the crystal structure, suggesting that there is some inherent flexibility in the kinase itself, likely at the C-terminal tail containing the AID/CBD. CaM also showed an envelope that was larger than its crystal structure, however, it is known that the linker that separates the two globular domains is inherently flexible which may account for this additional volume. The NMR structure of apo-CaM is shown in panel D of Figure 3.8 and its shape is consistent with that of the SAXS derived envelope. As expected, the CaM/CaMK1D complex exhibited the largest values for both R_g and D_{maxv} and an elongated, tri-lobed structure. It is likely that two of these lobes are from the kinase and one from CaM bound to the CaMK1D CBD. It appears as though CaM sits proximal to the C-terminal lobe of CaMK1D, potentially in a secondary binding site. Manual fitting of all data is shown in Figs. 3.6-3.9.

Table 3.1 Summary of experimental SAXS hydrodynamic parameters

Hydrodynamic Parameter	CaMK1D WT	CaMK1D S179E/T180E	CaM	CaM/CaMK1D Complex
R_g (nm)	2.4	2.5	2.2	2.6
D_{max} (nm)	8.5	8.7	7.2	10.2
Chi value	1.1	1.1	1.1	1.1
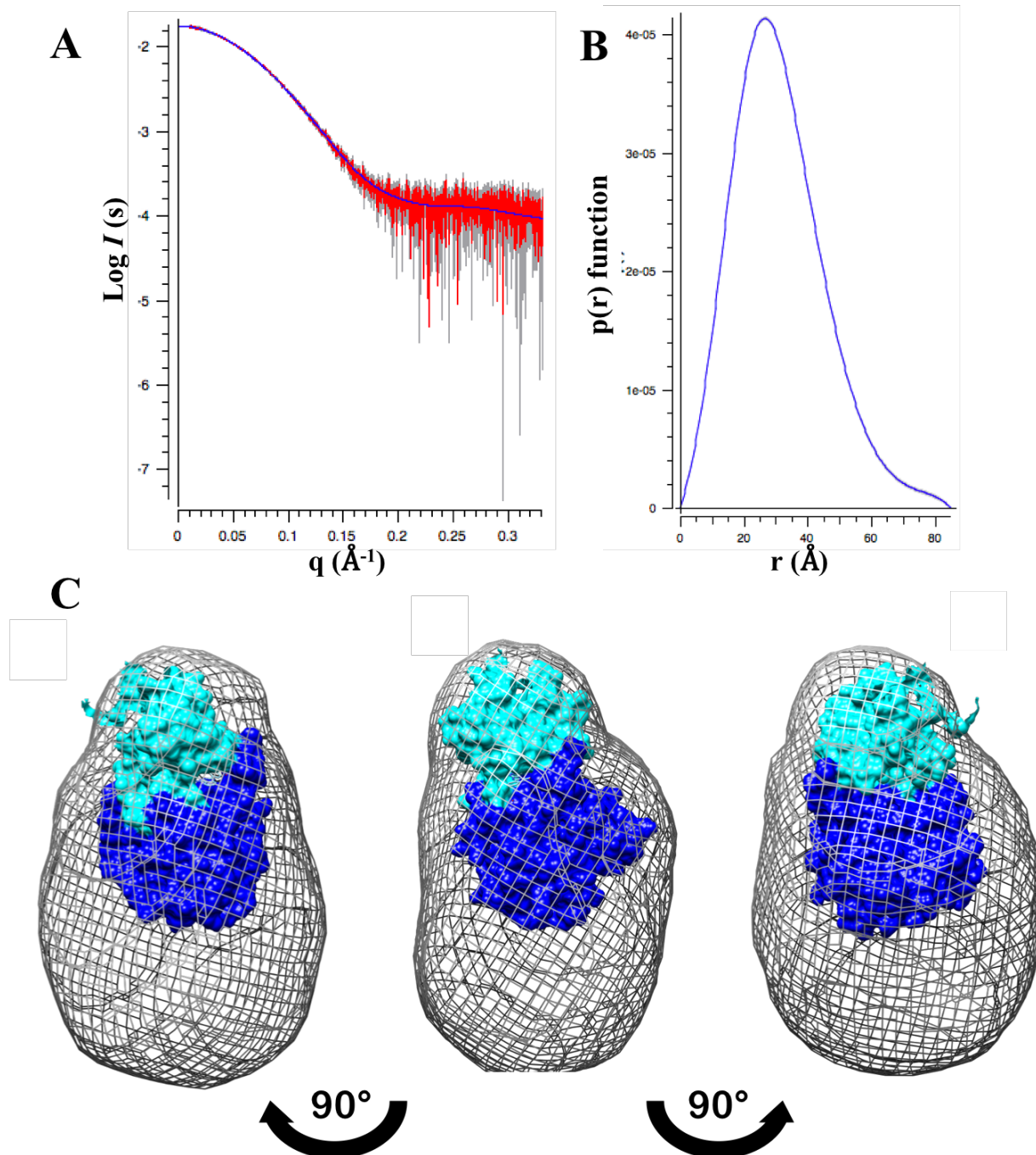
Figure 3.6 SAXS Solution Structure of CaMK1D Wildtype
9mgml^{-1} CaMK1D sample was run over a Superdex 200 column and immediately subjected to an X-ray source as described in section 2.6. A) 1D experimental X-ray scatter pattern of the complex (red) with the back calculation of the DAMMIN envelope fit overlaid (blue), B) p(r), reciprocal space, distribution plot, and C) density envelope of 10 *ab initio* bead models generated with DAMMIN and averaged with DAMAVER. Manual fitting of CaMK1D (PDB ID: 2JC6) done with Chimera. D_{max} = 8.5 nm, R_g = 2.4 nm
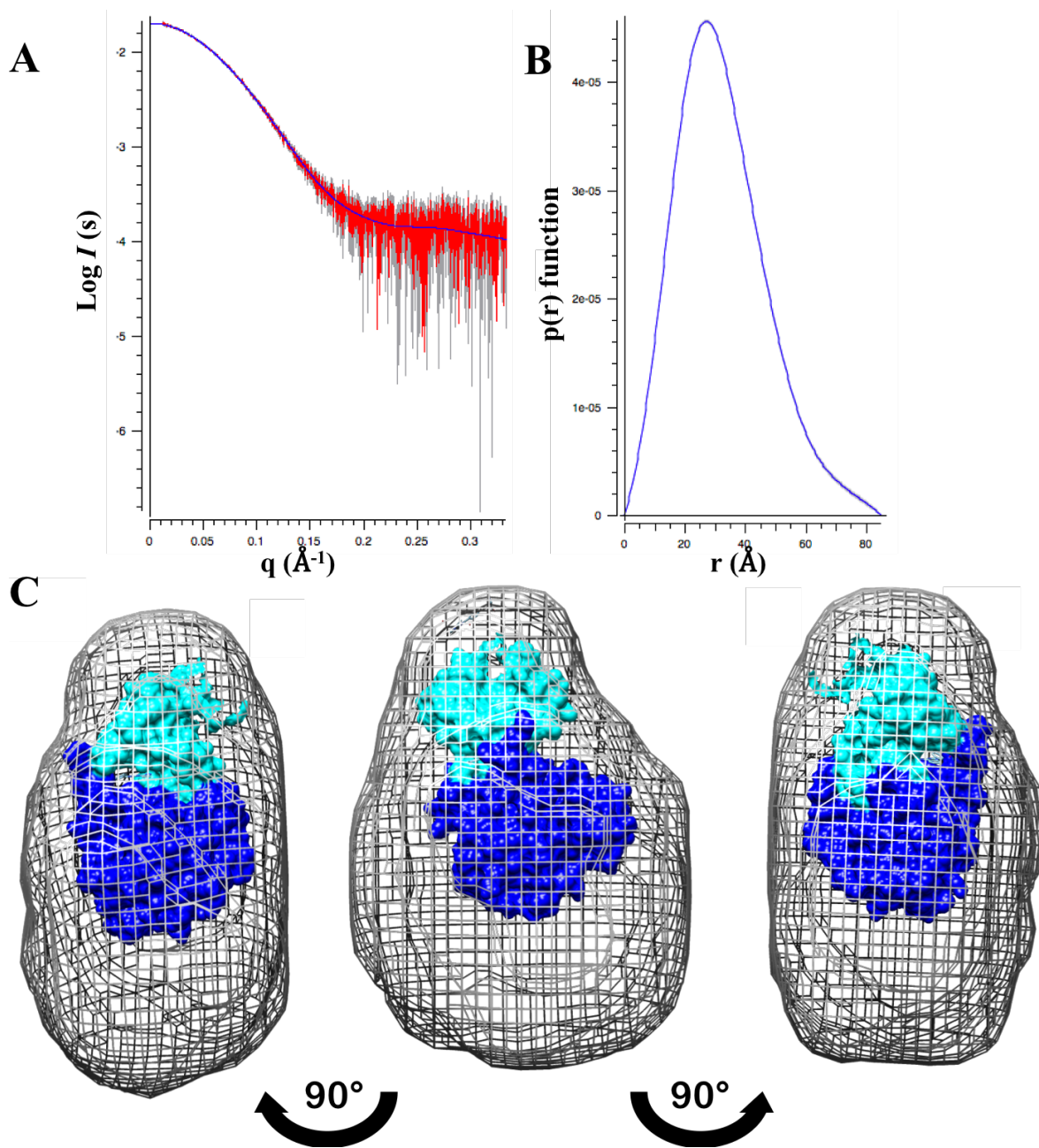
Figure 3.7 SAXS solution structure of CaMK1D S179E/T180E
9mgml^{-1} CaMK1D sample was run over a Superdex 200 column and immediately subjected to an X-ray source as described in section 2.6. A) 1D experimental X-ray scatter pattern of the complex (red) with the back calculation of the DAMMIN envelope fit overlaid (blue), B) p(r), reciprocal space, distribution plot, and C) density envelope of 10 *ab initio* bead models generated with DAMMIN and averaged with DAMAVER. Manual fitting of CaMK1D (PDB ID: 2JC6) done with Chimera. D_{max} = 8.7 nm, R_g = 2.5 nm
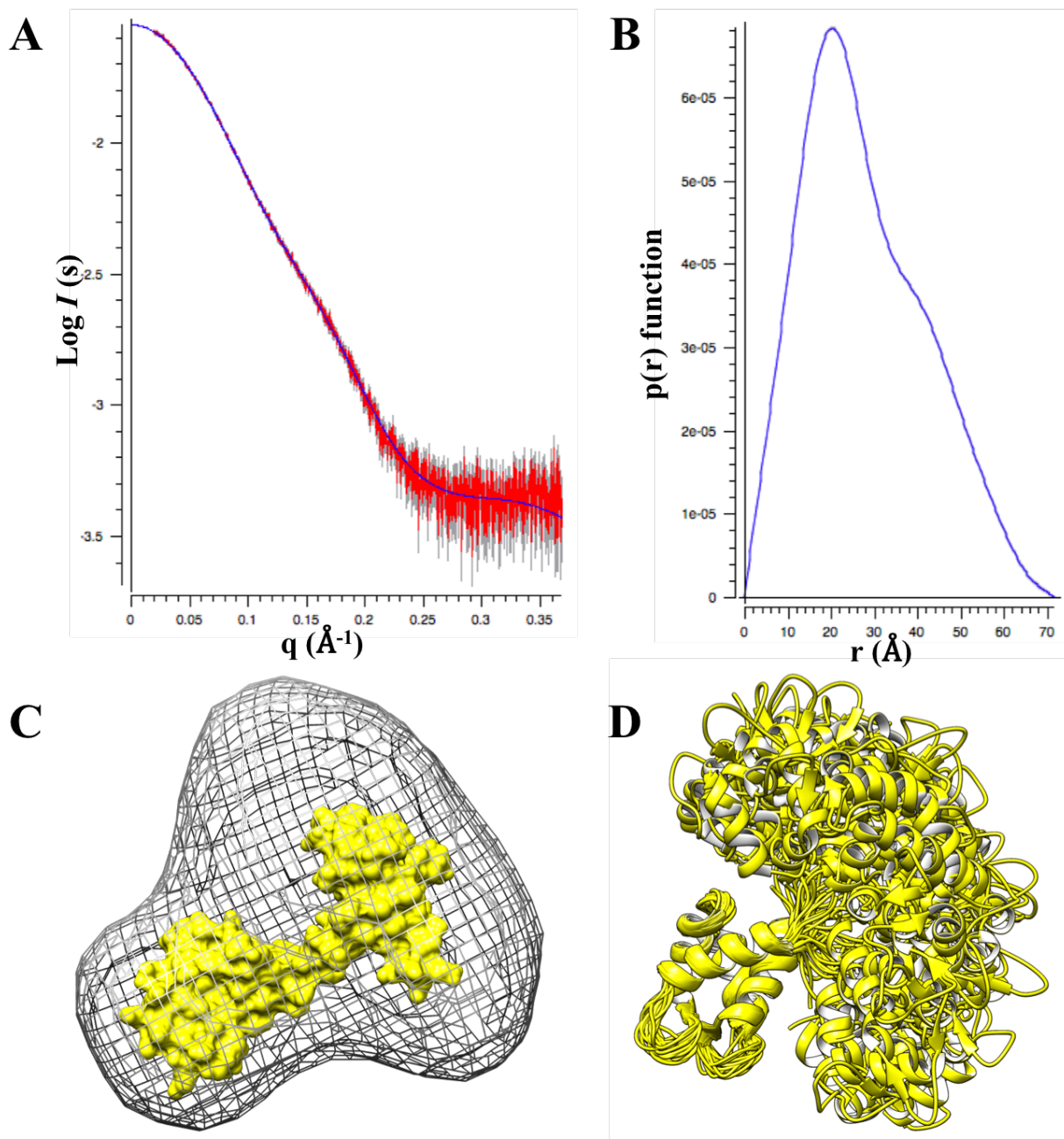
Figure 3.8 SAXS solution structure of calmodulin
9mgml^{-1} CaM sample was run over a Superdex 200 column and immediately subjected to an X-ray source as described in section 2.6. A) 1D experimental X-ray scatter pattern of the complex (red) with the back calculation of the DAMMIN envelope fit overlaid (blue), B) p(r), reciprocal space, distribution plot, and C) density envelope of 15 *ab initio* bead models generated with DAMMIN and averaged with DAMAVER. Manual fitting of CaM (PDB ID: 1IQ5) done with Chimera. D_{max} = 7.2 nm, R_g = 2.2 nm D) Ensemble of solution structures of CaM showing inherent flexibility in the linker domain (PDB ID: 1DMO)
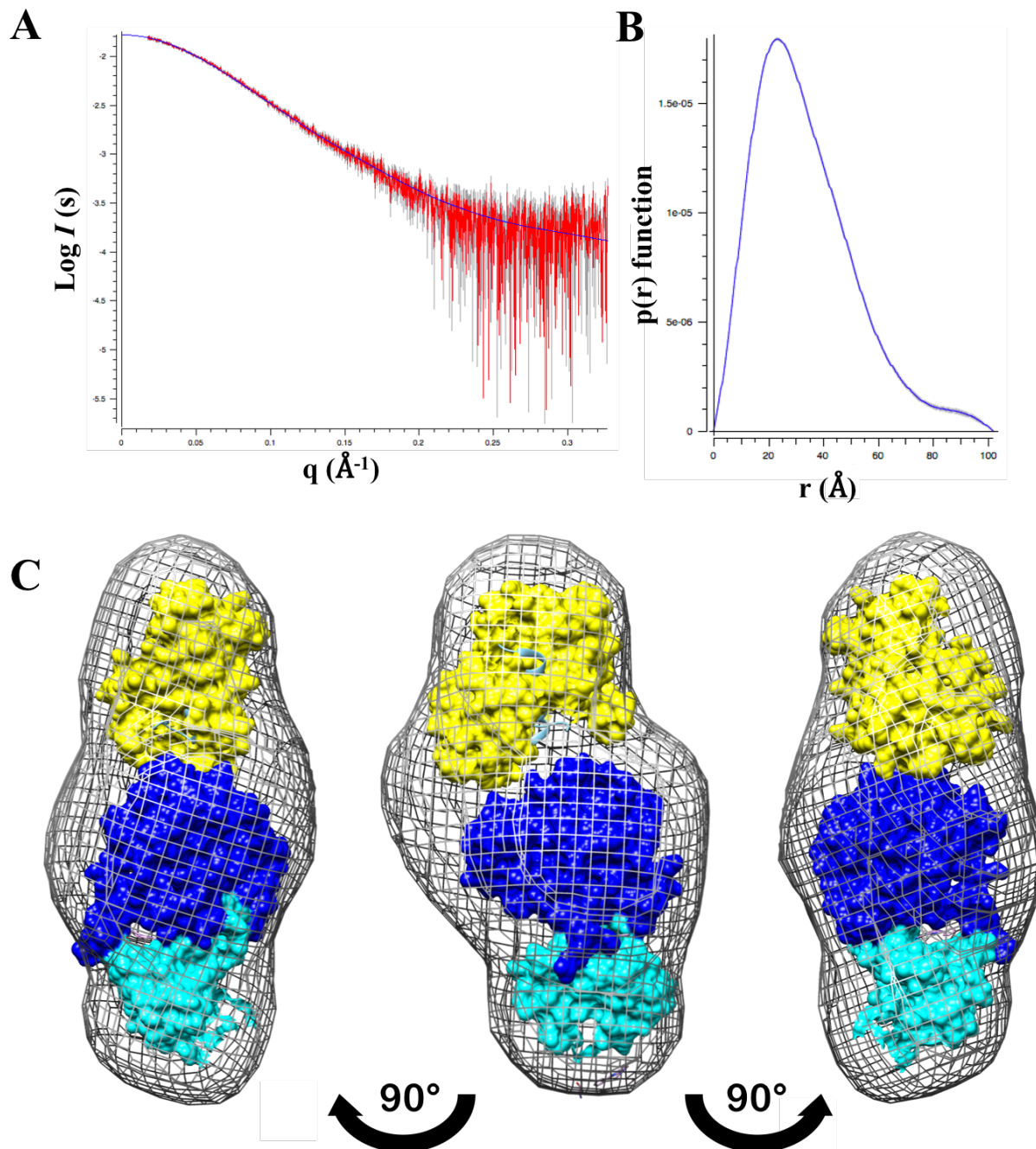
Figure 3.9 SAXS solution structure of the CaMK1D/CaM complex
Equimolar CaMK1D/CaM sample was run over a Shodex KW402.5 column and immediately subjected to an X-ray source as described in section 2.6. A) 1D experimental X-ray scatter pattern of the complex (red) with the back calculation of the DAMMIN envelope fit overlaid (blue), B) p(r), reciprocal space, distribution plot, and C) density envelope of 15 *ab initio* bead models generated with DAMMIN and averaged with DAMAVER. Manual fitting of CaMK1D (blue, PDB ID: 2JC6) and CaM bound to a CaMKII peptide (yellow, PDB ID: 1IQ5) done with Chimera. D_{max} = 10.2 nm, R_g = 2.6 nm

3.3.2 CaM/CaMK1D Interaction Kinetics with the Octet RED96 system

The ForteBio Octet RED96 system was used to determine the binding kinetics of the Ca^{2+}-CaM/CaMK1D interaction. 6 runs of the interaction were performed and the results are summarized in Table 3.2. A biphasic interaction was observed with both CaMK1D 1-333 and 10-329 (data not shown) constructs, denoting a binding interactions characterized by two binding events and suggesting that CaM may have multiple binding events with CaMK1D. The two interactions are characterized by K_{D1} and K_{D2} values of 0.53($\pm$0.01) nM and 3.89($\pm$0.15) nM, respectively. The tighter interaction, K_{D1}, showed a k_{on} = 8.60($\pm$0.08) x 10^5 $M^{-1}s^{-1}$ and a k_{off} = 4.57($\pm$0.07) x 10^{-4} s^{-1} whereas K_{D2} showed a k_{on} = 9.18($\pm$0.33) x 10^6 $M^{-1}s^{-1}$ and a k_{off} = 3.57($\pm$0.05) x 10^{-2} s^{-1}.
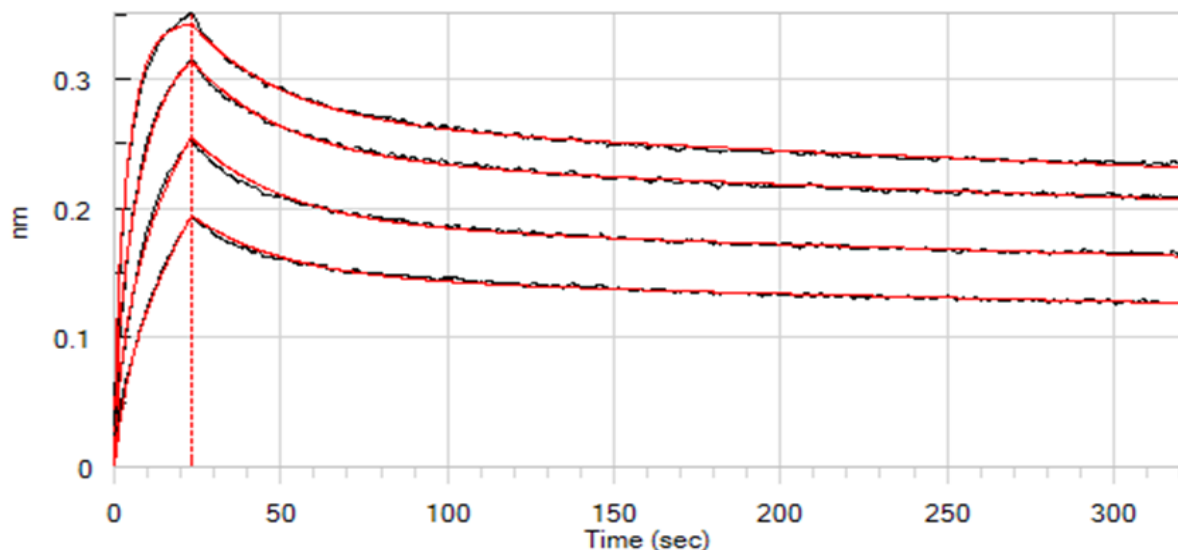
Figure 3.10 CaM/CaMK1D affinity determined by Octet RED96
200 nM His_6 WT CaMK1D was loaded onto a Ni-NTA sensor of the Octet RED96 system (ForteBio) and subjected to conditions with 30-1000 nM S-Tag CaM. The buffer contained 50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP. Data showed a biphasic interaction between CaMK1D and CaM with a K_{d}s of 0.53($\pm$.01) nM and 3.89($\pm$.15) nM for the two interactions. Data acquired and analyzed on Octet® Software (n=6).

Table 3.2 Summary of kinetic parameters of the CaM/CaMK1D binding interaction by the Octet Red system

K_{D1} (nM)	K_{on1} (1/Ms)	K_{off1} (s^{-1})
0.53($\pm$0.01)	8.60($\pm$0.08) x 10^5	4.57($\pm$0.07) x 10^{-4}
K_{D2} (nM)	**K_{on2} (1/Ms)**	**K_{off2} (s^{-1})**
3.89($\pm$0.15)	9.18($\pm$0.33) x 10^6	3.57($\pm$0.05) x 10^{-2}

3.3.3 CaM/CaMK1D Interaction Kinetics with MicroScale Thermophoresis

Microscale thermophoresis was completed as described in section 2.5. The binding affinity values disagreed with what was previously obtained from the Octet Red system, however, a comparative analysis between CaM binding to WT vs the A-loop S179E/T180E phosphomutant can still be performed. Binding of CaM to CaMK1D Wildtype showed a slightly lower affinity of 2.22($\pm$0.17) μM as compared to 1.50($\pm$0.09) μM for the A-loop phosphomutant (S179E/T180E) (Fig. 3.11). This suggests that A-loop phosphorylation does not play a significant role in changing the affinity of CaMK1D for CaM in the activation mechanism but likely only acts as a means to sustain long-term activity.

It is also interesting to note that many trials of this experiment were completed before successful binding was observed. Buffer conditions of 50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM CaCl$_2$, and 10 mM MgCl$_2$ was used in the first instance, however, binding checks on the Monolith NT.115 were not reproducible. A lower salt buffer containing 10mM NaCl was the only condition that revealed consistent, measurable binding between CaMK1D and the labeled CaM, suggesting that this interaction relies partially on an electrostatic mechanism.
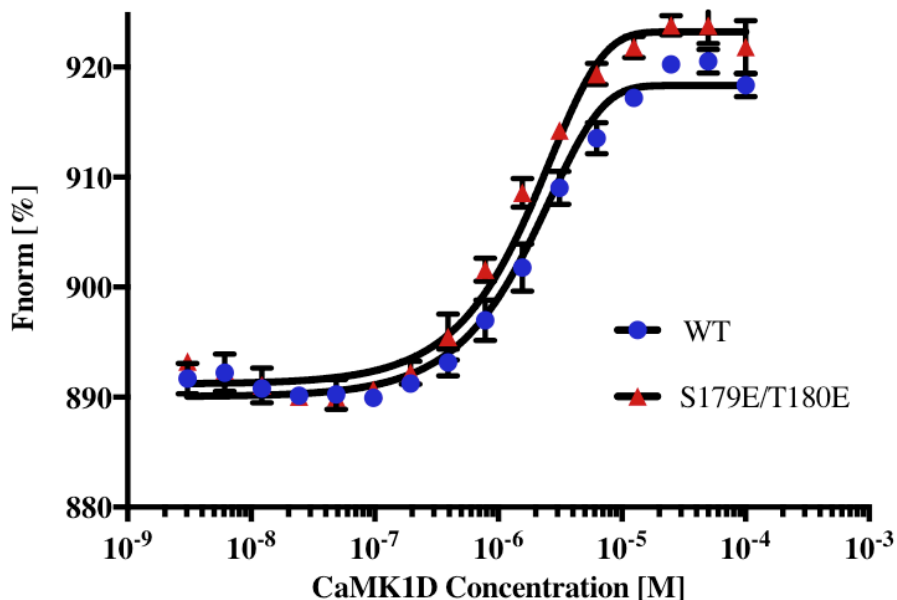
Figure 3.11 CaM/CaMK1D affinity determined by Microscale Thermophoresis
100 nM CaM-His$_6$ was labelled with RED-tris-NTA fluorescent dye (NanoTemper) and subjected
to a set of serial dilutions of CaMK1D. CaM showed an affinity of 2.22±0.17 μM for wildtype
CaMK1D and 1.50±0.09 μM for S179E/T180E (n=4). Data collected on a Nanotemper Monolith
NT.115 label system.

3.3.3 HSQC NMR Experiments to identify molecular contacts of CaMK1D interactions

NMR is technique that was employed to identify points of molecular contact between two binding

partners[67,85,86] and ultimately can be used to create docking models with programs such as

HADDOCK[87]. HSQC experiments were utilized to identify contact points between CaMK1D and

CaM because of the cost advantage provided by ^{15}N isotopes as compared ^{13}C glucose and the

relatively short experimental time required, since CaMK1D is prone to aggregation.

Fig. 3.12 shows HSQC spectra of ^{15}N WT CaMK1D alone (left) and ^{15}N WT CaMK1D with an

equimolar amount of unlabeled CaM (right). Backbone assignments of this CaMK1D construct

were previously completed to ~80%[88], though they were completed in a phosphate buffer with a

low salt concentration which was a problem since it was experimentally found that under low salt,

CaMK1D quickly and consistently precipitated out of solution, precluding the performance of lengthy experiments. Using a phosphate buffer has other potential implications with CaMK1D as it is an ATP binding protein and may have a tendency to bind phosphate. This could give us an unclear picture of the kinase and its interactions, which made 20 mM Tris pH 8.0, 150 mM NaCl, 0.5 mM TCEP a clear and safe buffer alternative as CaMK1D showed similar thermal stability in this condition as with the HEPES pH 7.5 buffer that was found to be most optimal with PTS assays outlined in section 3.1.1. The previously acquired phosphate buffer assignments overlap reasonably well with the Tris buffer condition at pH 8.0 and peak intensity and resolution is comparable. However, even without full assignments in this buffer condition, it is clear that line broadening is observed upon CaM addition, suggesting a slow exchange or tight binding interaction between CaM and CaMK1D. Simultaneous blank titration experiments of buffer instead of unlabeled CaM did not show this line broadening (data not shown), supporting this slow exchange interaction. The binding interaction between the Ca^{2+}-CaM/CaMK1D complex and an inhibitor known to bind to the ATP pocket of CaMK1D was also explored with NMR HSQC titrations (Figure 3.13). Upon the addition of 100μM of this inhibitor, there was significant line broadening observed, suggesting that this interaction is also characterized by slow exchange kinetics.

HSQC titrations of ^{15}N CaM with unlabeled CaMK1D are shown in Figure 3.14. The left most spectrum is ^{15}N CaM alone and the spectrum on the right is ^{15}N CaM titrated with a 1:1 molar ratio of CaMK1D:CaM. Clear line broadening of the CaM peaks indicates a slow exchange binding relationship between CaM and CaMK1D which supports the tight binding relationship observed in the reverse titration of ^{15}N CaMK1D with unlabeled CaM (Fig. 3.12). Again, simultaneous

buffer titration showed that the dilution effects of CaM with the addition of CaMK1D did not cause the observed line broadening indicating that the interaction is characterized by slow exchange kinetics. The interaction between CaMK1D and Ca^{2+}-CaM was further explored by overlaying previously acquired CaM assignments[89] and inferring suggestive assignments for the acquired spectra. The relative decreases in peak intensity were measured and are plotted in Figure 3.15. The C-terminal lobe of CaM seems to exhibit a more drastic relative decrease in peak intensity, however, it does appear to be a global effect. It is important to note that these assignments were not transferred to the spectra acquired in this thesis work, only overlaid to roughly estimate where binding was occurring.

Lastly, the optimization of NMR experiments to increase peak resolution by decreasing peak number for CaMK1D was completed by adding glycine to the M9 Minimal Media as described in section 2.3.3. Figure 3.16 demonstrates that there is a decrease in peak intensity corresponding to the region of the HSQC where glycine amide peaks would appear. This suggests that the unlabeled glycine addition to the media did facilitate the incorporation of unlabeled glycine into an otherwise uniformly ^{15}N labeled system.

A large amount of NMR data was collected on CaMK1D during this master's project. However, because transferring the phosphate buffer assignments to our new Tris pH 8.0 condition has been a non-trivial process, much of the data has not been used to its full extent. It is hoped that the transferred assignments can be used to determine the molecular contact points of natural ligands such as Ca^{2+}-CaM, ATP, and substrate peptides and to map lead molecules on the kinase surface.

Figure 3.12 HSQC Titration of ^{15}N CaMK1D with unlabeled CaM

The interaction between CaMK1D and its activator CaM is shown by ^{1}H, ^{15}N HSQC. The top panel (blue) represents CaMK1D alone and the bottom panel blue shows CaMK1D alone (blue) superimposed with CaMK1D with the addition of 1 molar ratio of unlabeled CaM (red). A noticeable reduction in peak intensity suggests that these binding partners interact via slow exchange kinetics and display a tight binding interaction. Spectra were acquired with 10% D_2O, 0.3mM CaMK1D in 20mM Tris pH 8.0, 150mM NaCl, 0.5mM TCEP, 10mM $CaCl_2$ on a Bruker 800MHz at 25°C.

CaMK1D-CaM 1:1
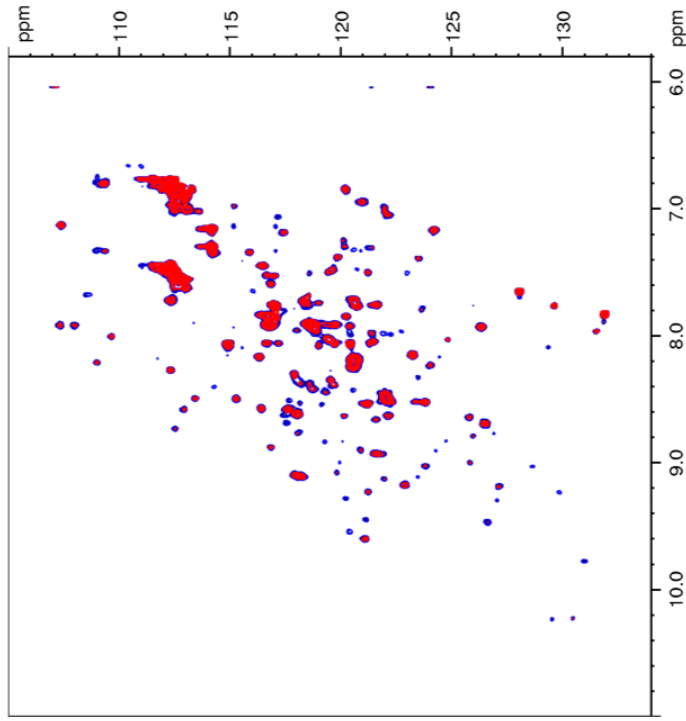
CaMK1D-CaM with inhibitor

¹⁵N Chemical Shift (ppm)
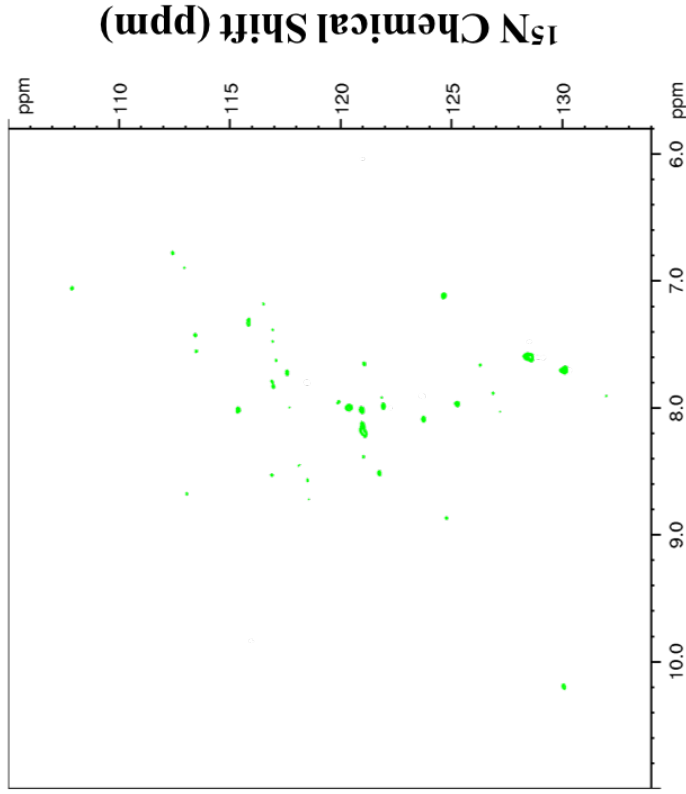
¹H Chemical Shift (ppm)

Figure 3.13 HSQC Titration of ¹⁵N CaMK1D bound with unlabeled Ca²⁺-CaM and an inhibitor
The interaction between Ca²⁺-CaM/CaMK1D and an inhibitor is shown by ¹H, ¹⁵N HSQC. The left panel represents Ca²⁺-CaM/CaMK1D and the right panel blue shows the Ca²⁺-CaM/CaMK1D complex with inhibitor present (green). A noticeable reduction in peak intensity suggests that these binding partners interact via slow exchange kinetics and display a tight binding interaction. Spectra were acquired with 10% D₂O, 0.1mM CaMK1D, 0.1mM CaM, 0.1mM inhibitor in 20mM Tris pH 8.0, 150mM NaCl, 0.5mM TCEP, 10mM CaCl₂ on a Bruker 800MHz at 25°C.

Figure 3.14 ^{15}N HSQC Titrations of 15N CaM with unlabeled CaMK1D

The interaction between CaM and its substrate CaMK1D is shown by ^{1}H, ^{15}N HSQC. The left panel (blue) represents CaM alone. The spectra on the right show CaM alone (blue) superimposed with CaM with the addition of a 1 molar ratio of CaMK1D (red). A noticeable reduction in peak intensity in the bottom panel suggests that these binding partners interact via slow exchange kinetics and display a tight binding interaction. Spectra were acquired with 10% D$_2$O, 0.3mM CaM in 20mM Tris pH 8.0, 150mM NaCl, 0.5mM TCEP, 10mM CaCl$_2$ on a Bruker 800MHz at 25° C.
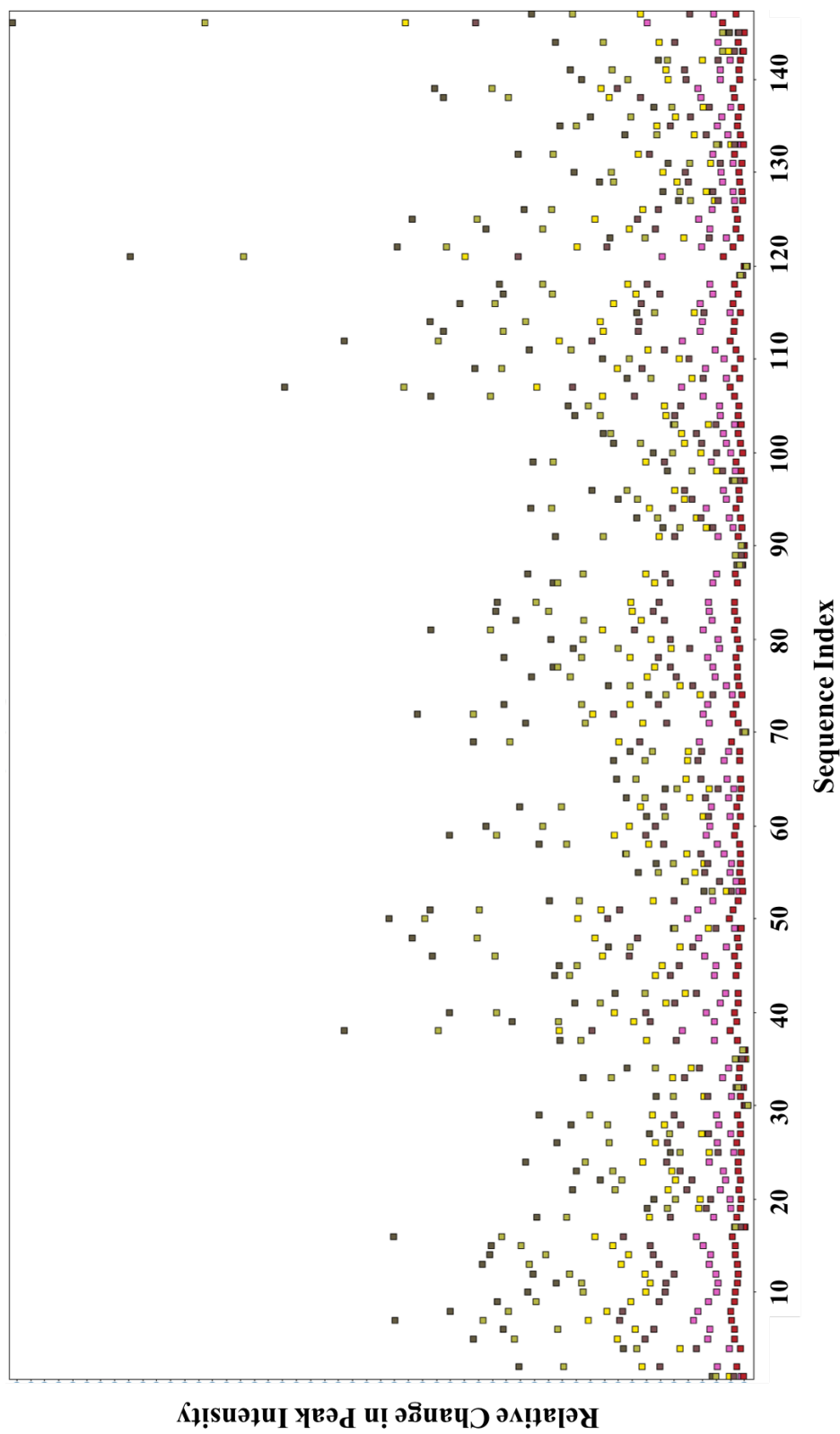
Figure 3.15 Perturbation plot of CaM titration
CaM chemical shifts from BMRB entry 4310 were overlaid on the HSQC titration data shown in Fig. 3.14 and suggestive assignments were inferred. There is a greater change in peak intensity upon CaMK1D addition in the C-terminal lobe of CaM, however, the peak intensity decrease seems like a global effect.
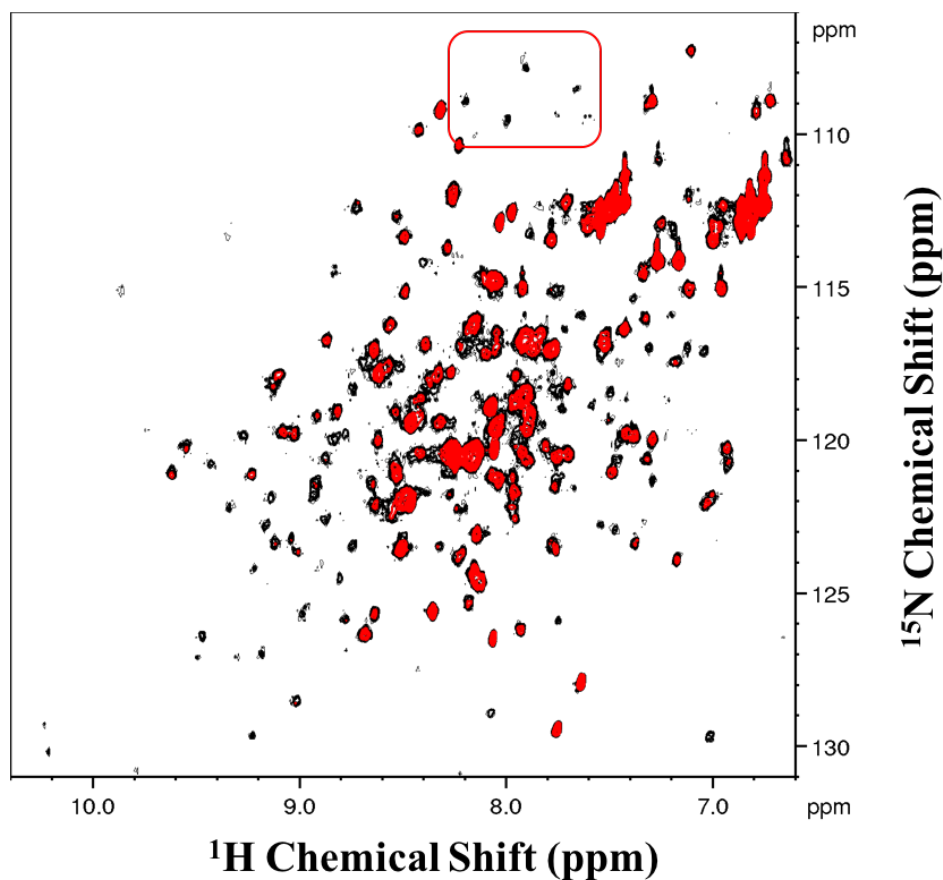
Figure 3.16 ^{15}N HSQC demonstrating Gly delabeling
To improve spectral quality by decreasing peak number, 1g glycine was added per L M9 Minimal Media. Glycine peaks generally appear in the upper portion of the HSQC fingerprint region (boxed in red). The black spectrum is CaMK1D fully ^{15}N labeled and the red spectrum is CaMK1D ^{15}N labeled but supplemented with unlabeled glycine. Other peaks that do not correspond to the glycine region of the spectrum are decreased and this is likely due to scrambling, as glycine can act as a precursor in the metabolism of other amino acids.

Chapter Four

Discussion

CHAPTER FOUR – DISCUSSION

4.1 Experimental Optimization

4.1.1 CaMK1D is stabilized by high sodium chloride concentrations

The optimization of experimental conditions was an important part of this project as CaMK1D quickly and consistently precipitated in phosphate buffers and any condition with a low salt concentration (<150mM NaCl). This was unsurprising as proteins that have large, charged surfaces often display instability in low salt conditions and it is thought that higher salt concentrations may act to reduce electrostatic interactions and improve protein stability of charged systems[90]. As demonstrated in Figure 4.1, CaMK1D displays many acidic and basic patches on its surface, which may contribute to this overall instability. It was found that NaCl was the only salt that conferred a stabilizing effect on CaMK1D, and the longevity of the sample was further increased in NMR experiments by utilizing Tris buffer at pH 8.0, which retained similar peak intensity as lower pH spectra.
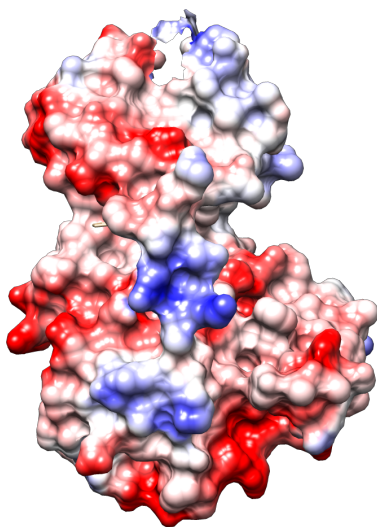
Figure 4.1 Electrostatic surface potential of CaMK1D
A schematic of the electrostatic surface potential of CaMK1D where red denotes acidic patches of negative charge and blue denotes basic patches of positive charge. Overall charges on a system are thought to contribute to instability in solution.

4.1.2 NMR experimental optimization

The CaMK1D construct utilized for NMR experiments produces a 37 kDa protein, which is on the upper end of what can be resolved by NMR for structural assignment purposes. T_2 (spin-spin) relaxation decreases with increasing molecular size, which is often a barrier in protein NMR. The size of protein kinases makes them difficult to work with, which is why only a relatively small number have been fully assigned[91–93]. However, new techniques such as specific labeling of the methyl groups of Ile, Leu, and Val residues have become more prevalent in recent years[94], as in the example of extracellular signal-regulated kinase 2 (ERK2)[95]. This specific labeling utilizes the high resolution of methyl peaks in ^{1}H-^{13}C correlated spectra and assumes an overall equal distribution of these residues on the protein surface and interior to give a representative picture of the kinase. Chemical shift perturbations can then be used to give general information about regions of ligand binding. However, the labeled precursors can be expensive and, therefore, this thesis explored glycine unlabeling as an inexpensive way to decrease peak number while increasing resolution for two-dimensional NMR experiments. The amide chemical shift for glycine generally appears in the upper portion of the fingerprint region in the HSQC spectrum, and Fig. 3.16 does reveal a decrease in peak intensity in this region. However, peaks that do not correlate to the glycine region are observed and this is likely due to scrambling of amino acids as glycine can be used a precursor in the synthesis of other amino acids. In the future, it is hoped that this work can contribute to the development of protocols that decrease the peak intensity of amino acids that are not involved in ligand binding, increasing resolution of peaks that can be used as reporters for ligand binding.

4.2 CaMK1D Regulation: the current understanding

Work presented in this thesis has aided in understanding how CaMK1D, an oncogenic kinase, is activated in the cell. By understanding the natural mechanism of CaMK1D activation, the drug discovery process can pinpoint where in this mechanism the kinase can be targeted for inhibition by either directly inhibiting kinase enzymatic activity or by disrupting protein-protein interactions necessary for activation[96]. Figure 4.4 summarizes our current understanding of CaMK1D regulation.

4.2.1 The two states of the C-terminal regulatory domain

It was understood that inactive CaMK1D existed in the cell as a monomeric kinase where the C-terminal regulatory domain, specifically the AID, acts to auto-inhibit the kinase by blocking the active site and substrate binding. However, SAXS data of the kinase alone revealed additional volume under the C-terminal lobe which likely correlates to the C-terminal regulatory domain of the kinase containing the CBD and AID. The current CaMK1D crystal structure (PDB ID: 2JC6) does not display density for the CBD as it is inherently flexible; the structure is missing approximately 30 amino acids on the C-terminal tail that could fill this volume under the C-lobe observed in the experimental SAXS envelope. From this SAXS envelope, we deduce that the CaMK1D C-terminus likely exists in an equilibrium between its auto-inhibitory position along the surface of the kinase and in a free state below the C-terminal lobe where it is accessible for CaM binding. Current crystallization studies are utilizing a shorter CaMK1D construct containing amino acids S10-A302 which displays higher stability than the commonly used M1-A333 construct, possibly due to the absence of the dynamic regulatory domain. Due to time constraints, a SAXS solution structure of this construct was not obtained, however, it would be expected that the

additional volume observed under the C-terminal lobe for the M1-A333 construct of CaMK1D would not be present in a structure of the S10-A302 construct.

4.2.2 The two-part CaM binding interaction

The first step of CaMK1D activation is the binding of CaM in a two-step mechanism. The kinetics of this interaction were initially investigated with the Octet RED96 Dip and Read Assay revealing a tight, biphasic interaction for the CaM/WT CaMK1D complex characterized by K_D values of 0.53($\pm$0.01) nM and 3.89($\pm$0.15) nM. These values are in the range of what was expected as CaM generally binds to its substrates with nanomolar affinity[97–101]. This biphasic interaction is of particular interest as it had been previously postulated that CaM may first bind to the CBD of CaMK1D but would have a final resting position elsewhere on the kinase surface.

A preliminary low-resolution structure of the Ca^{2+}-CaM/CaMK1D complex was revealed with SAXS which exhibits a novel, tri-lobed, elongated complex where it appears that CaM is situated on the lower portion of the C-terminal lobe (Fig. 3.9) which would correspond to its final resting position. Structurally, this is a unique interaction as CaM demonstrates distinctive interactions with different substrates (Fig. 4.2)[102,103]. A high-resolution structure of CaM bound to a CaMK protein was first revealed in the x-ray crystal structure of Ca^{2+}-CaM bound to Death Associated Protein Kinase (DAPK)[104] which exhibits complex architecture unlike what is experimentally observed with the Ca^{2+}-CaM/CaMK1D complex by SAXS. More recently, a 2010 study revealed a co-crystal of Ca^{2+}-CaM bound CaMKIIδ, which is a closer relative of CaMK1D[105]. This crystal structure demonstrates that Ca^{2+}-CaM binds the CBD, extending the AID and effectively removes it from its auto-inhibitory position. It would be expected that a kinase following the same,

canonical CaMK domain organization as CaMK1D would exhibit similar CaM binding. However, Ca^{2+}-CaM does not seem to have a secondary binding site on the CaMKII surface. This is in contrast to experimental SAXS data from CaMK1D which suggests that Ca^{2+}-CaM binds secondarily to its C-terminal lobe. Co-crystallization efforts of Ca^{2+}-CaM/CaMK1D have been unsuccessful up to this point, making this low-resolution SAXS solution structure particularly useful in understanding the CaMK1D activation mechanism.
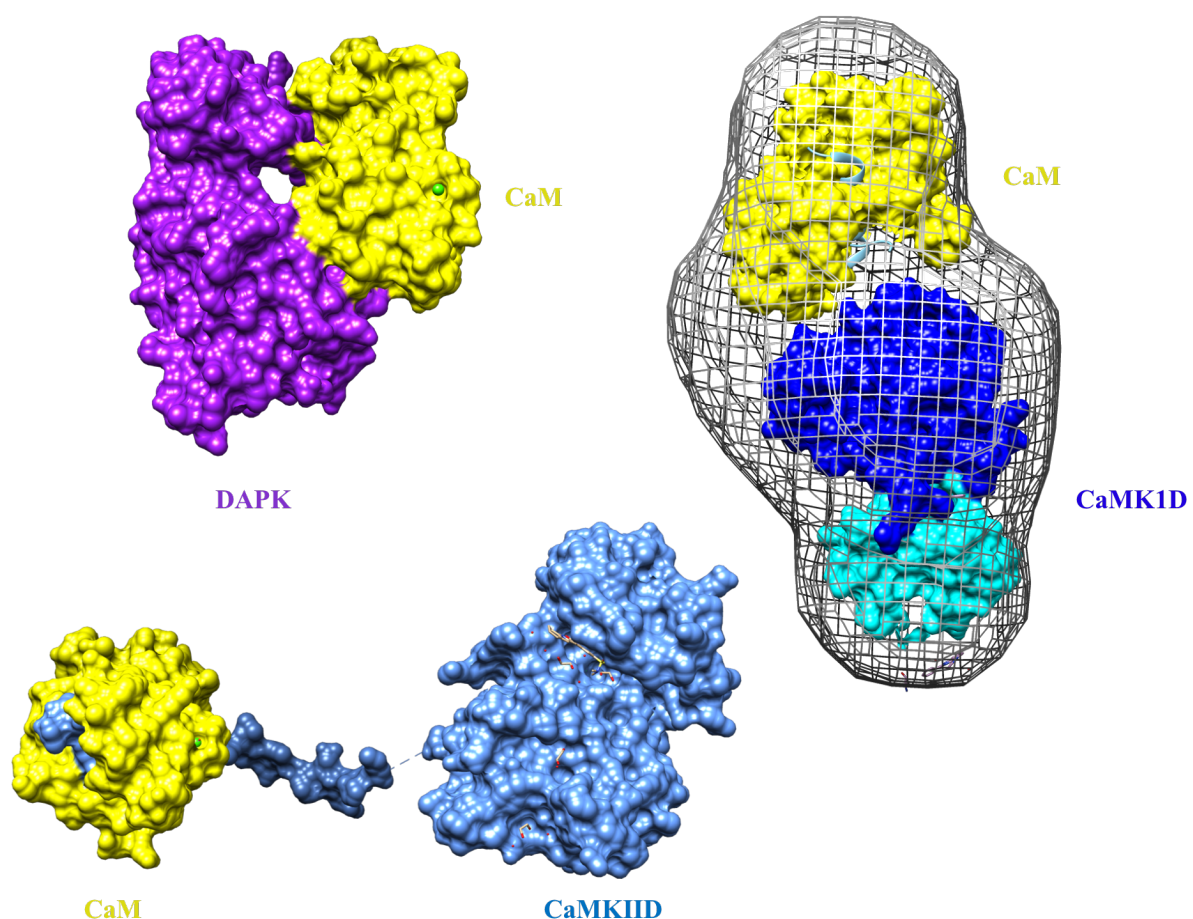
Figure 4.2 CaM displays unique interactions with different substrates
Though the CBD of different proteins are similar, unique CaM interactions are observed. CaM binds the regulatory domain of DAPK and then has a secondary binding site with contact points on both the N- and C-terminal lobes of the kinase (PDB ID: 2X0G). The more recently published crystal structure of Ca^{2+}-CaM bound CaMKIID demonstrates that CaM is only bound to the CBD of the kinase and does not have a secondary binding site on the kinase surface (PDB ID: 2WEL). The fit of the Ca^{2+}-CaM bound CaMK1D is shown to demonstrate its unique complex architecture.

NMR confirmed that the interaction between Ca^{2+}-CaM and CaMK1D is characterized by slow exchange kinetics, or a tight binding interaction, as line broadening was observed. Assignments have yet to be transferred to the Tris buffer condition from the phosphate buffer utilized in the original experiments. However, in light of this elongated complex, it would be likely that chemical shift perturbations and/or line broadening would be observed for peaks associated with both the regulatory domain (AID/CBD) and the C-terminal lobe of the kinase where a secondary binding site for CaM may reside.

MST experiments revealed a binding affinity that was in the micromolar range for the Ca^{2+}-CaM/CaMK1D interaction, 3 orders of magnitude greater than what was revealed by the Octet Red System. The affinity values are on the upper end of what is generally seen with the affinity of CaM binding interactions. It's possible that these values disagree with previous data due to the sensitivity of MST methods to salt, which was observed by other Overduin lab members as well as with the CaMK1D project. By lowering the NaCl concentration to <10mM, it's likely that this interaction would appeared to have strengthened. A sensitivity to salt often denotes an electrostatic component to the interaction[106], and Ca^{2+}-CaM interacts with its substrate through a combination of hydrophobic and electrostatic interactions[97,107–109]. The overall charge of CaM (-15 at pH 7.5)[110] is complemented by the positively charged CBD of CaMK1D from residues K303, K305, R307, R317, R320, and K321 which would facilitate an electrostatic interaction. The high likelihood of an electrostatic interaction is emphasized in Figure 4.3.

Furthermore, MST experiments performed in this work utilized a fluorophore that was bound to the N-terminal His$_6$ affinity tag of CaM. Since CaM undergoes a dramatic conformational change upon substrate binding, the fluorophore attachment may slightly disrupt CaM binding capabilities, resulting in a perceived lower affinity interaction. However, the likelihood of this is difficult to validate since NanoTemper does not release specifics of this compound.

A
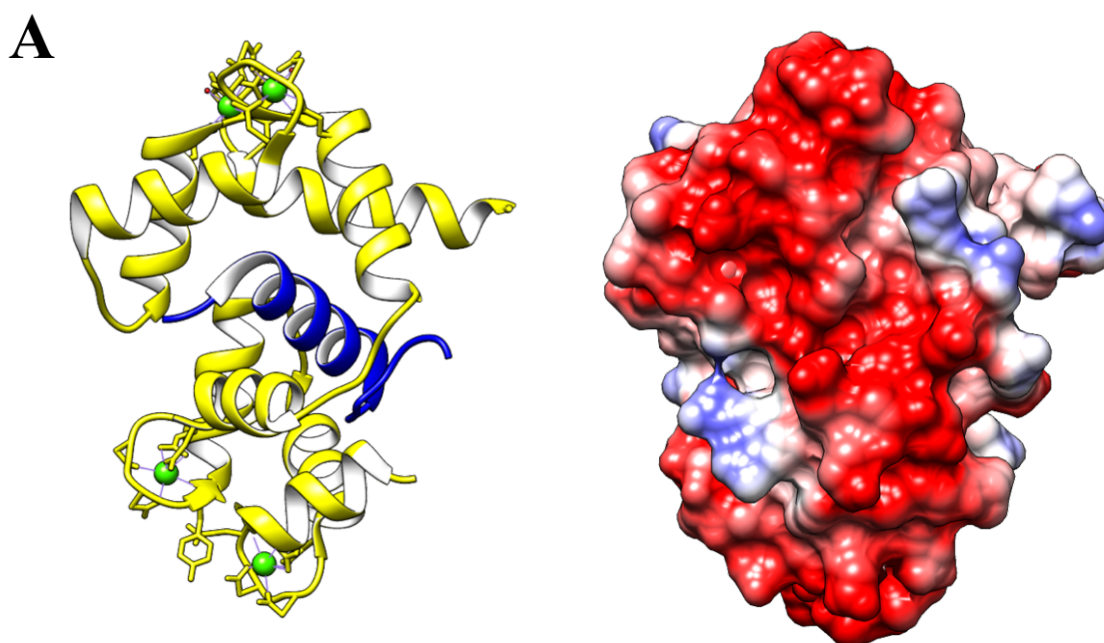
B

303-**KSK**WRQ<u>AF</u>N<u>A</u>T<u>AVV</u>RHM**RK**<u>L</u>-322

Figure 4.3 CaM interacts with substrates with electrostatic and hydrophobic interactions
A schematic of CaM interactions with a substrate peptide. A) Crystal structure of Ca^{2+}-CaM bound to CaMKII peptide (left) and the corresponding electrostatic potential where red denotes acidic patches of negative charge and blue denotes basic patches of positive charge. It is clear in this diagram that CaM has a large acidic patch which it uses to interact with substrates through electrostatic interactions. (PDB ID: 1IQ5) B) The sequence of the CaMK1D CBD where the interior of the peptide, which would be buried in the bound form, displays a high number of hydrophobic residues (underlined) whereas there is a large number of basic residues along the edges of the CBD that could interact with the acidic patch of the CaM surface.

4.2.3 The role of phosphorylation in CaMK1D regulation

It has been well established that CaMKs are regulated by phosphorylation[38,41,55], and CaMK1D is no exception. The phosphorylation of T180 is validated as a post-translational modification that serves to sustain activated kinase activity[55]. MST experiments were performed with an A-loop S179E/T180E phosphomutant where glutamic acids replaced the serine and threonine residues to mimic the negative charge of a phosphorylation event. It was postulated that T180 phosphorylation may increase the affinity of CaMK1D for CaM, however, it was determined that the binding affinities were not significantly different between the wildtype and mutant forms. This suggests that the charge state of the activation loop does not impact the binding of CaM, and this data aids in determining the order of events in the CaMK1D activation pathway. The K_D of CaM binding does not change with the phosphorylation of T180 and this supports the notion that the activation loop may not be fully accessible to CaMKK1 or CaMKK2 for phosphorylation until after Ca^{2+}-CaM is bound and removes the AID from the kinase surface.

More recently, the story of CaMK1D phosphorylation in enzymatic regulation has become more developed where collaborator Colin Kenyon is working with mass spectrometry to identify potential phosphorylation sites. Of particular interest is the phosphorylation of T313, which sits within the CaMK1D CBD. The centre of the CBD is generally hydrophobic as outlined in Fig. 4.3B, and the addition of a negative charge at this position would disrupt any hydrophobic interactions in this area. Moreover, Fig. 4.3A demonstrates the vast negative charge observed on the CaM surface which would likely exhibit a repulsive charge effect with a negatively charge phosphate. As outlined in section 1.2.2, the phosphorylation of T305/306/307 of CaMKII, which

is located in its CBD, acts as an inhibitory event that blocks CaM binding. Preliminary MST experiments with a T313E CaMK1D mutant are not showing Ca^{2+}-CaM binding and is currently being investigated further.

4.2.4 Summary of CaMK1D regulation

The current understanding of CaMK1D enzymatic regulation is summarized in a schematic in Figure 4.4. First, the C-terminal regulatory domain exists in an equilibrium between its auto-inhibitory position along the kinase surface and a free state accessible to Ca^{2+}-CaM binding. CaM then binds in a two-part mechanism that is characterized by a first binding interaction to the CBD of CaMK1D and then a secondary binding site on the bottom of the C-terminal lobe. Only once Ca^{2+}-CaM is bound is the activation loop accessible for phosphorylation of T180 by CaMKK1 or CaMKK2. We have also proposed a secondary phosphorylation site, T313, that is homologous to T305/306/307 of CaMKII isoforms that is an inhibitory phosphorylation site that blocks Ca^{2+}-CaM binding and subsequent activation.
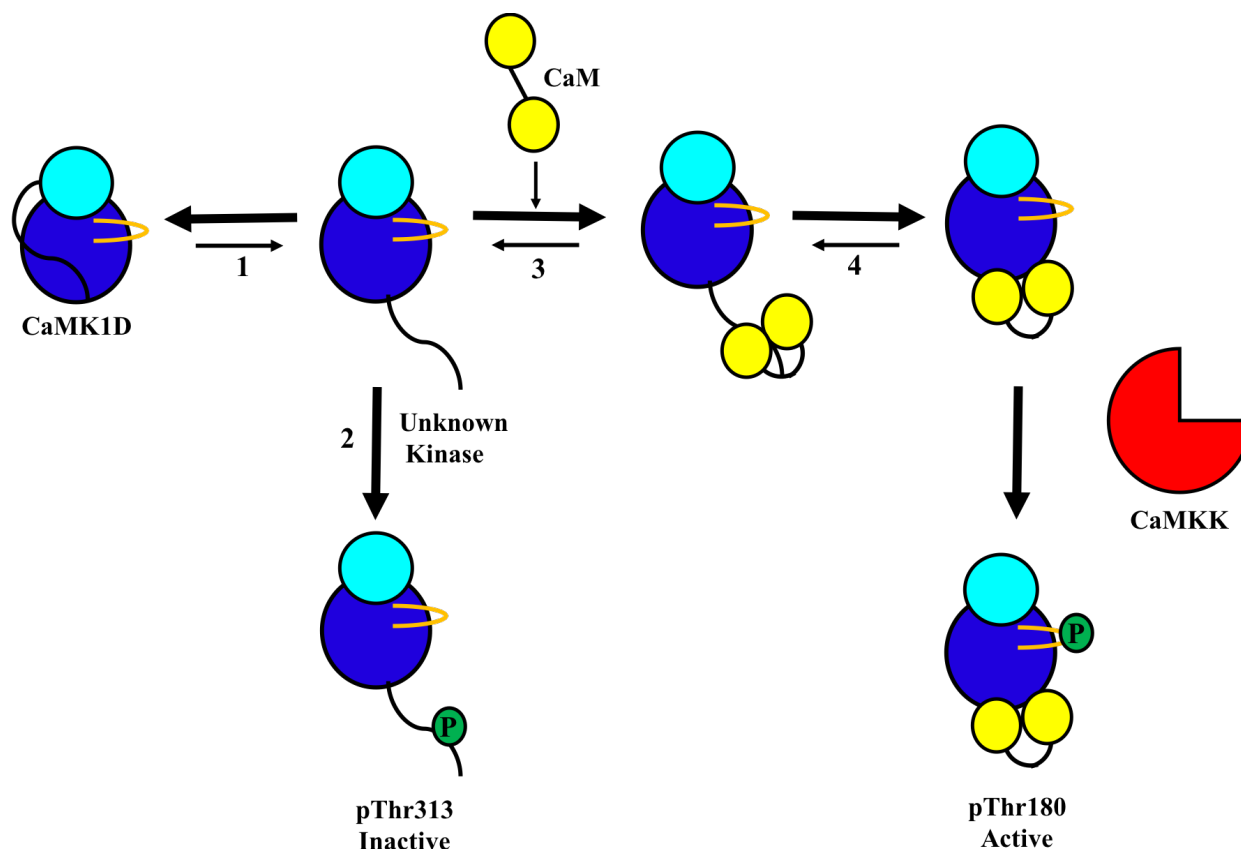
Figure 4.4 CaMK1D Regulation
A schematic of the currently understanding of the CaMK1D regulation mechanism. 1) the C-terminal regulatory domain exists in an equilibrium between its auto-inhibitory state and a free state accessible to Ca^{2+}-CaM binding. 2) Phosphorylation of T313 by an unknown kinase inhibits Ca^{2+}-CaM binding to the CaMK1D CBD. 3) Ca^{2+}-CaM binds to the CBD of CaMK1D. 4) Ca^{2+}-CaM binds to its secondary binding site on the C-terminal lobe of CaMK1D. 5) CaMKK1 or CaMKK2 phosphorylates T180 on the activation loop of CaMK1D, rendering it fully active.

4.3 Drug Fragment Screening and Drug Discovery

Protein Thermal Shift assays were performed to verify this method as a cost effective, quick way to screen drug fragments in the drug discovery pipeline. Theoretically the thermal stability of a protein system, CaMK1D in this case, should be changed upon ligand binding and one would expect a concentration dependent shift in T_m with an increasing ligand concentration. However, when deisopropylatrazine and 6-chloro-1,3-benzothiazol-2-amine (fragments 1 and 2 outlined in Table 2.6) titrations were completed with WT CaMK1D 1-333, there was only a concentration

dependent shift in T_m observed with 6-chloro-1,3-benzothiazol-2-amine. This suggests that although PTS experiments may be a good initial screening method to determine if a drug fragment binds, generating a binding curve that may be indicative of the binding affinity is not feasible at this point.

Though fragment screening was not a main focus of this thesis, the structural data obtained on Ca^{2+}-CaM/CaMK1D complex formation is of particular interest in drug targeting. The secondary binding site of Ca^{2+}-CaM on the bottom of the CaMK1D C-terminal lobe observed in the SAXS solution structure uncovered a protein-protein interaction site that could be targeted in drug discovery efforts. Targeting protein-protein interaction interfaces for enzymes has been proven effective in the example of cyclin-dependent kinases, where synthetic peptides were designed to competitively inhibit the recruitment of the cyclin subunit. Furthermore, the Bcr-Abl kinase that was discussed in section 1.1.3 of this thesis was originally inhibited by the compound imatinib, an ATP-competitive inhibitor[15,16]. However, it was found that resistance became a problem with this compound and the development of a non-ATP-competitive inhibitor was shown to be more effective in treating CML by inhibiting the kinase[112]. These two examples emphasize why the targeting of unique architecture and protein-protein interaction interfaces are effective in the development of pharmaceuticals, and why targeting of the interaction surface between Ca^{2+}-CaM and CaMK1D may be a viable, druggable surface.

Chapter Five

Conclusions and Future Perspectives

CHAPTER FIVE – CONCLUSIONS AND FUTURE PERSPECTIVES

5.1 Conclusions

The work in this thesis explores the natural regulation of CaMK1D, a putative driver of Triple Negative Breast Cancer (TNBC). There is much left to be explored in terms of CaMK1D regulation and ligand binding, however, the three hypotheses posed at the beginning of this project were, at least in part, addressed.

Firstly, drug fragment binding was explored with PTS assays to determine if it could be used to cheaply and effectively estimate the binding affinity of drug fragment molecules and to cross validate results from other methods in the screening process. A concentration dependent shift in thermal stability, measured by T_m, was observed with one of the two fragments tested. Nonetheless, a full sigmoidal binding curve was not observed concluding that PTS is not a reliable way to determine the relative binding affinity of drug fragment molecules. PTS is, however, able to cross-validate drug fragment molecules that bind to protein targets as a reproducible shift in thermal stability was observed with each fragment.

CaMK1D regulation and interactions with natural ligands was a major focus of this thesis work. An overview of the current understanding of CaMK1D regulation is demonstrated in Fig. 4.3. It was hypothesized that CaM would bind to the CBD of CaMK1D to regulate its enzymatic activity. This hypothesis was supported by kinetic assays that determined Ca^{2+}-CaM binds the CBD of CaMK1D with nanomolar affinity in a biphasic interaction, suggesting that this interaction is

characterized by two binding events. We postulate that the first, tighter binding interaction is that of Ca^{2+}-CaM binding to the CBD of CaMK1D and that the second is an interaction between bound Ca^{2+}-CaM and the C-terminal lobe of the kinase which secures CaM in its final binding site.

Lastly, phosphorylation has been shown to play a role in the regulation of CaMK family members, including CaMK1D. T180 was hypothesized to be phosphorylated post Ca^{2+}-CaM binding, which was supported by MST experiments as there was no significant change in CaM binding affinity between wildtype and the S179E/T180E phosphomutant. This supports the hypothesis that T180 is sterically inaccessible to CaMKK1 or CaMKK2 for phosphorylation until Ca^{2+}-CaM binds to and removes the regulatory domain from the kinase surface, revealing the active site.

5.2 Future Perspectives

In the future, the structural study of CaMK1D and its interaction partners aims to support drug discovery efforts that will eventually lead to a targeted, non-toxic treatment option for patients diagnosed with Triple Negative Breast Cancer. There is considerable untouched spaced in regard to understanding CaMK1D regulation and its role in TNBC. Phosphorylation of CaMK1D at sites other than T180 have recently become of interest, particularly T313 in the CBD which may be homologous to T305/306/307 of CaMKII. MST experiments have already begun to take place with the T313E/A mutants and are showing promising results to support this claim. However, NMR titration experiments showing that CaM does not bind to the CaMK1D CBD with this phosphorylation event would strengthen this argument as slow exchange kinetics were observed for both the wildtype and S179E/T180E isoforms. Moreover, it is thought that CaMK1D may act in a Ca^{2+}-CaM independent manner in its oncogenic form, which makes the CaMK1D

phosphorylation story particularly interesting. Collaborator Colin Kenyon is working on the identification of phosphorylation sites of CaMK1D and hopefully future experiments with phosphomimetic mutants will reveal a phosphorylation state that can support this hypothesis.

Though the interaction between CaM and CaMK1D has been made clearer through the work in this thesis, there is still only a low-resolution SAXS solution structure available. Techniques such as NMR and XRC can be used to develop structures with angstrom level resolution. So far, co-crystallization efforts of Ca^{2+}-CaM bound CaMK1D have resulted in poorly diffracting crystals that cannot be used for high resolution structure determination. Continued work needs to be done in this area to support the current SAXS structure. NMR has not, to this point, been used to determine a structure of the complex, however, data for these purposes have been collected and can be used in the future. NMR data may be difficult to work with in the case of this complex, as CaM displays dynamic properties, and this could convolute the spectra. For this reason, a crystal structure would be of highest priority to elucidate the Ca^{2+}-CaM/CaMK1D complex structure.

The overall goal of the CaMK1D project is to develop a therapeutic agent that inhibits CaMK1D to specifically treat TNBC. Drug discovery efforts are ongoing, and the University of Alberta group is responsible for the biophysical analysis of drug molecules to determine how and where they bind the kinase. To ensure that novel binding space is utilized, specific labeling techniques such as ILV labeling discussed in section 4.1.2 should be employed to identify, with well-resolved NMR spectra, where these potential drug molecules bind. This would be beneficial to pinpoint the molecular points of contact as the backbone does not always act as a good reporter to show what is occurring on the surface of the protein. The ^{13}C chemical shift perturbations of the Ile, Leu, and

Val side chains will be much more sensitive to small molecule binding, and better resolved than peaks in the fingerprint region of an HSQC. Furthermore, though NMR can reveal a rough affinity through line shape analysis of fragment titration experiments, other techniques such as Isothermal Calorimetry (ITC) and/or MST should be implemented. MST experiments are much more cost effective as the sample size is mere microliters as compared to the 1 mL standard chamber for ITC experiments, however, ITC experiments use unlabeled sample in their native states which may mimic the intracellular state much more closely.

Bibliography

BIBLIOGRAPHY

1. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298,** 1912–1934 (2002).

2. Bollen, M., Peti, W., Ragusa, M. J. & Beullens, M. The extended PP1 toolkit: designed to create specificity. *Trends Biochem. Sci.* **35,** 450–458 (2010).

3. Krebs, E. G. & Fischer, E. H. Phosphorylase activity of skeletal muscle extracts. *J. Biol. Chem.* **216,** 113–120 (1955).

4. Fischer, E. H. & Krebs, E. G. Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J. Biol. Chem.* **216,** 121–132 (1955).

5. Hayes, J. S. & Mayer, S. E. Regulation of guinea pig heart phosphorylase kinase by cAMP, protein kinase, and calcium. *Am. J. Physiol.-Endocrinol. Metab.* **240,** E340–E349 (1981).

6. E G Krebs & Beavo, and J. A. Phosphorylation-Dephosphorylation of Enzymes. *Annu. Rev. Biochem.* **48,** 923–959 (1979).

7. Cohen, P. The role of protein phosphorylation in neural and hormonal control of cellular activity. *Nature* **296,** 613–620 (1982).

8. Shchemelinin, I., Sefc, L. & Necas, E. Protein kinases, their function and implication in cancer and other diseases. *Folia Biol. (Praha)* **52,** 81–100 (2006).

9. Zheng, H. *et al.* Gain-of-function mutations in the gene encoding the tyrosine phosphatase SHP2 induce hydrocephalus in a catalytically dependent manner. *Sci. Signal.* **11,** (2018).

10. Knighton, D. R. *et al.* Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253,** 414–420 (1991).

11. Hanks, S. K., Quinn, A. M. & Hunter, T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241,** 42–52 (1988).

12. Guimarães, C. R. W. *et al.* Understanding the Impact of the P-loop Conformation on Kinase Selectivity. *J. Chem. Inf. Model.* **51,** 1199–1204 (2011).

13. Marin, O. *et al.* Tyrosine Versus Serine/Threonine Phosphorylation by Protein Kinase Casein Kinase-2 A STUDY WITH PEPTIDE SUBSTRATES DERIVED FROM IMMUNOPHILIN Fpr3. *J. Biol. Chem.* **274,** 29260–29265 (1999).

14. Wood, L. D. *et al.* The genomic landscapes of human breast and colorectal cancers. *Science* **318,** 1108–1113 (2007).

15. Buchdunger, E., Matter, A. & Druker, B. J. Bcr-Abl inhibition as a modality of CML therapeutics. *Biochim. Biophys. Acta* **1551,** M11-18 (2001).

16. Druker, B. J. & Lydon, N. B. Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J. Clin. Invest.* **105,** 3–7 (2000).

17. Vieth, M., Sutherland, J. J., Robertson, D. H. & Campbell, R. M. Kinomics: characterizing the therapeutically validated kinase space. *Drug Discov. Today* **10,** 839–846 (2005).

18. Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today* **11,** 580–594 (2006).

19. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144,** 646–674 (2011).

20. Ferlay, J. *et al.* Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *Eur. J. Cancer* **49,** 1374–1403 (2013).

21. Inwald, E. C. *et al.* Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry. *Breast Cancer Res. Treat.* **139,** 539–552 (2013).

22. Dai, X. *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **5,** 2929–2943 (2015).

23. Dai, X., Chen, A. & Bai, Z. Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling. *Sci. Rep.* **4,** 6566 (2014).

24. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406,** 747–752 (2000).

25. Sotiriou, C. *et al.* Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 10393–10398 (2003).

26. Loo, L. W. *et al.* Genome-wide Copy Number Alterations in Subtypes of Invasive Breast Cancers in Young White and African American Women. *Breast Cancer Res. Treat.* **127,** 297–308 (2011).

27. Carey, L. A. *et al.* Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* **295,** 2492–2502 (2006).

28. Hu, M. *et al.* Distinct epigenetic changes in the stromal cells of breast cancers. *Nat. Genet.* **37,** 899–905 (2005).

29. Allinen, M. *et al.* Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* **6,** 17–32 (2004).

30. Lerwill, M. F. Current practical applications of diagnostic immunohistochemistry in breast pathology. *Am. J. Surg. Pathol.* **28,** 1076–1091 (2004).

31. Polyak, K. Breast cancer: origins and evolution. *J. Clin. Invest.* **117,** 3155–3163 (2007).

32. Cui, C., Merritt, R., Fu, L. & Pan, Z. Targeting calcium signaling in cancer therapy. *Acta Pharm. Sin. B* **7,** 3–17 (2017).

33. Zhu, H. *et al.* Elevated Orai1 expression mediates tumor-promoting intracellular Ca2+ oscillations in human esophageal squamous cell carcinoma. *Oncotarget* **5,** 3455–3471 (2014).

34. Parekh, A. B. Decoding cytosolic Ca2+ oscillations. *Trends Biochem. Sci.* **36,** 78–87 (2011).

35. Berridge, M. J. The AM and FM of calcium signalling. *Nature* **386,** 759–760 (1997).

36. Zlotnik, A. Involvement of chemokine receptors in organ-specific metastasis. *Contrib. Microbiol.* **13,** 191–199 (2006).

37. Azimi, I., Roberts-Thomson, S. J. & Monteith, G. R. Calcium influx pathways in breast cancer: opportunities for pharmacological intervention. *Br. J. Pharmacol.* **171,** 945–960 (2014).

38. Swulius, M. T. & Waxham, M. N. Ca2+/Calmodulin-dependent Protein Kinases. *Cell. Mol. Life Sci. CMLS* **65,** 2637–2657 (2008).

39. Hook, S. S. & Means, A. R. Ca2+/CaM-Dependent Kinases: From Activation to Function. *Annu. Rev. Pharmacol. Toxicol.* **41,** 471–505 (2001).

40. Soderling, T. R. & Stull, J. T. Structure and Regulation of Calcium/Calmodulin-Dependent Protein Kinases. *Chem. Rev.* **101,** 2341–2352 (2001).

41. Ma, H. *et al.* γCaMKII shuttles Ca2+/CaM to the nucleus to trigger CREB phosphorylation and gene expression. *Cell* **159,** 281–294 (2014).

42. Bergamaschi, A. *et al.* CAMK1D amplification implicated in epithelial–mesenchymal transition in basal-like breast cancer. *Mol. Oncol.* **2,** 327–339 (2008).

43. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486,** 346–352 (2012).

44. Haney, S. *et al.* RNAi Screening in Primary Human Hepatocytes of Genes Implicated in Genome-Wide Association Studies for Roles in Type 2 Diabetes Identifies Roles for CAMK1D and CDKAL1, among Others, in Hepatic Glucose Regulation. *PLoS ONE* **8,** e64946 (2013).

45. Lawson, J. *et al.* Selective secretion of microRNAs from lung cancer cells via extracellular vesicles promotes CAMK1D-mediated tube formation in endothelial cells. *Oncotarget* **8,** 83913–83924 (2017).

46. Imamura, M. *et al.* Genetic variants at CDC123/CAMK1D and SPRY2 are associated with susceptibility to type 2 diabetes in the Japanese population. *Diabetologia* **54,** 3071–3077 (2011).

47. Gamboa-Meléndez, M. A. *et al.* Contribution of common genetic variation to the risk of type 2 diabetes in the Mexican Mestizo population. *Diabetes* **61,** 3314–3321 (2012).

48. Liang, X. *et al.* Effect of heterogeneity on the chromosome 10 risk in late-onset Alzheimer disease. *Hum. Mutat.* **28,** 1065–1073 (2007).

49. Desikan, R. S. *et al.* Polygenic Overlap Between C-Reactive Protein, Plasma Lipids, and Alzheimer Disease. *Circulation* **131,** 2061–2069 (2015).

50. Müller, M., Cárdenas, C., Mei, L., Cheung, K.-H. & Foskett, J. K. Constitutive cAMP response element binding protein (CREB) activation by Alzheimer's disease presenilin-driven inositol trisphosphate receptor (InsP3R) Ca2+ signaling. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 13293–13298 (2011).

51. Yamauchi, T. Neuronal Ca2+/calmodulin-dependent protein kinase II--discovery, progress in a quarter of a century, and perspective: implication for learning and memory. *Biol. Pharm. Bull.* **28,** 1342–1354 (2005).

52. Kuret, J. *et al.* Pathways of tau fibrillization. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1739,** 167–178 (2005).

53. Verploegen, S., Lammers, J. W., Koenderman, L. & Coffer, P. J. Identification and characterization of CKLiK, a novel granulocyte Ca(++)/calmodulin-dependent kinase. *Blood* **96,** 3215–3223 (2000).

54. Verploegen, S. *et al.* Characterization of the role of CaMKI-like kinase (CKLiK) in human granulocyte function. *Blood* **106,** 1076–1083 (2005).

55. Ishikawa, Y. *et al.* Identification and characterization of novel components of a Ca2+/calmodulin-dependent protein kinase cascade in HeLa cells. *FEBS Lett.* **550,** 57–63 (2003).

56. Swulius, M. T. & Waxham, M. N. Ca2+/Calmodulin-dependent Protein Kinases. *Cell. Mol. Life Sci. CMLS* **65,** 2637–2657 (2008).

57. Sakagami, H. *et al.* Prominent expression and activity-dependent nuclear translocation of Ca2+/calmodulin-dependent protein kinase Idelta in hippocampal neurons. *Eur. J. Neurosci.* **22,** 2697–2707 (2005).

58. Qin, X. *et al.* Eukaryotic initiation factor 5A and Ca2+ /calmodulin-dependent protein kinase 1D modulate trophoblast cell function. *Am. J. Reprod. Immunol. N. Y. N 1989* **80,** e12845 (2018).

59. Kahl, C. R. & Means, A. R. Regulation of cell cycle progression by calcium/calmodulin-dependent pathways. *Endocr. Rev.* **24,** 719–736 (2003).

60. Crotti, L. *et al.* Calmodulin mutations associated with recurrent cardiac arrest in infants. *Circulation* **127,** 1009–1017 (2013).

61. Esteras, N. *et al.* Calmodulin levels in blood cells as a potential biomarker of Alzheimer's disease. *Alzheimers Res. Ther.* **5,** 55 (2013).

62. Babu, Y. S., Bugg, C. E. & Cook, W. J. Structure of calmodulin refined at 2.2 A resolution. *J. Mol. Biol.* **204,** 191–204 (1988).

63. Kretsinger, R. H. The linker of calmodulin — to helix or not to helix. *Cell Calcium* **13,** 363–376 (1992).

64. Wüthrich, K. NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture). *Angew. Chem. Int. Ed.* **42,** 3340–3363 (2003).

65. Wishart, D. S., Sykes, B. D. & Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31,** 1647–1651 (1992).

66. Neuhaus, D. Nuclear Overhauser Effect. in *eMagRes* (American Cancer Society, 2011). doi:10.1002/9780470034590.emrstm0350.pub2

67. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73,** 1–16 (2013).

68. Svergun, D. I. & Koch, M. H. J. Advances in structure analysis using small-angle scattering in solution. *Curr. Opin. Struct. Biol.* **12,** 654–660 (2002).

69. Patel, T. R. *et al.* Structural elucidation of full-length nidogen and the laminin–nidogen complex in solution. *Matrix Biol.* **33,** 60–67 (2014).

70. Kozak, S. *et al.* Optimization of protein samples for NMR using thermal shift assays. *J. Biomol. Nmr* **64,** 281–289 (2016).

71. Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357,** 289–298 (2006).

72. Raines, R. T., McCormick, M., Van Oosbree, T. R. & Mierendorf, R. C. The S.Tag fusion system for protein purification. *Methods Enzymol.* **326,** 362–376 (2000).

73. Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **1,** 100 (2010).

74. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50,** 1212–1225 (2017).

75. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36,** 1277–1282 (2003).

76. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496,** 477–481 (2013).

77. Tuukkanen, A. T. & Svergun, D. I. Weak protein–ligand interactions studied by small-angle X-ray scattering. *FEBS J.* **281,** 1974–1987

78. Koch, M. H., Vachette, P. & Svergun, D. I. Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q. Rev. Biophys.* **36,** 147–227 (2003).

79. Petoukhov, M. V. & Svergun, D. I. Applications of small-angle X-ray scattering to biomacromolecular solutions. *Int. J. Biochem. Cell Biol.* **45,** 429–437 (2013).

80. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40,** 191–285 (2007).

81. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76,** 2879–2886 (1999).

82. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42,** 342–346 (2009).

83. Volkov, V. V., Svergun, D. I. & IUCr. Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* (2003). doi:10.1107/S0021889803000268

84. UCSF Chimera—A visualization system for exploratory research and analysis - Pettersen - 2004 - Journal of Computational Chemistry - Wiley Online Library. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20084. (Accessed: 26th June 2018)

85. Meyer, B. & Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew. Chem. Int. Ed.* **42,** 864–890

86. Medek, A., Hajduk, P. J., Mack, J. & Fesik, S. W. The Use of Differential Chemical Shifts for Determining the Binding Site Location and Orientation of Protein-Bound Ligands. *J. Am. Chem. Soc.* **122,** 1241–1242 (2000).

87. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. *J. Am. Chem. Soc.* **125,** 1731–1737 (2003).

88. Tong, M. Evaluation of protein kinases for solution NMR spectroscopy and the structural mechanism of inhibition and activation of an oncogenic calcium calmodulin dependent protein kinase. (University of Birmingham, 2012).

89. Larsson, G. *et al.* A novel target recognition revealed by calmodulin in complex with the basic helix--loop--helix transcription factor SEF2-1/E2-2. *Protein Sci. Publ. Protein Soc.* **10,** 169–186 (2001).

90. Lindman, S. *et al.* Salting the Charged Surface: pH and Salt Dependence of Protein G B1 Stability. *Biophys. J.* **90,** 2911–2921 (2006).

91. Masterson, L. R. *et al.* Backbone NMR resonance assignment of the catalytic subunit of cAMP-dependent protein kinase A in complex with AMP-PNP. *Biomol. NMR Assign.* **3,** 115–117 (2009).

92. Tong, M. *et al.* Survey of solution dynamics in Src kinase reveals allosteric cross talk between the ligand binding and regulatory sites. *Nat. Commun.* **8,** 2160 (2017).

93. Bunney, T. D. *et al.* Disease Variants of FGFR3 Reveal Molecular Basis for the Recognition and Additional Roles for Cdc37 in Hsp90 Chaperone System. *Struct. Lond. Engl. 1993* **26,** 446-458.e8 (2018).

94. Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C. & Kay, L. E. A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. *J. Biomol. NMR* **13,** 369–374 (1999).

95. Xiao, Y., Warner, L. R., Latham, M. P., Ahn, N. G. & Pardi, A. Structure-Based Assignment of Ile, Leu, and Val Methyl Groups in the Active and Inactive Forms of the Mitogen-Activated Protein Kinase Extracellular Signal-Regulated Kinase 2. *Biochemistry* **54,** 4307–4319 (2015).

96. Modell, A. E., Blosser, S. L. & Arora, P. S. Systematic Targeting of Protein-Protein Interactions. *Trends Pharmacol. Sci.* **37,** 702–713 (2016).

97. Clapperton, J. A., Martin, S. R., Smerdon, S. J., Gamblin, S. J. & Bayley, P. M. Structure of the Complex of Calmodulin with the Target Sequence of Calmodulin-Dependent Protein Kinase I: Studies of the Kinase Activation Mechanism. *Biochemistry* **41,** 14669–14679 (2002).

98. Yuan, T., Walsh, M. P., Sutherland, C., Fabian, H. & Vogel, H. J. Calcium-Dependent and - Independent Interactions of the Calmodulin-Binding Domain of Cyclic Nucleotide Phosphodiesterase with Calmodulin. *Biochemistry* **38,** 1446–1455 (1999).

99. Evans, T. I. A. & Shea, M. A. Energetics of calmodulin domain interactions with the calmodulin binding domain of CaMKII. *Proteins* **76,** 47–61 (2009).

100. Quintana, A. R., Wang, D., Forbes, J. E. & Waxham, M. N. Kinetics of calmodulin binding to calcineurin. *Biochem. Biophys. Res. Commun.* **334,** 674–680 (2005).

101. Bayley, P. M., Findlay, W. A. & Martin, S. R. Target recognition by calmodulin: dissecting the kinetics and affinity of interaction using short peptide sequences. *Protein Sci. Publ. Protein Soc.* **5,** 1215–1228 (1996).

102. Seeger, C., Talibov, V. O. & Danielson, U. H. Biophysical analysis of the dynamics of calmodulin interactions with neurogranin and Ca2+/calmodulin-dependent kinase II. *J. Mol. Recognit.* **30,** (2017).

103. Tidow, H. & Nissen, P. Structural diversity of calmodulin binding to its target sites. *FEBS J.* **280,** 5551–5565 (2013).

104. de Diego, I., Kuper, J., Bakalova, N., Kursula, P. & Wilmanns, M. Molecular Basis of the Death-Associated Protein Kinase-Calcium/Calmodulin Regulator Complex. *Sci. Signal.* **3,** ra6–ra6 (2010).

105. Rellos, P. *et al.* Structure of the CaMKIIdelta/calmodulin complex reveals the molecular mechanism of CaMKII kinase activation. *PLoS Biol.* **8,** e1000426 (2010).

106. Bertonati, C., Honig, B. & Alexov, E. Poisson-Boltzmann Calculations of Nonspecific Salt Effects on Protein-Protein Binding Free Energies. *Biophys. J.* **92,** 1891–1899 (2007).

107. McLaughlin, S., Hangyás-Mihályné, G., Zaitseva, I. & Golebiewska, U. Reversible - through calmodulin - electrostatic interactions between basic residues on proteins and acidic lipids in the plasma membrane. *Biochem. Soc. Symp.* 189–198 (2005).

108. Wang, X. & Putkey, J. A. PEP-19 modulates calcium binding to calmodulin by electrostatic steering. *Nat. Commun.* **7,** 13583 (2016).

109. Liu, F., Chu, X., Lu, H. P. & Wang, J. Molecular mechanism of multispecific recognition of Calmodulin through conformational changes. *Proc. Natl. Acad. Sci.* **114,** E3927–E3934 (2017).

110. André, I., Kesvatera, T., Jönsson, B., Akerfeldt, K. S. & Linse, S. The role of electrostatic interactions in calmodulin-peptide complex formation. *Biophys. J.* **87,** 1929–1938 (2004).

111. McInnes, C., Andrews, M. J. I., Zheleva, D. I., Lane, D. P. & Fischer, P. M. Peptidomimetic design of CDK inhibitors targeting the recruitment site of the cyclin subunit. *Curr. Med. Chem. Anti-Cancer Agents* **3,** 57–69 (2003).

112. Gumireddy, K. *et al.* A non-ATP-competitive inhibitor of BCR-ABL overrides imatinib resistance. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 1992–1997 (2005).