Evaluation of Protein Kinases for Solution NMR Spectroscopy and the Structural Mechanism of Inhibition and Activation of an Oncogenic Calcium Calmodulin Dependent Protein Kinase

By
Michael Tong

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

School of Cancer Sciences

University of Birmingham

May 2011

Abstract

Protein kinases are important mediators of cell signalling that are often implicated in disease when deregulation occurs. The catalytic kinase domain is highly conserved with 518 identified members in the super family. Kinase structural studies are mainly based on relatively static crystal structures. However protein kinases are inherently dynamic entities in solution. Several Ser/Thr protein kinases were evaluated by NMR in order to obtain an amenable target for solution structure and functional characterisation. Subsequently a calcium calmodulin dependent protein kinase dubbed CaMK1D was identified as the optimal system. CaMK1D normally mediates intracellular signalling downstream of chemokines. It is amplified in breast cancer, and induces cell proliferation, migration and invasion. Here we report the backbone resonance assignments for the 38 kDa human autoinhibited CaMK1D in its free state, encompassing a canonical bi-lobed kinase fold and autoinhibitory and calmodulin binding domains. These assignments allowed us to probe the binding mode of CaMK1D with small molecule ligands and refine the crystal structure via dihedral angle restraints for a more complete structure. Furthermore we investigated the solution structure of the CaMK1D·Ca^{2+}/CaM complex and propose a model of the activation mechanism and establish a key residue implicated in complex formation.

Acknowledgments

" Endurance is one of the most difficult disciplines, but it is to the one who endures that the final victory comes. "

-Buddha

Table of contents

Table of contents

Table of contents

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,	Threonine
Trp, W,	Tryptophan
Tyr, Y,	Tyrosine
Val, V,	Val

Abbreviations

Kinase Groups

AGC

Named after cAMP dependent kinase, cGMP dependent kinase, and protein kinase C families. This group includes kinases that are typically modulated by cyclic nucleotides, phospholipids, and calcium.

Atypical

Named atypical because this group represents kinases which lack sequence similarity to the eukaryotic protein kinase domain but exhibit kinase actitivity.

CAMK

Named after the Ca^{2+}/CaM dependent kinase family. Consists of kinases that are regulated by Ca^{2+}/CaM dependent but also those that not.

CK1

Named after cell kinase 1, although originally named after casein kinase 1. This is a small ancient group of kinases that are highly similar to each other but are the most divergent of the groups related to the eukaryotic protein KD.

CMGC

Named after the cyclin dependent protein kinase, mitogen dependent kinase, gylocgen synthase kinase, and cyclin dependent like kinase families. This group includes kinases involved in cell cycle control, MAPK signalling, metabolic control splicing, stress response and other unknown functions.

Other

This group is named other because it consists of several kinase families that include unique kinases which are have sequence similarity to the conserved eukaryotic protein KD but do not fit into any of the other groups.

PKL

Named after protein kinase-like because these kinases have a protein kinase-like fold and catalytic mechanism but lack substantial sequence similarity.

Abbreviations

RGC	Named after Receptor guanylate cyclises. This group consists of kinases that feature an active guanylate cyclase domain that produces cGMP. Furthermore these kinases have a catalytically inactive KD. However the KD appears to have a regulatory function instead.
STE	Named after homologs of yeast sterile 7, sterile 11, sterile 20 genes. These kinases are involved in the MAPK cascade.
TK	Named after tyrosine kinase, because this group consists of families of kinases that phosphorylate almost exclusively on tyrosine residues.
TKL	Named after tyrosine kinase-like because this group is most similar to the TK group but instead typically phosphorylate on Ser/Thr residues.

Domains, Genes, Motifs and Proteins

AD	Association domain
AID	Autoinhibitory domain
APE motif	Alanine, Proline, Glutamic acid, motif
A-loop	Activation loop
CAAX motif	Cysteine, Alanine, Alanine, any amino acid, motif
CaM	Calmodulin
CAMK/CaMK	Ca^{2+}/CaM dependent protein kinase
CaMK1A	Calcium/calmodulin dependent protein kinase α
CaMK1B/Pnck	Calcium/calmodulin dependent protein kinase β/pregnancy up-regulated non-uqbuitiously expressed CaM kinase

Abbreviations

CaMK1G/CLICK-III	Calcium/calmodulin dependent protein kinase γ/ CaMK-like CREB like kinase 3
CaMK1D	Calcium/calmodulin dependent protein kinase δ
CaMKK	Calcium/calmodulin dependent protein kinase kinase
CBD	Calmodulin binding domain
CDK	Cyclin-dependent kinases
CK1	Cell kinase 1
CKLiK	Calcium/calmodulin dependent protein kinase kinase
CLK	Cyclin-dependent-like kinases
C-lobe	C-terminal lobe
CREB	cAMP response binding protein 1
CRIB	Cdc42/Rac1 interactive binding domain
CSNK1	Casein kinase 1
CSNK1G3	Casein kinase 1 γ isoform 3
DFG motif	Aspartic acid, Phenylalanine, Glycine, motif
GSK	Glycogen synthase kinases
KD	Kinase domain
MAPK	Mitogen-dependent kinase or mitogen activated protein kinase
N-lobe	N-terminal lobe
PAK4	p21-activated kinase 4
PAK5	p21-activated kinase 5
PBD	p21-binding domain
PKA	protein kinase A or cAMP dependent kinase

Abbreviations

PKB	protein kinase B
PKC/Akt	protein kinase C/retovirus strain Ak thymoma
PKG	Protein kinase G or cGMP dependent kinase
Pim1	Proviral insertion site of Maloney murine virus leukemia induced 1
p-loop	Phosphate binding loop/Gly rich loop
RD	Regulatory domain
RGC/GYC	Receptor guanylate cyclases
TK	Tyrosine kinase
TKL	Tyrosine kinase-like
TOPK/PBK	T-cell originated protein kinase/PDZ binding kinase
STE	Homologs of yeast sterile 7, sterile 11, sterile 20 genes
VRK	Vaccinia related kinase

Biological metabolites and reagents

ADP	Adenosine 5'-diphosphate
AMP	Adenosine 5'-monophosphate
AMPPNP	Adenosine 5'-(β,γ-imido)triphosphate
ATP	Adenosine 5'-triphosphate
BIM-1	Bisindolylmaleimide I, hydrochloride
	2-[1-(3-Dimethylaminopropyl)-1H-indol-3-yl]-3-(1H-indol-3-yl)maleimide, HCl
cAMP	cyclic adenosine monophosphate
cGMP	cyclic guanosine monophosphate

Abbreviations

DMSO   Dimethly sulfoxide

GSK-3 XIII   Glycogen synthase kinase 3 XIII inhibitor

              (5-methyl-1H-pyrazol-3-yl)-(2-phenylquinazolin-4-yl)amine

GTP   Guanosine triphosphate

^2H_2O Deuterium oxide

HEPES   4-(2-hydroxethyl)piperazine-1-ethanesulfonic acid

IPTG   Isopropyl-β-D-1-thiogalactopyranoside

LB   Luria broth

MES   2-(N-morpholino)ethanesulfonic acid

MOPS   3-(N-morpholino)propanesulfonic acid

Na_2HPO_4 Sodium phosphate

NaN_3 Sodium azide

TCEP   Tris(2-caroboxyethyl)phosphine hydrochloride

TMAO   Trimethylamine N-oxide

Tris   2-Amino-2-(hydroxymethyl)-1,3-propanediol

Methods

BLAST   Basic alignment search tool

HSQC   Heteronuclear single quantum coherence

ITC   Isothermal titration calorimetry

RT-PCR   Real time polymerase chain reaction

SAXS   Small angle x-ray scattering

SEC   Size exclusion chromatography

Abbreviations

SDM   Site directed mutagenesis

SDS PAGE   Sodium dodecyl sulfate polyacrylamide gel electrophoresis

TROSY   Transverse Relaxation Optimised Spectroscopy

SDM   Site directed mutagenesis

List of Figures

List of Figures

List of Figures

List of Figures

Chapter I Introduction

Objectives

Protein function is governed by its structure and dynamics, all these phenomena are inherently connected. Protein kinases can interconvert between multiple states reflecting distinct conformations along their reaction trajectory. Despite a highly conserved ATP binding site and global protein kinase domain fold, specific inhibition and differential modes of regulation and activation has been demonstrated across many protein kinases. Due to their dynamic behaviour often a ligand is required to stabiliize a particular conformation inorder to promote crystallization for structural study by protein crystallography. In collaboration with the Structural Genomics Consortium (SGC) the aim of this thesis is to determine the structure, mechanism of regulation, activation, and small molecule ligand mediated inhibition of a protein by solution NMR. The initial target dubbed TOPK/PBK, a cancer relevant protein kinase that has to date proven to be unreceptive to crystallisation. Structure determination of TOPK/PBK will help to elucidate the mechanism of its regulation and activation because a high throughput screen against a library of kinase ligands conducted by the SGC did not indicate any binders suggesting an unknown mode of autoinhibition. The requirements for activation and relief of autoinhibition will be investigated inaddition to the discovery and characterisation of ligand binding sites that modulate activity. In a wider context, a kinase by NMR approach will complement the wealth of data available on kinase crystal structures by yielding structural, functional, binding, and dynamic information in solution. Furthermore an advantage of NMR over crystallography is its application as a ligand screening tool that can readily identify low affinity substrates, inhibitor fragments, and novel allosteric binding sites, thus providing a foundation for the rational design of more specific inhibitors.

In parallel with TOPK/PBK, several different Ser/Thr protein kinase clones whose structures have been solved by the SGC will be evaluated to establish their amenability to solution NMR. Here a similar general principle of study to TOPK/PBK will be applied in terms of characterising the structure, its mechanism of regulation and activation, and interactions in solution. Since crystal structures of these

alternative kinases are already available only a backbone solution structure will be necessary initially as a basis for solution NMR interaction studies.

The general criteria for an optimised kinase target includes: expression of iostopically labeled soluble monomeric folded protein at high yields (typically 5-10 mg quantites or more of protein), long term solubility and stability of atleast several days to one week preferably at room temperature or higher ($\geq$25°C) for triple resonance data acquisition. Spectral quality must also be met sufficiently in terms of high peak intensity, peak count, and resolution. Evaluating a kinase that satisfies these criteria will likely require optimisation of its solution conditions and construct design.

Once optimal conditions are established, appropriate isotope labelling methods will be implemented in order to acquire relevant experiments for straightforward sequential backbone assignment. As a minimum requirement the backbone assignment of a protein kinase will be necessary in order to characterise its backbone solution structure which can then be used as a reference spectrum for binding studies and probing its regulation and activation. Perturbation of the reference state via interactions with ligands, proteins, or mutagenesis will be monitored in solution by ^{1}H-^{15}N HSQC TROSY experiments.

Phosphorylation and the kinase superfamily

Protein kinases by definition are enzymes which modify other proteins through covalent attachment of a gamma phosphate group from adenosine triphosphate (ATP) to the hydroxyl group on a protein substrate. This post translational modification is described as protein phosphorylation and is reversible by dephosphorylation, a separate, counteractive, process mediated by protein phosphatases (Figure 1). Studies of kinase mediated phosphorylation can be traced back to 1954, when casein was found to be phosphorylated by a liver enzyme (Burnett and Kennedy 1954; Cohen 2002). Shortly after in 1955, it became evident that phosphorylation and dephosphorylation were implicated in the interconversion of glycogen phosphorylase b and glycogen phosphorylase a and was dependent on ATP, Mg^{2+} and an enzyme termed phosphorylase kinase (Fischer and Krebs 1955; Krebs and Fischer 1956; Fischer, Graves et al. 1959; Cohen 2002). During this period it was incorrectly thought that phosphorylation was limited to the regulation of glycogenolysis. However, studies on protein kinase A (PKA) and the discovery of various phosphorylated substrates including both enzymes and proteins in the 1970s and 1980s revealed that phosphorylation was actually a widespread regulatory mechanism that affected many cellular processes (Langan 1969; Linn, Pettit et al. 1969; Cohen 2002).

Research on protein kinases has advanced progressively since the 1950s such that protein phosphorylation is now recognised as a major signalling and regulatory mechanism that can modulate protein activity, stability, protein-protein interactions and cellular localisation. Phosphorylation occurs extensively throughout the cell with approximately 30 % of all human proteins being phosphorylated (Cohen 2001). Regulated kinase activity is therefore critical for normal cellular function because kinases can orchestrate many diverse cellular processes such as cell proliferation, gene expression, apoptosis and metabolism. Therefore deregulation of kinase activity often contributes to cancer and many other diseases and is a major focus of drug discovery efforts through the desgine of kinase inhibitors (Cohen 2002; Vlahovic and Crawford 2003; Melnikova and Golden 2004; Paul and Mukhopadhyay 2004; Baselga 2006; Fedorov, Marsden et al. 2007; Marsden and Knapp 2008).
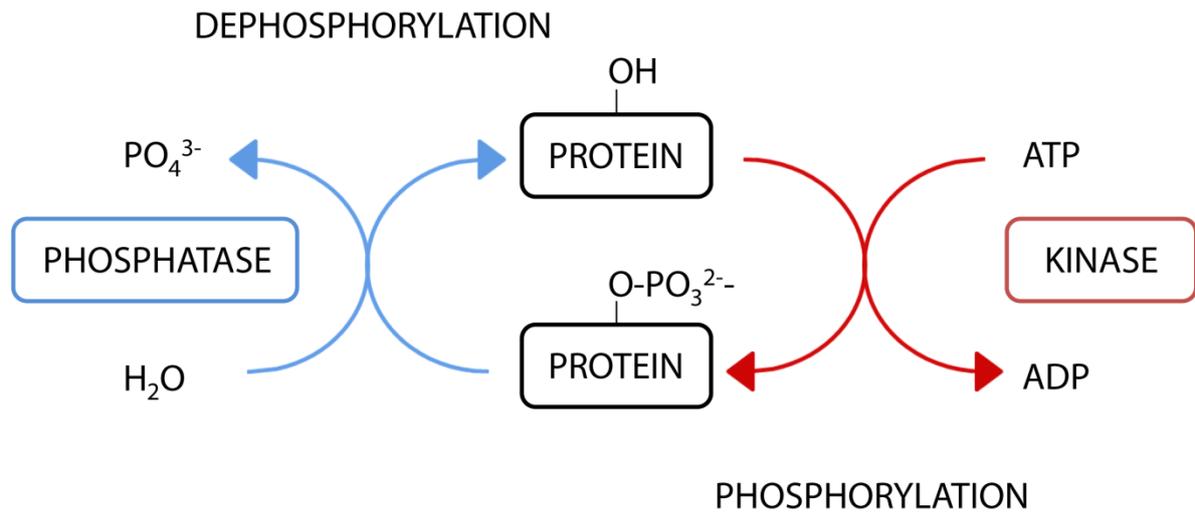
Figure 1.0. Reversible phosphorylation.

Protein kinases catalyse the phosphorylation of substrate proteins by transferring the gamma phosphate group from ATP to the hydroxyl group of the substrate protein whilst generating ADP in the process. The reaction usually requires ATP in complex with Mg^{2+} although certain kinases have evolved to acquire Mg^{2+} independent kinase activity (Mukherjee, Sharma et al. 2008). Protein phosphatases catalyse the dephosphorylation of phosphorylated proteins via hydrolysis of the phosphoric acid monoester bond to produce the hydroxyl group of the substrate protein and a free phosphate ion.

Currently, 518 protein kinase members have been identified in the human kinase superfamily dubbed the human kinome, and based on their ability to phosphorylate Ser/Thr or Tyr substrates they can be broadly sub divided into two main catagories: protein Ser/Thr kinases and the protein Tyr kinases. These can be further classified into ten groups with distinct sequences based on sequence alignments and phylogenetic trees using the conserved eukaryotic protein kinase domain (KD) as a reference whilst also supported by knowledge of sequence similarity and domain structure outside of the KD, known biological function, and the classification of yeast, worm, and fly kinomes (Manning, Whyte et al. 2002). Nine of the defined groups including: AGC, CAMK, CK1, CMGC, STE, other, TK, TKL, and RGC/GYC, classify the 478 kinases that are related to the eukaryotic protein KD (Figure 3). The 40 kinases which lack sequence similarity to the KD but show kinase activity and structural similarity are classified under the atypical group. Most of the groups are further subdivided into a heirachy of families and subfamilies (Manning, Whyte et al. 2002).

The human kinome tree (dendrogram of protein kinase families): TK, TKL, STE, CK1, AGC, CAMK, CMGC, and GYC groups.
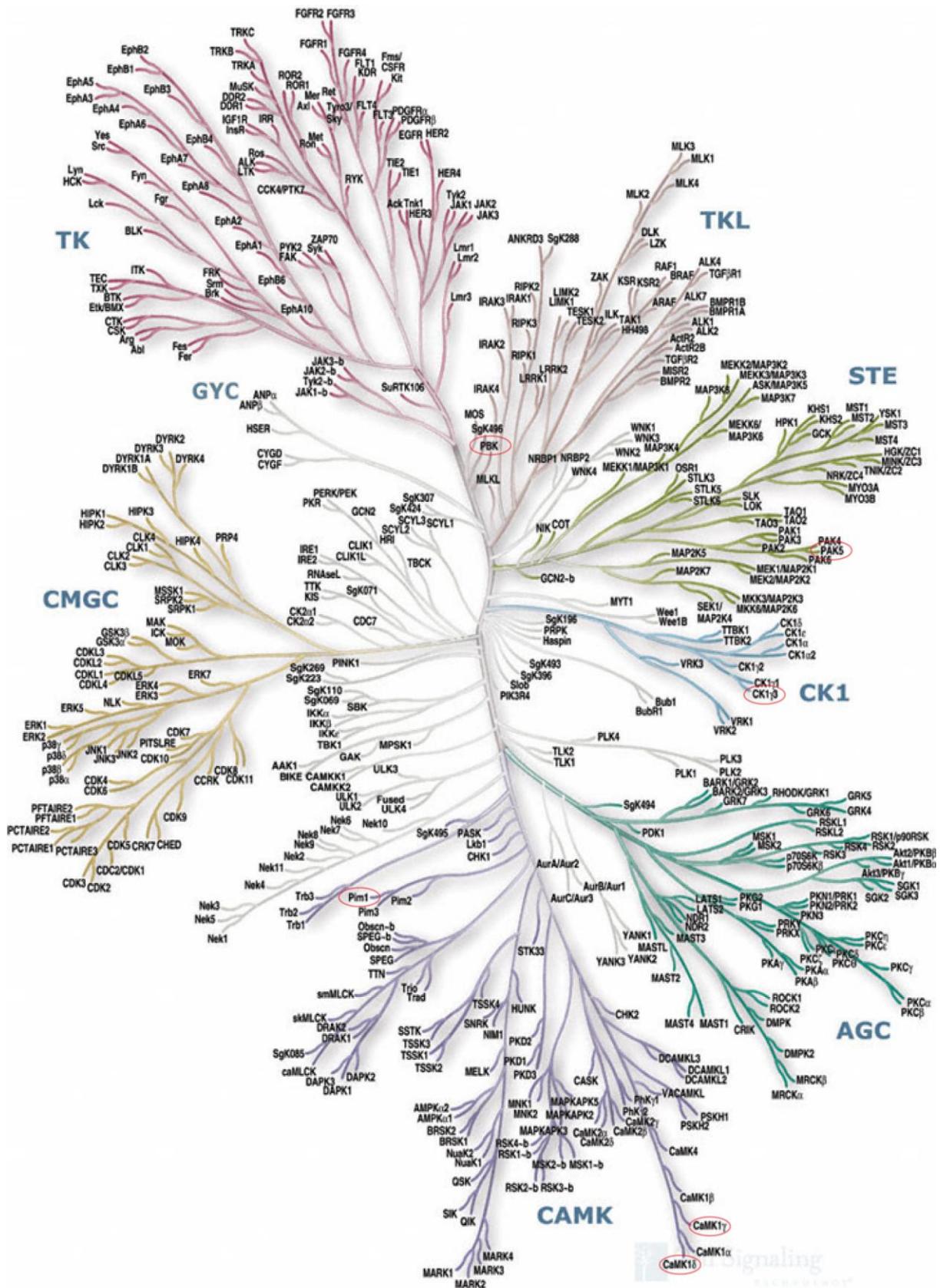
Figure 1.1. Human kinome.

A dendogram of the human kinase superfamily illustrates the classification and evolutionary relationship of 478 protein kinases which show homology to the conserved eukaryotic protein KD. Classification for these particular protein kinases is defined by nine major groups: AGC, CAMK, CK1, CMGC, other RGC/GYC, STE, TK and TKL. Only the other group is not labelled but is represented by branches separate from the main branch of the other eight groups. Here distances along the branch between two kinases is proportional to the divergence between their sequences. The atypical group, representing 40 divergent human protein kinases lacking sequence homology to the conserved eukaryotic KD is not shown. Kinases that were evaluated for NMR studies in this thesis are circled in red. This figure was taken from cellsignal.com.

The Protein Kinase Domain

The structure of the first KD to be determined was murine PKA (Knighton, Zheng et al. 1991; Knighton, Zheng et al. 1991), revealing a bilobal fold that comprised of a small ATP binding β-sheet N-terminal lobe (N-lobe) and a large substrate binding α-helical C-terminal lobe (C-lobe). Both lobes are linked together by a short loop dubbed the hinge which also allows inter lobe movment (Figure 1.2). Furthermore it revealed the structure of a protein kinase in an active conformation poised function of the many key conserved residues and motifs first identified and described by Hanks et al in their kinase classification (Hanks, Quinn et al. 1988; Hanks and Hunter 1995). These functional motifs implicated in binding and catalysis of ATP-Mg^{2+} are described below

Within the N-lobe there is a conserved motif, known as the p-loop or Gly rich loop whose consensus sequence is: GXGXXG. This loop behaves as a flexible clamp at the front of the active site cleft to cover and anchor the phosphates of ATP serving to position the nucleotide for catalysis. Capable of adopting open and closed conformations the distance of the p-loop relative to the catalytic loop has been described as a way to gauge the open and closed conformations of the loop. For example in the presence of nucleotide the tip of the p-loop is brought closer to the catalytic loop indicative of a closed conformation but in the absence of nucleotide the p-loop is further away and often disordered.

A pair of invariant residues found in the N-lobe of most kinases is the catalytic Lys and αC helix associated Glu, together these residues form a salt bridge that is an archetypal hallmark for the kinase active conformation. The catalytic Lys interacts with the α and β phosphates of ATP, helping to neutralise these negative charges, but it is only optimially configured to achieve this when the αC helix Glu forms the conserved salt bridge with the catalytic Lys. The αC helix is therefore a key regulatory element in many kinases because it can adopt an active in inactive conformation (in which the αC helix Glu is oritentated away from the catalytic Lys

hence kinase activation will require rearrangement of the αC helix (Huse and Kuriyan 2002).

In the upper region of the C-lobe several important loops are found that include the catalytic loop described by the consensus sequence: DXXXXN. Here the conserved Asp acts as a catalytic base via its carboxylic acid side chain nucleophilic attack of the substrate hydroxyl group. Another significant region that encompasses several motifs is the activation segment which includes the DFG, activation loop, and P+1 loop. The magnesium coordinating DFG, in particular its conformation is critical for catalytic activity whereas the activation loop is often disordered but typically features a phosphoryalation site which can stabilize the segment. Phosphoryaltion promotes kinase activation by coordinating several key catalytic residues into their catalytically competetant conformation (Figure 1.3). At the C-terminus of this segment is the P+1 loop, a short stretch of residues that contribute towards substrate binding near the phosphoylation site.

Situated between the lobes is the active site cleft also known as the catalytic cleft. Here the native phosphate donating substrate ATP-Mg^{2+} binds. Typically the catalytic cleft can be described in terms of six main features known as the hydrophobic pocket, hydrophobic channel, and adenosine, sugar, hinge and phosphate binding regions (Figure 1.4). ATP-Mg^{2+} binds with its adenosine moiety facing towards the hinge and engages in hydrogen bonding with this flexible loop whilst the ribose and phosphate moieties are orientated into their respective sugar and phosphate regions of the cleft (Fabbro, Ruetz et al. 2002).

Active and inactive conformations

Kinases can adopt multiple conformations but are often described simply as inactive or active depending on the hallmarks of the DFG conformation and catalytic Lys αC helix Glu salt bridge formation. Another model has been proposed to describe the active and inactive conformation interms of their dynamic assembly of internal hydrophobic spines (Taylor and Kornev ; Kornev, Taylor et al. 2008). Key conserved internal residues that constitute two spines dubbed the regulatory (R-spine) and

catalytic (C-spine) span from the N-lobe to the C-lobe have been identified as being central to the dynamic interconversion between active and inactive kinase state (Figure 1.5). In the active conformation of many kinase structures the R-spine is assembled satisfying the classic hallmarks of the active state. However the active conformation poised for competent catalysis is also dependent on the complete assembly of the C-spine which requires ATP to be bound in the catalytic cleft. Here the adenosine moiety becomes part of the spine.
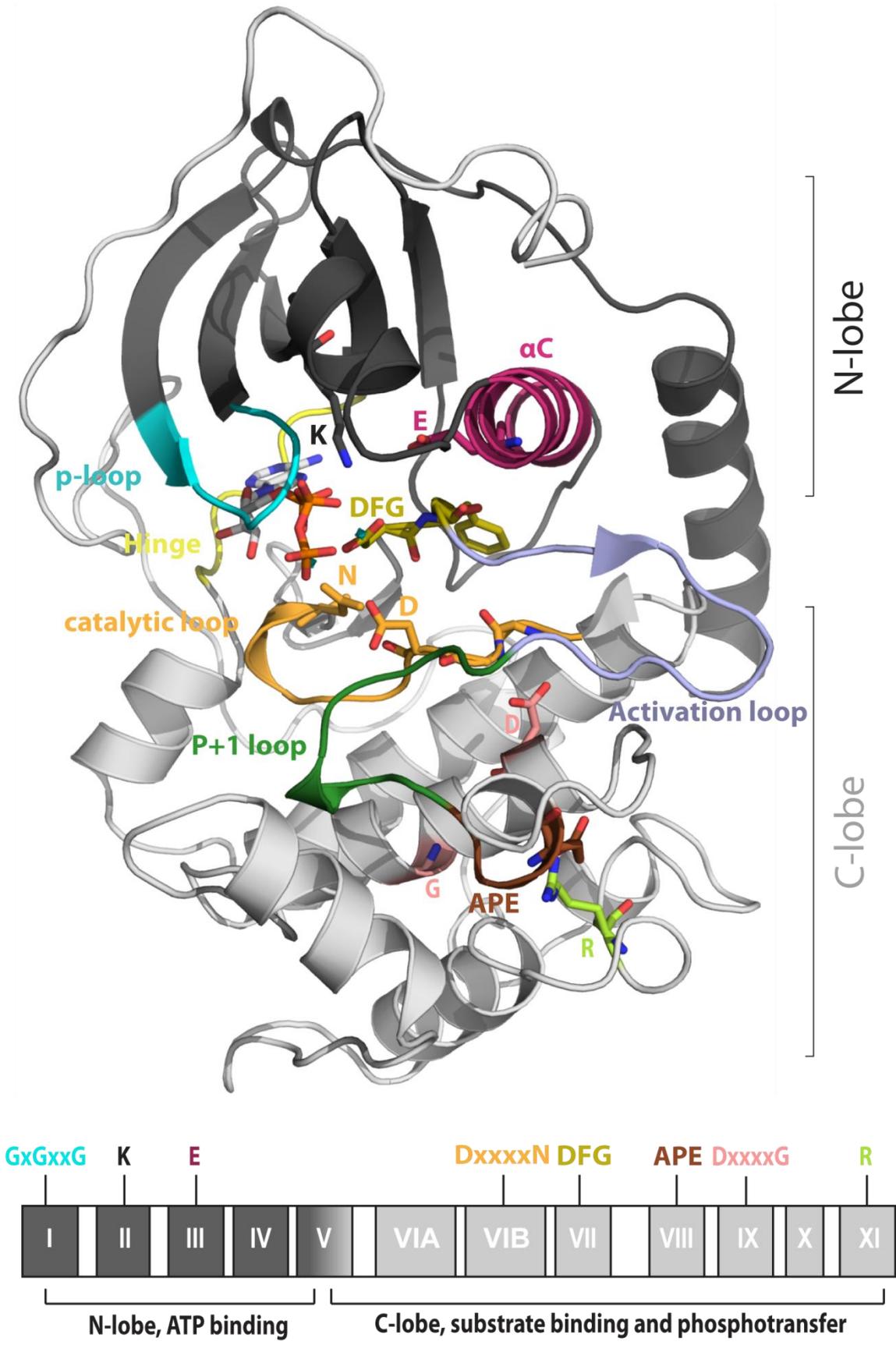
Figure 1.2. The eukaryotic catalytic kinase domain.

PKA (PDB ID: 1ATP) the archetypal KD illustrates canonical bilobal fold consists of a small N-lobe (black) and a large C-lobe (gray). Here PKA adopts an active and catalytically competent conformation poised for catalysis. Key functional motifs that contribute to substrate binding and catalysis are highlighted in colour. The p-loop, hinge, DFG, catalytic Lys, and catalytic loop interact with ATP-Mg^{2+}.
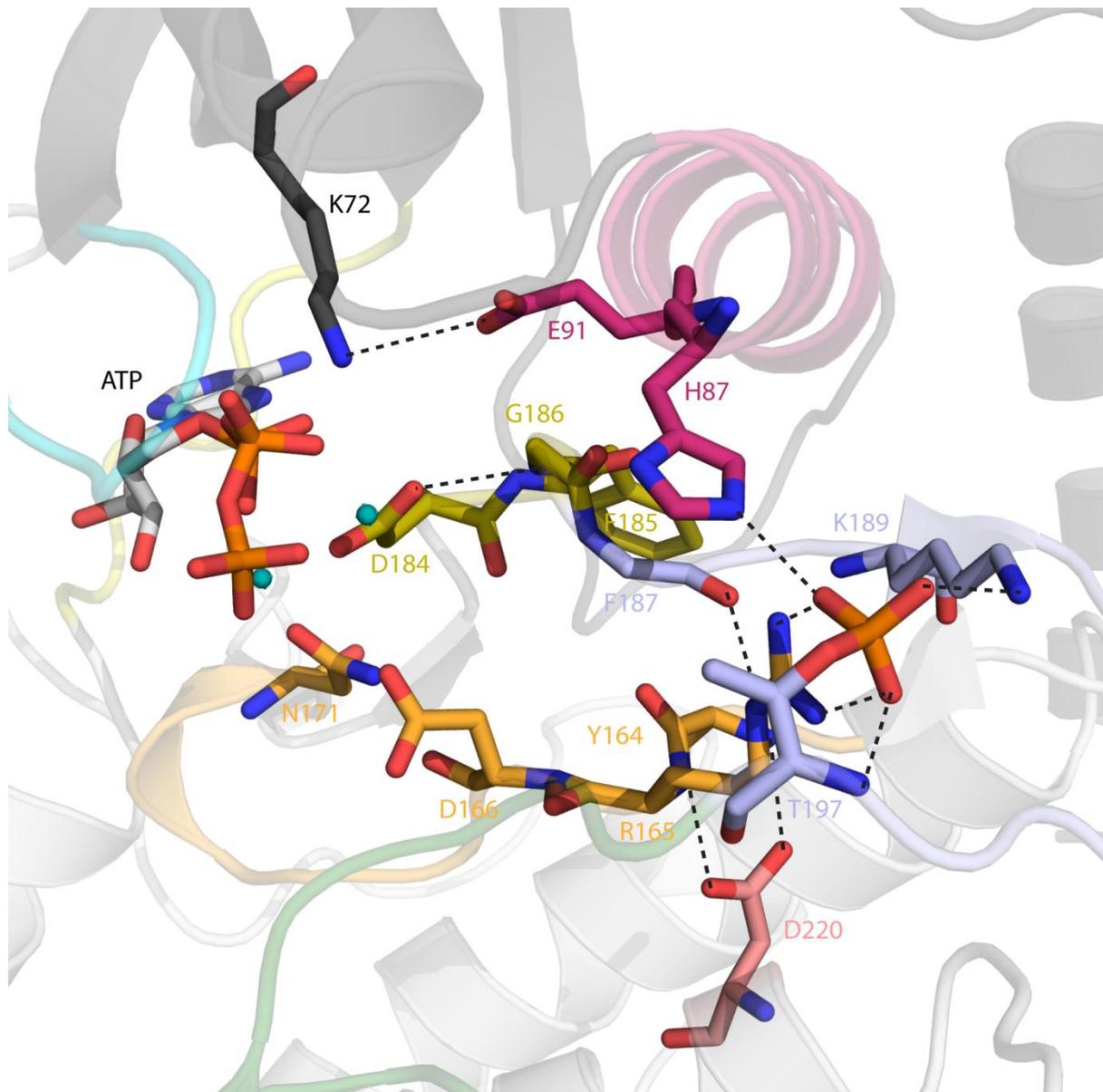
Figure 1.3. Key residues in the catalytic cleft.

PKA is itself regulated by phosphorylation, which increases its activity. Residues are shown in the PKA active conformation of the catalytic cleft poised for catalysis. In the activation loop phospho Thr 197 forms a network of interactions that coordinates and stabilizes the catalytic loop, DFG motif, and αC helix in their catalytically competent confirguation.
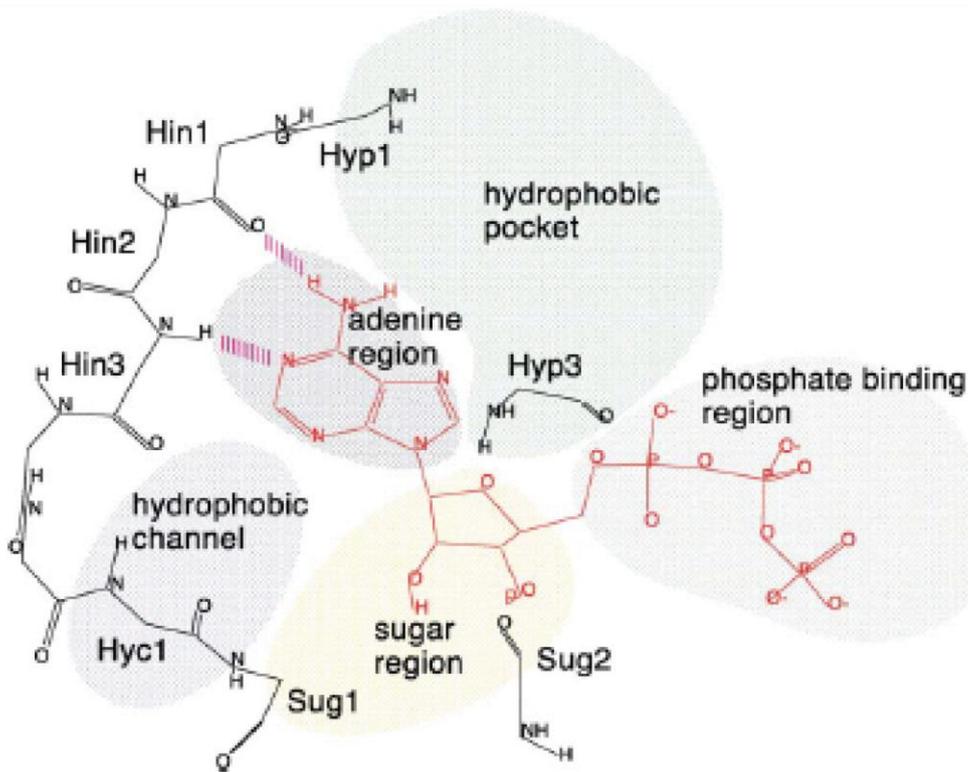
Figure 1.4. Key regions within the catalytic cleft.

Binding of ATP-Mg^{2+} to a protein kinase active site cleft can be described in terms of several specific regions dubbed the hinge, adenine, sugar, and phosphate binding regions, the hydrophobic channel, and the hydrophobic pocket.
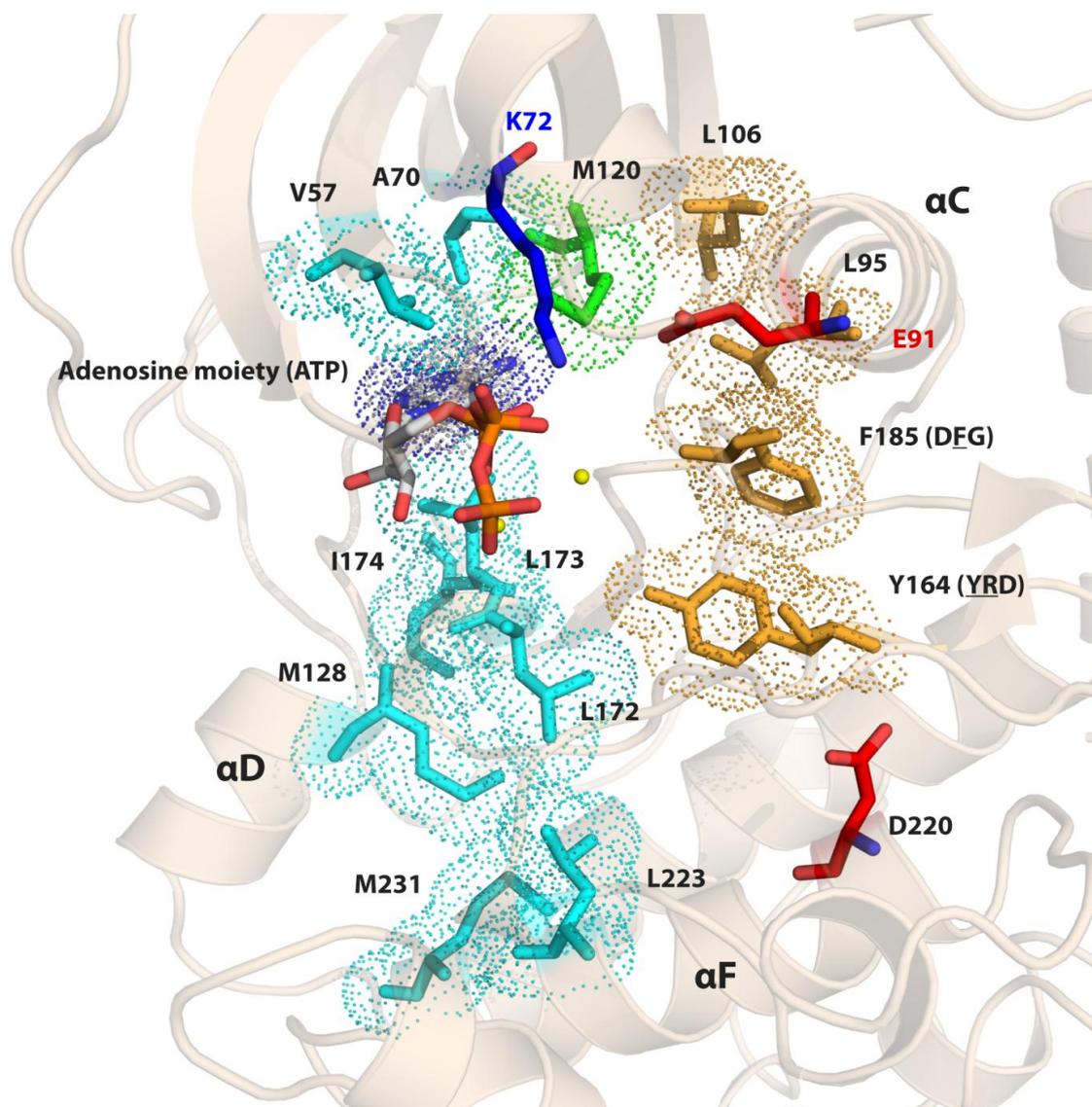
Figure 1.5. Internal hydrophobic spines.

The conformation of an active kinase poised for competent catalysis can be described in terms of conserved R-spine (orange) and C-spine (cyan) assembly. Here the both spines are completely assembled. Notable hallmarks of a catalytically competent conformation such as the conserved Lys-Glu salt bridge and DFG-in conformation are satisfied.

TOPK

T-cell originated protein kinase (TOPK) also known as PDZ-binding kinase (PBK), is a 36 kDa Ser/Thr kinase classified under the other group and was first identified in a yeast two hybrid screen interacting with a tumor suppressor protein termed the PDZ2 domain of human homologue of the drosophila discs-large (hDlg). Subsequent analysis of TOPK revealed that it features a KD followed by a C-terminal T/S-X-V PDZ interacting motif (Gaudet, Branton et al. 2000). Shortly after its discovery, TOPK was also identified in lymphokine-activated killer T (T-LAK) cells and testis. BLAST (basic local alignment search tool) homology searching and phylogenetic tree analysis revealed that TOPK is related to the dual specific kinase mitogen activated protein kinase kinase (MAPKK) whilst sequence analysis revealed divergence in some conserved functional motifs (Abe, Matsumoto et al. 2000).

Studies have indicated that TOPK is up-regulated during mitosis, meiosis, and is involved in the promotion of cell proliferation and cytokinesis. This is coupled with phosphorylation by cdk1/cyclin B at the T9 position to form a complex that associates with protein regulator of cytokinesis (PRC1) which interacts with microtubules and enhances cytokinesis. Interesringly the suppression of TOPK disrupts cytokinesis (Matsumoto, Abe et al. 2004; Abe, Takeuchi et al. 2007; Park, Nishidate et al. 2009). Furthermore a role for TOPK in spermatogenesis has also been described (Fujibuchi, Abe et al. 2005). Activation of TOPK requires phosphorylation on T198 but the kinase that mediates this is unknown, although TOPK activity is believed to be down-regulated by protein phosphatase 1α (PP1α) (Abe, Matsumoto et al. 2000; Matsumoto, Abe et al. 2004; Fujibuchi, Abe et al. 2005; Park, Nishidate et al. 2009).

Up-regulated TOPK expression is evident in T-LAK cells, activated lymphoid cells and tumors (Abe, Matsumoto et al. 2000), and hematologic tumors that include leukemia, lymphoma, and myeloma. Furthermore, TOPK up-regulation correlates with the malignancy of these tumors (Simons-Evelyn, Bailey-Dell et al. 2001; Cote, Simard et al. 2002; Nandi, Tidwell et al. 2004). A role for TOPK is suggested in UVB-

mediated skin carcinogenesis (Oh, Zhu et al. 2007), colorectoral cancer (Zhu, Zykova et al. 2007), and HeLa cervical cancer cells (Kwon, Lee et al. 2010). In melanoma cells TOPK phosphorylation of histone H2AX prevents apoptosis induced by arsenite (Zykova, Zhu et al. 2006). TOPK is highly expressed in breast cancer and promotes cell growth via Histone H3 phosphorylation but siRNA mediated gene silencing of TOPK inhibits breast cancer cell growth (Park, Lin et al. 2006).

Downstream phosphorylation targets of TOPK include Leu-Gly-Asn repeat enriched protein/G-protein signalling modulator 2 (LGN/GSM2) whose activation mediates cell division in breast cancer cells (Fukukawa, Ueda et al. 2010) Moreover, activated TOPK can phosphorylate p38 MAPK (Abe, Matsumoto et al. 2000). Downstream of the insulin-like growth factor 1/insulin-like growth factor receptor 1 (IGF/IGF-IR) signalling pathway TOPK mediates tumor cell proliferation via p38 MAPK (Herrero-Martin, Osuna et al. 2009). Additional TOPK interacting proteins include the tumor suppression protein p53, adapter protein p47 of the p97 AAA ATPase, and ERK2 which TOPK can phosphorylate and vice versa (Zhu, Zykova et al. 2007; Park, Nishidate et al. 2009; Hu, Gartenhaus et al. 2010).

Pim1

First discovered as the preferential proviral integration site in Maloney murine leukemia virus (MoMuLV) induced T-cell lymphoma (Cuypers, Selten et al. 1984), Pim1 is a member of the Ser/Thr Pim kinase family that also constitutes Pim2 and Pim3. The Pim kinases are catagorized under the CAMK group. Pim1 is expressed predominantly as a 33 kDa protein although a 44 kDa isoform also exists in human prostate cancer cells and contains an N-terminal proline-rich motif that binds to the SH3 domain of tyrosine kinase Etk. The 33 kDa Pim1 is found within the cytosol and nucleus whereas the 44 kDa Pim1 localises around the plasma membrane (Xie, Xu et al. 2006). Due to different upstream initiation codons in the gene, murine Pim1 also exists as 33 and 44 kDa isoforms (Padma and Nagarajan 1991). Pim1 (1XWS) adopts a constitutively active conformation and contains a Pro insertion in its hinge

region resulting in novel hydrogen-bonding patterns and an expanded ATP binding pocket (Qian, Wang et al. 2005).

Tissue specificity of Pim1 is wide spread, with PIm1 occurring in hematopoietic cells, lymphoid cells, prostate, testis, thymus, and bone marrow. Regulation of Pim1 activity occurs at transcriptional, post transcriptional, translational and post translational levels but is predominantly transcriptional. Pim1 activity correlates with its protein levels (Zhu, Ramirez et al. 2002). Physiologically Pim1 is implicated in the proliferation and survival of cells (Wang, Bhattacharya et al. 2001; Wang, Bhattacharya et al. 2002). Expression of Pim1 is induced by cytokines, mitogens, and growth factors (Bachmann and Moroy 2005). Pim1, 2 and 3 deficient mice exhibit reduced body size at birth and during the postnatal period (Mikkers, Nawijn et al. 2004). Transgenic mice over expressing Pim1 become more susceptible to tumor incidence, carcinogens and MoMuLV induced lymphomagenesis (Domen, van der Lugt et al. 1993). Activation of Pim1 is co-operative with the oncogene c-myc (van Lohuizen, Verbeek et al. 1989). Double transgenic mice over expressing both Pim1 and c-myc display pre-B-cell leukemia *in utero* (Verbeek, van Lohuizen et al. 1991).

Abberant over expression of Pim1 is associated with human lymphomas, acute leukemia, and neoplastic prostate cancer (Amson, Sigaux et al. 1989; Dhanasekaran, Barrette et al. 2001). Furthermore Pim1 over expression due to chromosomal translocation occurs in the common form of non-Hodgkin's lymphoma known as diffuse large cell lymphoma (DLCL) (Akasaka, Akasaka et al. 2000). Aberrant somatic hypermutations associated with Pim1 have also been identified in DLCL (Pasqualucci, Neumeister et al. 2001).

PAK5

p21-activated kinase 5 (PAK5) is a member of the PAK family of Ser/Thr kinases which are catagorised under the STE group. PAKs are downstream effectors of the small GTPases Cdc42 and Rac1. Six isoforms constitute the PAK family and are classified into group 1 (PAK1, 2, 3) and 2 (PAK 4, 5, 6) based on their domain architecture and regulation. Both groups feature a C-terminal KD, a central domain containing Proline-rich motifs which act as docking sites for SH3 domain containing proteins, and a N-terminal regulatory domain (RD) composed of a p21-binding domain (PBD) also known as Cdc42/Rac1 interactive binding domain (CRIB). Group 1 members have an additional autoinhibitory domain (AID) that overlaps the PBD and require GTP•Cdc42/Rac1 binding to the PDB/CRIB to relieve inhibition and initiate activation. In contrast group 2 PAKs are constitutively active and do not require GTP•Cdc42/Rac1 binding (Jaffer and Chernoff 2002; Wells and Jones 2010). Instead binding of small GTPases regulates their subcellular localisation (Abo, Qu et al. 1998). Group 2 PAKs express a high level of domain plasticity and experience many conformational changes during catalysis suggesting that their mechanism of transition from a catalytically non-productive to productive conformation is different from group 1 members (Eswaran, Lee et al. 2007).

PAK5 was first identified from an RT-PCR screen using degenerate primers based on PAK4 and found to express predominantly in the brain as an 80 kDa protein. The KD features a Ste20 motif that is highly conserved in all Ste20 kinases. PAK5 is closely related to its other group II members, showing 85 % and 80 % sequence identity with PAK4 and PAK6 respectively but only about 50 % sequence identity with group I members. Unlike group I PAKs which bind both GTP-Cdc42/Rac1, PAK5 can only bind to GTP-Cdc42 (Pandey, Dan et al. 2002). Similarly PAK4 preferentially binds GTP-Cdc42 but only weakly to Rac1 (Abo, Qu et al. 1998). Interestingly although PAK5 is constitutively active it is reported to contain an AID in the N-terminal domain (Ching, Leong et al. 2003).

Physiologically the PAK kinases are implicated in cell motility, morphology, cell survival, hormone signalling, and gene transcription (Wells and Jones 2010). Current evidence suggests that PAK5 is implicated in cell survival and neurite extension. Localisation of PAK5 to the mitochondria from the nucleus and its activatiion of anti-apoptotic pathways involving Raf1 and BAD promote cell survival (Cotteret, Jaffer et al. 2003; Cotteret and Chernoff 2006; Wu, Carr et al. 2008). Over expression of PAK5 in neuronal cells promote neurite outgrowth (Dan, Nath et al. 2002).

CSNK1G3

Casein kinase 1 γ isoform 3 (CSNK1G3) is one of 7 known human isoforms that constitute the Casein kinase 1 (CK1) family of Ser/Thr kinases. CSNK1G3 is 51 kDa in size and is the largest CK1 isoform. CK1 kinases are monomeric and constitutively active and highly conserved across the kinase catalytic domain. They possess variable N-terminal and C-terminal domains. CK1 kinases phosphorylate a wide spectrum of substrate targets that include for example transcription factors, splice factors, translation factors, cytoskeletal proteins, and vesicle associated proteins (Knippschild, Gocht et al. 2005). Splice variants of the CK1 isoforms result in the modulation of their activities, subcellular localisation, function, and biochemical properties. The rat homolog of CSNK1G3 is highly expressed in brain, heart, liver, kidney, and lung (Zhai, Graves et al. 1995). In kidney cancer cells the activity of CK1 γ isoforms is elevated (Elias, Li et al. 1981).

CK1 kinases have a substrate consensus sequence that to some extent requires a priming phosphorylation step involving another kinase phopsphorylating a Ser/Thr near the substrate consensus sequence (Flotow, Graves et al. 1990; Meggio, Perich et al. 1991; Meggio, Perich et al. 1992). CSNK1G3 has a large regulatory C-terminal domain that features an inhibitory domain. Autophosphorylation of the inhibitory domain results in autoinhibition via a pseudosubstrate mechanism (Graves, Haas et al. 1993; Fish, Cegielska et al. 1995; Graves and Roach 1995; Zhai, Graves et al. 1995; Longenecker, Roach et al. 1998). Truncation of the C-terminal domain increases kinase activity (Carmel, Leichus et al. 1994; Zhai, Graves et al. 1995). A

CK1 kinase inhibited homodimer model has been postulated (Longenecker, Roach et al. 1998). Another described mechanism of regulating the constitutively active CK1 kinases involves changing its subcellular localisation and compartmentalisation (Ho, Mason et al. 1997).

multifunctional or restricted. Multifunctional types exhibit broad substrate specificity and include CaMKK, CaMK1, CaMK2, and CaMK4. Restricted types exhibit single substrate specificity and include CaMK3 (also known as eEF-2K, or elongation factor 2), myosin-light chain kinase (MLCK), and phoshorylase kinase (Hook and Means 2001; Soderling and Stull 2001).

With the exception of CaMK3, all CaMK family members a share similar domain organisation (Figure 1.6), comprising of an N-terminal catalytic KD followed by a C-terminal RD which maintains the kinase in an autoinhibited and inactive state via the AID component. A calmodulin binding domain (CBD) overlaps the C-terminal end of the AID. CaMK2 and phosphorylase kinase feature an additional association domain (AD) C-terminal of RD which enables multimerization. All CaMK members regulated by the RD are activated by the binding of Ca^{2+}/Calmodulin (Ca^{2+}/CaM) to the CBD which relieves autoinhibition and exposes the substrate binding site and activates the catalytic KD. Once activated CaMKs can phosphorylate Ser/Thr residues in protein substrates to alter their functionality (Hook and Means 2001; Soderling and Stull 2001).

Some CaMKs are further regulated by phosphorylation within their activation loop by another upstream CaMK. For example CaMKK can phosphorylate CaMK1 and CaMK4. Both CaMK2 and CaMK4 are capable of autophosphorylation after Ca^{2+}/CaM dependent activation. CaMK2 autophosphorylation can occur at two sites within the AID. Depending on which site and the sequential order of autophosphorylation, CaMK2 can undergo autonomous activity or inhibition (Figure 1.7). Likewise CaMK4 exhibits Ca^{2+}/CaM independent activity if autophosphorylation occurs within the N-terminus. If autophosphorylation is within the AID then this

causes CaMKIV inhibition. Phosphorylation can therefore lead to activation and inactivation (Swulius and Waxham 2008).
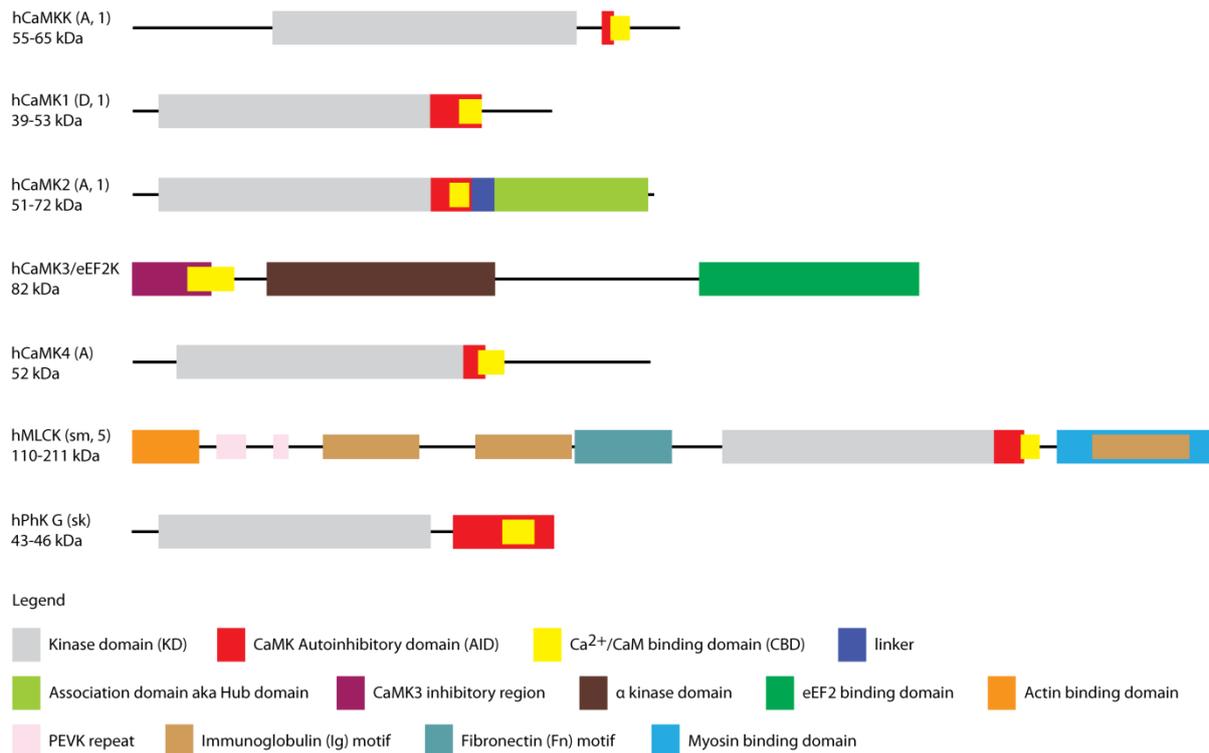
Figure 1.6. Domain organisation of the CaMK family.

A common feature for all the members of the CaMK family except for CaMK3 is a C-termina regulatory domain comprised of the AID which imparts autoinhibition on the kinase, and the CBD whic relieves autoinhibition upon Ca^{2+}/CaM binding. The KD of CaMK3 lacks sequence homology to the conserved eukaryotic KD and its 12 conserved subdomains and is defined as the αKD (Swulius and Waxham 2008).
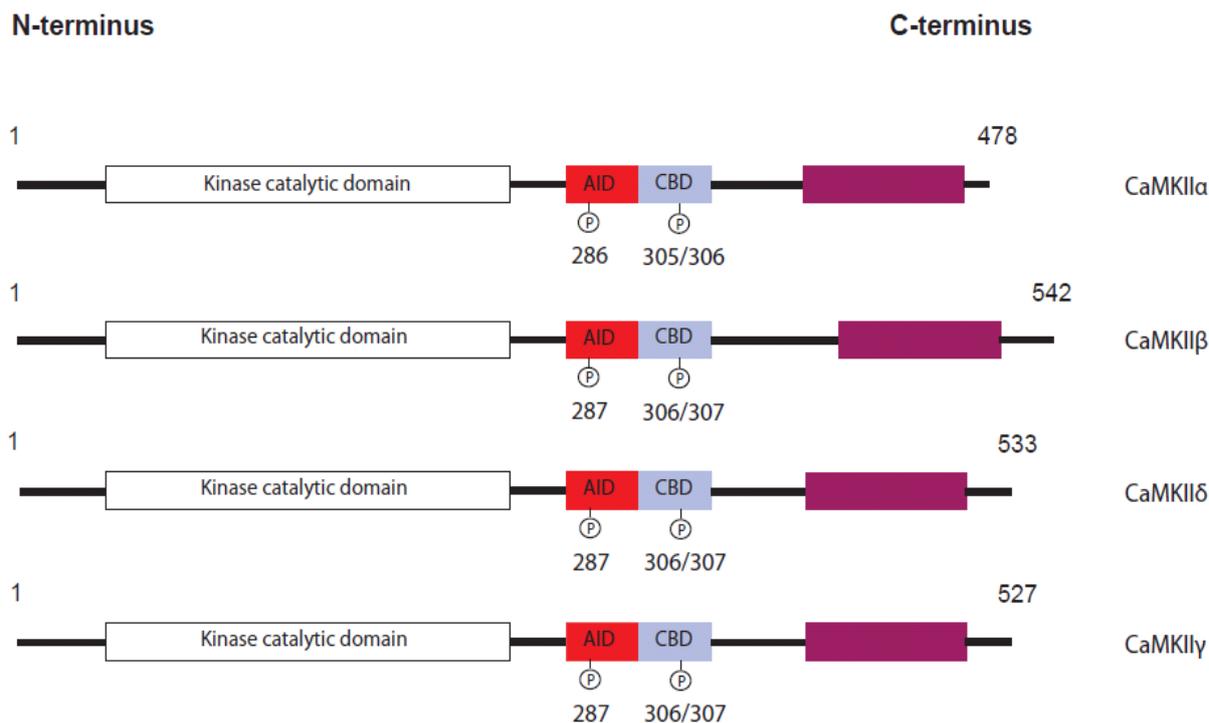
Figure 1.7. Domain organisation of CaMK2 subfamily members.

Amongst the CaMK2 isoforms the linker region between the CBD and the AD exhibits the most sequence diversity as a result of splice variants. CaMK2 regulation is also modulated by autophosphorylation in the AID or CBD, an intermolercular event that occurs within the same oligomeric holoenzyme dependent on the association domain. Both the activating and substrate CaMK2 require Ca^{2+}/CaM. Autophosphorylation of T286/287 site enables autonomous activity by permitting Ca^{2+} independence dubbed CaM trapping. Conversely autophosphorylation of T305/306/307 is inhibitory, abolishing Ca^{2+}/CaM binding.

(Figure 1.8). Sequence diversity is most evident in the C-terminal stretch of residues after the RD. In CaMK1A (also known as CaMK1 because it was the first CaMK1 studied), CaMK1B (also known as PncK) and CaMK1D isoforms, this stretch comprises of 30-60 C-terminal residues of unknown function. However, with CaMK1G (also known as CLICK-III) the C-terminal stretch of residues extends 92 residues after AID and terminates with a CAAX (where C is Cys, A is any aliphatic residue, and X is any residue) motif. Prenylation occurs on this motif and mediates CaMK1G targeting to the membranes of the sarcoplasmic reticulum and golgi (Takemoto-Kimura, Terai et al. 2003). Initial activation of CaMK1 members requires Ca^{2+}/CaM and maximal activity requires phosphorylation of a Thr residue within the activation loop that is mediated by CaMKK. The latter process requires the binding of Ca^{2+}/CaM to both CaMKI and the upstream CaMKK (Haribabu, Hook et al. 1995).

CaMK1D

Calcium calmodulin dependent protein kinase I δ (CaMK1D) is a 43 kDa Ser/Thr protein kinase that belongs to the CaMK1 subfamily from the CaMK family consisting of four isoforms (α, β, γ, δ) which are expressed from four different genes and can produce splice variants. all exhibit the canonical domain organisation of the CaMK family. CaMK1D was originally isolated from a HeLa cDNA library and found to share 80% sequence identity with CaMK1A (Ishikawa, Tokumitsu et al. 2003). A splice of CaMK1D dubbed CKLiK variant atory domains of both are identical in sequence showing diversity only after Leu 348 of CaMK1D. Therefore CKLiK is believed to be a splice variant of CaMK1D (Ishikawa, Tokumitsu et al. 2003). CKLiK is a granulocyte specific kinase (Verploegen, Lammers et al. 2000; Verploegen, Ulfman et al. 2005)

Expression of CaMK1D is widespread in other tissues including brain, liver, pancreas, spleen, thymus, prostate, testis, ovary, and colon (Ishikawa, Tokumitsu et al. 2003). CKLiK is specifically expressed in polymorphonuclear leukocytes including neutrophilic and eosinophilic granulocytes (Verploegen, Lammers et al. 2000). CaMK1B exhibits similar tissue specificity to CaMK1D, whereas CaMK1G expression appears highest in neurons but is low in skeletal muscle, kidney, liver and spleen. In contrast CaMK1A expression is ubiquitous. Heterogeneous

expression of CaMKI isoforms occurs in different parts of the brain (Haribabu, Hook et al. 1995; Takemoto-Kimura, Terai et al. 2003; Colomer and Means 2007). CaMKI isoforms are predominantly cytosolic. Although mouse CaMKI homologs and splice variants demonstrate translocation to the nucleus after cell stimulation (Ueda, Sakagami et al. 1999; Rina, Jusuf et al. 2001; Sakagami, Kamata et al. 2005).

In neurons CaMK1D exhibits a two-fold increase in translocation following neuron activation with Ca^{2+} mobilizing stimuli. CaMK1D nuclear translocation is dependent on its activation because two inactive forms of CaMK1D, K52A and T180A abolish translocation. K52A renders CaMK1D catalytically inactive within the ATP binding site whereas T180A prevents activation loop phosphorylation by CaMKK. Furthermore nuclear translocation of CaMK1D is blocked in stimulated neurons preincubated with STO-609, a CaMKK inhibitor (Sakagami, Kamata et al. 2005). Interestingly a constitutively active construct of CKLiK identical in sequence to the CaMK1D kinase catalytic domain demonstrates localisation to the nucleus (Verploegen, Lammers et al. 2000).

Activation of CaMKI members requires Ca^{2+}/CaM and enhanced activity requires activation loop phosphorylation by CaMKK. Ishikawa et al assessed the activity of fully active CaMK1D in vitro by incubating CaMK1D with CaMKK and Ca^{2+}/CaM and found CaMK1D to be active. Activation of CaMK1D by Ca^{2+}/CaM alone produces 30-fold less activity due to the absence of Thr 180 phosphorylation mediated by CaMKK. Furthermore CaMK1D could not be phosphorylated by either CaMKKα or the autonomously active isoform CaMKKβ3 in the absence of Ca^{2+}/CaM. This suggested that the activation loop is not accessible to CaMKKs unless Ca^{2+}/CaM is present (Ishikawa, Tokumitsu et al. 2003). CaMK1D can also be activated using HeLa cell extract suggesting that CaMKK is expressed in the extract. A GST-CaMK1D fusion exhibited Ca^{2+}/CaM dependent autophosphorylation but curiously had no impact on CaMk1D activity (Ishikawa, Tokumitsu et al. 2003). Evidence of CaMKI activation loop autophosphorylation or autophosphorylation of another residue was lacking but may have occurred on GST or on an unidentified residue that does not directly modulate CaMK1D activity.

In further experiments, an increase in the intracellular Ca^{2+} concentration was induced with inomycin and produced a 1.5-fold increase in CaMK1D activity in HeLa cells compared to the absence of inomycin. Their data suggested that CaMK1D in HeLa cells exhibit high activity even without increasing the intracellular Ca^{2+} concentration. A high level of sustained CaMK1D activity was presumed to be due to the upstream activating CaMKK even at resting Ca^{2+} concentrations. Consequently CaMK1D activity in resting HeLa cells was found to be reduced when the CaMKK inhibitor STO-609 was applied. In summary the HeLa cells contained CaMKK that could activate CaMK1D in both the presence and absence of inomycin and that sustained CaMK1D activity was due in part to the Ca^{2+} independent activity of CaMKK (Ishikawa, Tokumitsu et al. 2003). Interestingly Ca^{2+} independent activation of CaMKIs is demonstrated in human T lymphocyte cells under conditions of oxidative stress suggesting a novel mechanism of CaMK1 activation (Howe, LaHair et al. 2002).

CaMK1D and other CaMK1 iosforms are currently not well characterised in terms of their biological role but in vitro studies have shown CaMK1 activity towards substrates such as CREB (cyclic AMP response element binding protein 1), synapsin, cystic fibrosis transmembrane regulator (CFTR), and eukaryotic translation initiation factor 4 gamma 2 (eIF4GII). Perhaps the best characterised and physiologically relevant substrates are CREB and eIF4GII (Picciotto, Cohn et al. 1992; Picciotto, Czernik et al. 1993; Sun, Lou et al. 1996; Qin, Raught et al. 2003). CKLiK is better characterised and shown to be expressed in human polymorphonuclear leukocytes (PMNs) that include neutrophilic and eosinophilic granulocytes. PMNs are implicated in defending the host from foreign invading microorganisms, damaged cells and cellular debris. During inflammation, PMNs are recruited to the site of inflammation through the release of chemokines at the site of inflammation and upon activation PMNs mediate functions such as phagocytosis. CKLiK is therefore proposed to play a major role in signalling such cellular responses in PMNs (Verploegen, Lammers et al. 2000; Verploegen, Ulfman et al. 2005).

CaMK1G

Calcium calmodulin dependent protein kinase I γ (CaMK1G) is the largest isoform of the CaMK1 subfamily and was first identified in a screen for novel putative CaMK-like CREB kinases (CLICKs) using degenerate PCR strategies and subsequently showed high sequence homology to CLICK I and II (Takemoto-Kimura, Terai et al. 2003). In vivo studies demonstrate that CaMK1G has the ability to anchor to the golgi and plasma membrane in a CAAX motif and prenylation dependent manner. Mutagenesis of the CAAX motif Cys prenylation site abolished the recruitment of CaMK1G to the golgi and plasma membrane. Likewise reduction of prenylation via inhibition of HMG-CoA produced a similar effect (Takemoto-Kimura, Terai et al. 2003). CaMK1G mRNA transcripts express predominantly in the brain with lower amounts evident in skeletal muscle, spleen, liver and kidney. Furthermore expression is particularly high in neurons with CaMK1G implicated in dendrite formation via lipid raft association. Actin polymerisation is also promoted by CaMK1G via downstream activation of the MEK/erk pathway (Takemoto-Kimura, Terai et al. 2003; Takemoto-Kimura, Ageta-Ishihara et al. 2007)
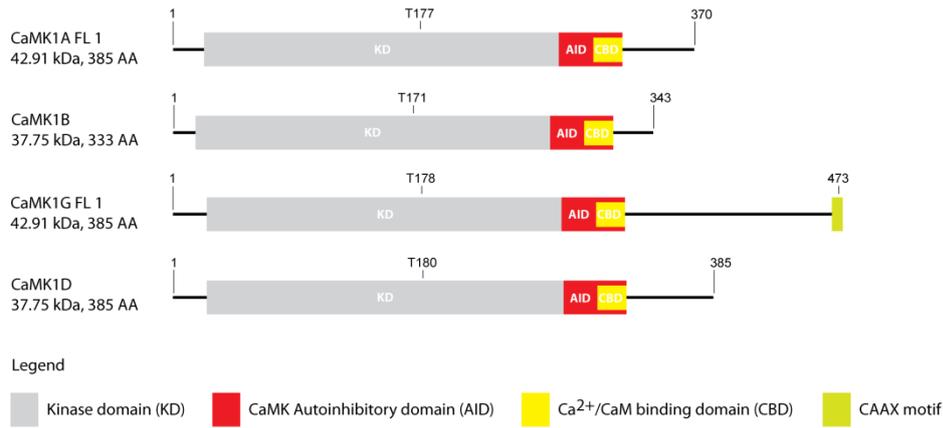
Figure 1.8. CaMK1 domain organisation.

This schematic illustrates the domain organisation of the CaMK1 subfamily members. They consist of a KD and C-terminal RD composed of the AID and overlapping CBD. Additional C-terminal residues after the RD are undefined in terms of function except in CaMK1G, which features a CAAX motif for membrane targeting. Ca^{2+}/CaM binding to the CBD relieves autoinhibition. Further activation can follow via activation loop phosphorylation by CaMKK as indicated by the phosphorylated residues.

Secondary structure elements (top): β1 — αB — β2 — p-loop — β3 — β4

```
CaMK1A    1  ---MLGAVEGPRWKQAEDIRDIYDFRDVLGTGAFSEVILAEDKRTQKLVAIKCIAKEALE
CaMK1D    1  MARENGESSSSWKKQAEDIKKIFEFKETLGTGAFSEVVLAEEKATGKLFAVKCIPKKALK
CaMK1B    1  --------MLLLKKHTEDISSVYEIRERLGSGAFSEVVLAQERGSAHLVALKCIPKKALR
CaMK1G    1  MGRKEEDDCSSWKKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA
```

Secondary structure elements: αC — β5 — β6 — hinge — αD

```
CaMK1A   58  GKEGSMENEIAVLHKIKHPNIVALDDIYESGGHLYLIMQLVSGGELFDRIVEKGFYTERD
CaMK1D   61  GKESSIENEIAVLRKIKHENIVALEDIYESPNHLYLVMQLVSGGELFDRIVEKGFYTEKD
CaMK1B   53  GKEALVENEIAVLRRISHPNIVALEDVHESPSHLYLAMELVTGGELFDRIMERGSYTEKD
CaMK1G   60  FRDSSLENEIAVLKKIKHENIVTLEDIYESTTHYYLVMQLVSGGELFDRILERGVYTEKD
```

Secondary structure elements: αE — C-loop — αF — β7 — β8 — A-loop

```
CaMK1A  118  ASRLIFQVLDAVKYLHDLGIVHRDLKPENLLYYSLDEDSKIMISDFGLSKMEDPGSVLST
CaMK1D  121  ASTLIRQVLDAVYYLHRMGIVHRDLKPENLLYYSQDEESKIMISDFGLSKMEGKGDVMST
CaMK1B  113  ASHLVGQVLGAVSYLHSLGIVHRDLKPENLLYATPFEDSKIMVSDFGLSKIQ-AGNMLGT
CaMK1G  120  ASLVIQQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMST
```

Secondary structure elements: A-loop — αG — αH — αI

```
CaMK1A  178  ACGTPGYVAPEVLAQKPYSKAVDCWSIGVIAYILLCGYPPFYDENDAKLFEQILKAEYEF
CaMK1D  181  ACGTPGYVAPEVLAQKPYSKAVDCWSIGVIAYILLCGYPPFYDENDSKLFEQILKAEYEF
CaMK1B  172  ACGTPGYVAPELLEQKPYGKAVDVWALGVISYILLCGYPPFYDESDPELFSQILRASYEF
CaMK1G  179  ACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPPFYEETESKLFEKIKEGYYEF
```

Secondary structure elements: αJ — αK — αL

```
CaMK1A  238  DSPYWDDISDSAKDFIRHLMEKDPEKRFTCEQALQHPWIAGDTALDKNIHQSVSEQIKKN
CaMK1D  241  DSPYWDDISDSAKDFIRNLMEKDPNKRYTCEQAARHPWIAGDTALNKNIHESVSAQIRKN
CaMK1B  232  DSPFWDDISESAKDFIRHLLERDPQKRFTCQQALRHLWISGDTAFDRDILGSVSEQIRKN
CaMK1G  239  ESPFWDDISESAKDFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKN
```

Secondary structure elements: αM

```
CaMK1A  298  FAKSKWKQAFNATAVVRHMRKLQLGTSCEGQGQTASHG----------------------
CaMK1D  301  FAKSKWRQAFNATAVVRHMRKLHLGSSLDSSNASVSSS----------------------
CaMK1B  292  FARTHWKRAFNATSFLRHIRKLGQIPEGEGASECGMAR----------------------
CaMK1G  299  FAKSKWRQAFNAAAVVHHMRKLHVNLHSPGVRPEVENRPPETQASETSRPSSPEITITEA
```

```
CaMK1A  336  ----------------------------------------------ELLTPVAGGPAAGCCCRD
CaMK1D  339  -----------------------------------------------LSLASQKDCLAPSTLCSF
CaMK1B  330  -----------------------------------------------HSHSGLRAGQPPKW----
CaMK1G  359  PVLDHSVALPALTQLPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPNHQGSLAAGPCGC
```

```
CaMK1A  354  CCVEP-----G-TELSP-----TLPHQL---------------------------------
CaMK1D  357  ISSSSGVSGVG-AERRPRPTTVTAVHSGSK-------------------------------
CaMK1B       ------------------------------------------------------------
CaMK1G  419  CSSCLNIGSKGKSSYCSEPTLLKKANKKQNFKSEVMVPVKASGSSHCRAGQTGVCLIM
```

Figure 1.9. Sequence alignment of CaMK1 isoforms.

The KD and RD of the CaMK1 isoforms show high sequence homology. Non-homologous regions arise mainly in the C-terminal end after the RD. Conserved residues, partially conserved, and similar residues are coloured in red, yellow and green respectively. The alignment and was generated using ClustalW (Thompson, Higgins et al. 1994), whilst colouring was done using TEXSHADE (Beitz 2000) from the biology workbench software (Subramaniam 1998). The gray, brown, and blue bars above the sequence alignment correspond to the KD, regulatory AID, and CBD respectively. Important loop regions such as the p-loop, hinge, catalytic loop, and activation loop are shown by the denoted by text whilst the illustrated secondary structure elements correspond to CaMK1D determined from its crystal structure (PDB: 2JC6). The arrow denotes the activating phosphorylation site.

The CaMK cascade

CaMKK can modulate the activities of CaMK1, and CaMK4 via phosphorylation. These multifunctional CaMKs constitute the CaMK cascade, with CaMKK as the upstream kinase component. CaMKK can promote maximal activity and modulation of substrate specificity and affinity of CaMK1 and CaMK4 provided all CaMK cascade members are already Ca^{2+}/CaM activated. It is postulated that the activation loop of CaMK1s and CaMK4s is only accessible to CaMKK when bound to Ca^{2+}/CaM (Soderling 1999; Means 2000; Means 2008). The CaMK cascade is a physiological Ca^{2+}/CaM dependent signalling cascade that is involved in many cellular processes. Established findings indicate that the cascade exhibits cross talk with other kinases (Figure 1.10). For example, CaMKK can also interact with PKA, PKB, and AMPK. PKA inhibits CaMKK whilst CaMKK can active PKB and AMPK, suggesting that the CaMK cascade is implicated in regulating cell survival and levels cellular ATP. Also illustrated is the indirect activation of MAPKs such as ERK, JNK and p38 mediated by CaMKIV (Soderling 1999; Means 2000; Means 2008). Physiological substrates of CaMKI are not well characterised due to limited evidence and uncertainty associated with its translocation to the nucleus, although studies suggest CaMK1 interacts with CREB, a substrate that is also common to CaMK4. A more detailed introduction to CaMKK, CaMK1, and CaMK4 is discussed in their respective sections with particular emphasis on CaMK1 as the bulk of this thesis focuses on the CaMK1D isoform.
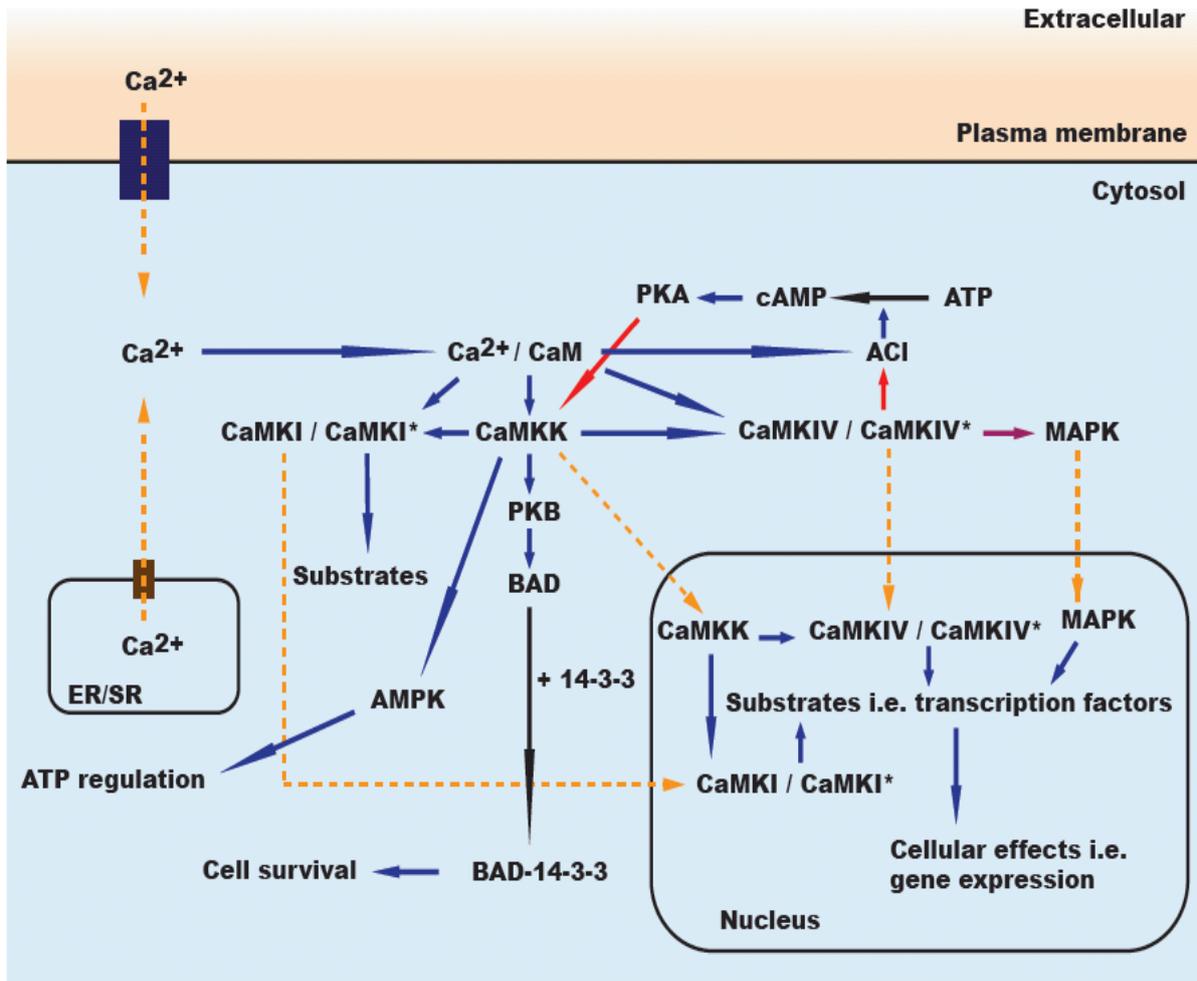
Figure 1.10. The CaMK cascade and putative signalling pathways.

The CaMK cascade is implicated in many signalling pathways and cellular processes such as gene expression but in this simplified schematic only a few of the best characterised examples are illustrated. For example it is shown that CaMKK can also activate PKB and AMPK in addition to its CaMK cascade members. CaMK1 resides predominantly within the cytosol although limited evidence suggests CaMK1 might translocate to the nucleus. With CaMKK and CaMK4 it is well established that both localise in the cytosol and nucleus. Blue arrows represent activation/phosphorylation, purple represents indirect activation, red represents inhibition, yellow represents translocation, and black represents formation/conversion.

CaMKK

Calcium calmodulin dependent protein kinase kinase, also known as CaMKK is a monomeric 55-65kDa protein kinase that exists as two isoforms (Figure 1.11), CaMKKα and CaMKKβ (Tokumitsu, Enslen et al. 1995; Kitani, Okuno et al. 1997; Anderson, Means et al. 1998). CaMKK isoforms are expressed from different genes with CaMKKα and CaMKKβ comprised of 505 and 587 residues respectively. CaMKKβ can also produce splice variants depending on the cell type. One unique splice variant in Hela cells exhibits Ca^{2+}/CaM independent activity completely (Hsu 2001(Ishikawa, Tokumitsu et al. 2003). In addition to regulation via the AID and Ca^{2+}/CaM, CaMKK is inhibited by PKA mediated phosphorylation at ser 458 of CaMKKα and ser 495 of CaMKKβ, a mechanism of inhibition equivalent to CaMK2 (Wayman, Tokumitsu et al. 1997; Matsushita and Nairn 1999). CaMKK contains a putative phosphorylation site in its activation loop and it is speculated that an upstream kinase yet to be identified might active CaMKK. Curiously CaMKK can autophosphorylate but without any effect on catalytic activity. Differences between the two CaMKK isoforms arise in the N-terminal region in which CaMKKβ features an N-terminal 23 residue sequence (residues 129-151) that imparts Ca^{2+}/CaM independent activity after activation by Ca^{2+}/CaM. This enables autonomous activity beyond transient increases in Ca^{2+} concentration (Tokumitsu, Iwabu et al. 2001). Both CaMK2 and CaMK4 are also capable of autonomous activity although via different mechanisms involving autophosphorylation.

Notably the KD of CaMKK lacks conserved acidic residues from subdomains V and VI which are generally required for the recognition of basic residues near the phosphorylation site of substrates. Instead CaMKK contains an Arg, Pro, rich insertion (RP insert) located between subdomain II and III. This insert allows CaMKK recognition of CaMK1s and CaMK4s. Deletion of the insert abolishes the ability of CaMKK to phosphorylate CaMKIs and CaMKIVs but not protein kinase B (PKB) (Tokumitsu, Takahashi et al. 1999). Despite deletion of the RP insert, CaMKK activity is retained against PKB and involves different substrate recognition determinants with PKB compared to CaMK1 and CaMK4. Generally kinase recognition of substrates is dependent on the primary sequence around the

phosphorylation site. However, with CaMKK, the tertiary structure of its substrate appears to be paramount because it cannot phosphorylate unfolded CaMK1s and CaMK4 (Hook and Means 2001).

CaMKK is expressed mainly in the brain, neurons, thymus, testis, and spleen. On a cellular level CaMKK localises in both the cytosol and nucleus (Nakamura, Okuno et al. 1996; Sakagami, Saito et al. 1998) sakagami). Expression of CaMKK seems to parallel CaMK4 unlike CaMK1 which is more broadly expressed. It is suggested that in cell types where CaMKK is not expressed, CaMK1 may exhibit substrate recognition that is independent of its activation by CaMKK or there are further isoforms of CaMKK yet to be identified. Over expression studies of CaMKK show an increase in CaMK1 and CaMK4 activity coupled with an increase in CREB mediated transcription. This correlates positively with evidence of CaMK4 residing within the nucleus. However CaMK1 localises predominantly in the cytosol. Evidence of CaMK1 residing in the nucleus is limited. Examples of CaMK1 nuclear localisation are discussed in the following CaMK1 section. Thus the mechanism in which CaMK1 phosphorylates CREB remains unresolved.
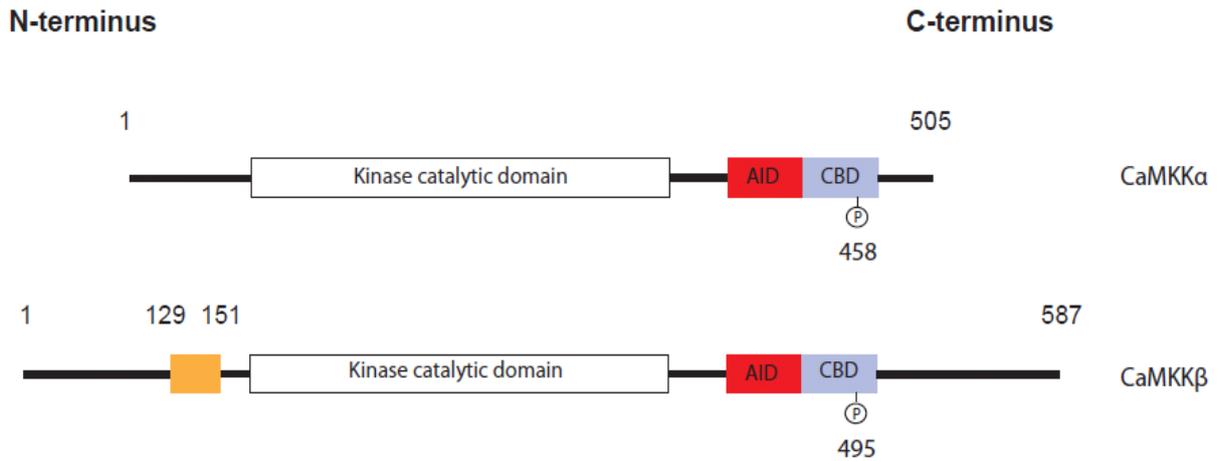
Figure 1.11. CaMKK domain organisation.

Two main isoforms of CaMKK both possess a canonical KD followed by a C-terminal RD representative of the CaMK family. CaMKK activation occurs through Ca^{2+}/CaM binding to the CBD. An active splice variant independent of Ca^{2+}/CaM also exists. CaMMKβ contains additional stretch of N-terminal residues (orange) that facilitates Ca^{2+}/CaM independent activity after Ca^{2+}/CaM activation. Inhibitory phosphorylation sites are indicated within the RD (Swulius and Waxham 2008).

CaMK4

Another member of the CaMK cascade is CaMK4, which is expressed as two isoforms from a single gene. Both CaMK4α and β isoforms are monomeric with molecular weights of 65 and 67 kDa respectively. CaMKIVβ is larger due to an additional 28 N-terminal residues (Figure 1.12). Like all CaMKs, the CaMK4s are activated by Ca^{2+}/CaM but possess a second activation step analogous to CaMK1s. Maximal activation occurs in a Ca^{2+}/CaM dependent manner when CaMK4s are phosphorylated within their activation loop by CaMKK. In contrast to CaMK1s maximally activated CaMK4 can then undergo autophosphorylation of its N-terminus at residues Ser 12 and 13 enabling Ca^{2+}/CaM independent activity. Mutatagenesis of these phosphorylation sites abolish activation by CaMKK but deletion of 20 N-terminal residues permits activation by CaMKK. The N-terminal region is suspected to function as a secondary AID that is relieved upon phosphorylation or truncation (Hook and Means 2001). Additional regulation can occur with autophosphorylation occurring on a Ser within the AID that prevents Ca^{2+}/CaM binding and activation. This inhibitory mechanism is comparable with CaMKK and CaMK2. Dephosphorylation of CaMK4s is mediated by protein phosphatase 2A (PP2A).

CaMK4 is predominantly expressed in the brain but is also expressed in hematopoietic stem cells, progenitor T cells, testis, and the ovaries (Hook and Means 2001). Within the cell CaMK4 resides in the cytosol and nucleus but information about its targets within the cytosol is limited (Jensen, Ohmstede et al. 1991). Contrastingly, CaMK4 nuclear substrates are better characterised. For example it is well established that CaMK4 is involved in Ca^{2+} dependent gene transcription and demonstrates activity towards CREB and CREB binding protein (CBP) (Enslen, Sun et al. 1994; Matthews, Guthrie et al. 1994). Other transcription factors that CaMK4 can phosphorylate include serum response factor (SRF) and histone deactylase 4 (HDAC4).
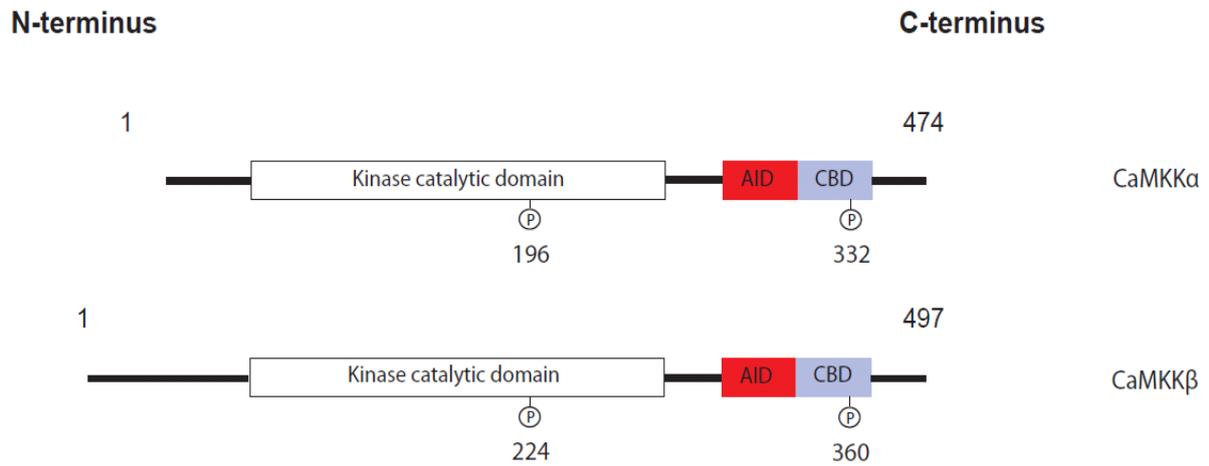
Figure 1.12. CaMK4 domain organisation.

CaMK4 isoforms feature a KD adjacent to a C-terminal RD characteristic of CaMKs. In the RD an inhibitory phosphorylation site prevents Ca^{2+}/CaM binding therefore down regulating activity whereas activation loop phosphorylation promotes activity. CaMKIVβ activation loop phosphorylation leads to autophosphorylation within the N-terminus and results in Ca^{2+}/CaM independent activity (Swulius and Waxham 2008).

The autoinhibitory domain

All CaMK1 members are maintained in an inactive state due to the regulatory AID. The domain boundaries of the AID and the overlapping CBD for CaMKIs were first defined through serial C-terminal truncations of CaMK1A which decreased the ability of CaMK1A to bind to Ca^{2+}/CaM and decreased Ca^{2+}/CaM dependent activity (Hook and Means 2001). Deletion analysis beyond the CBD boundary up to residue Phe 298 (CaMK1A_1-298), revealed that this construct was still autoinhibited, however an additional single residue truncation of Phe 298 up to Asn 297 (CaMK1A1-297), generated a constitutively active construct, thereby suggesting the importance of Phe 298 in mediating autoinhibition (Hook and Means 2001). Although this construct exhibited Ca^{2+}/CaM independent activity it was reported to be only 70% active in comparison to the activated wild type CaMK1A suggesting that residues N-terminal of Asn 297 remained that could potentially exert an autoinhibitory effect. Similarly truncation of another construct up to Val 294 was also found to be constitutively active (Hook and Means 2001). However, in the context of full length CaMK1A, mutagenesis of Phe 298, and Val 294 to Ala only resulted in a small increase in basal levels of Ca^{2+}/CaM independent activity, thus implying the role of other residues mediating autoinhibitory contacts. For example, Phe 307 was found to impart an autoinhibitory effect because mutagenesis of this residue to Ala resulted in Ca^{2+}/CaM independent activity that was 30% active in comparison to the activated wild type (Hook and Means 2001).

The pseudosubstrate hypothesis

The general model for the autoinhibition of CaMK family members was based on a pseudosubstrate hypothesis derived from the observation that Myosin light chain kinase (MYLK) possessed a sequence within its AID with similarity to the recognition motif in its substrate Myosin light chain. From the pseudosubstrate hypothesis, it was postulated that in the absence of Ca^{2+}/CaM, the AID would occupy the substrate binding site. Upon cleavage of the AID, MYLK was found to exhibit Ca^{2+}/CaM independent activity (Kemp, Pearson et al. 1987; Pearson, Wettenhall et al. 1988). Autoinhibition of MLCK was restored upon addition of a peptide corresponding to the AID of MYLK. This peptide could also bind to Ca^{2+}/CaM thus reinforcing the concept

of an overlapping AID and CBD within the regulatory domain. Other evidence consistent with the pseudosubstrate hypothesis was exemplified by the crystal structure of Twitchin, which had a 60 residue AID segment comprised of a helix and loop that blocked the substrate binding site and active site (Hu, Parker et al. 1994).

In the context of CaMK1s, Yokokura et al demonstrated that a peptide corresponding to the AID of CaMK1A could competitively inhibit CaMK1A with respect to Ca^{2+}/CaM. Similarly a constitutively active CaMK1A construct was also inhibited by the peptide AID and showed competitive inhibition with respect to the substrate peptide syntide-2 but was non-competitive with ATP, thus implying that the peptide AID bound to the substrate binding site of the constitutively active CaMK1A. Sequence analysis and alignment of the CaMK1A AID and CaMK1A substrates such as synapsin-1, syntide-2, and CREB reveals key basic and hydrophobic residues within the AID that satisfy the substrate consensus sequence of CaMK1s of Hyd-X-R-X-X-S/T-X-X-X-Hyd (where Hyd is any hydrophobic residue and X is any residue). Collectively the results suggested that CaMK1A was regulated by a pseudosubstrate mechanism (Yokokura, Picciotto et al. 1995).

Prior to the first structure determination of a CaMK1 member, it was presumed that the pseudosubstrate model of autoinhibition observed in other CaMKs held true for CaMK1s, however the structural basis for CaMK1 autoinhibition was not as simple as this. Back in 1996 the first apo CaMK1 structure to be solved was the rat form of CaMK1A, a construct which included the kinase catalytic domain and the AID (Goldberg, Nairn et al. 1996). Autoinhibited rat CaMK1A exhibited the typical bilobal kinase fold comprising of an N-terminal β-sheet lobe and a larger C-terminal α-helical lobe and significantly revealed the structure of the AID to be composed of a helix-loop-helix segment that folded over both the C-terminal and N-terminal lobes (Figure 1.13). The first and second helices of the AID were designated as αR1 and αR2 respectively and shown to mediate CaMK1A autoinhibition through distinct mechanisms. Helix αR1 and some parts of the AID loop occupy and occlude the substrate binding site through intramolecular interactions characteristic of a pseudosubstrate mechanism.

This pseudosubstrate interaction is mainly driven by hydrophobic packing with residues from α helix D. Notably near the C-terminal end of helix αR1 the side chain of Phe 298 which mimics the P-5 position, appears buried inside a hydrophobic pocket formed by Phe 104, Ile 210, and Pro-216 within the substrate binding site. N-terminal of Phe 298, further extensive hydrophobic packing is evident with helix αR1 residues Ile 286, Val 290, and Ile 294, all packing against a hydrophobic interface from the C-terminal lobe of the kinase catalytic domain (Goldberg, Nairn et al. 1996). Interestingly these autoinhibitory contacts mediated by Ile 286, Val 290, and Val 294 have been mutated by Matsushita and co-workers which resulted in an increased basal level of Ca^{2+}/CaM independent activity (Matsushita and Nairn 1998). Although the pseudosubstrate contacts in CaMK1A are predominantly hydrophobic, electrostatic interactions are also evident. The electrostatic contribution to the pseudosubstrate mechanism is mediated by Lys 300, a residue located in the loop region of the helix-loop-helix segment. This residue interacts with Glu 102 most likely by mimicking the P-3 position because mutation of Glu 102 reportedly alters the P-3 specificity of CaMK1A (Goldberg, Nairn et al. 1996).

The first mechanism of autoinhibtion described in CaMK1A satisfies the pseudosubstrate hypothesis that was predicted to occur in CaMK1s. A second mechanism is also exemplified in CaMK1A by the latter part of the AID which imparts additional autoinhibitiory effects beyond the substrate binding site of the C-terminal lobe. Notably helix αR2 forms autoinhibitory interactions with the N-terminal ATP binding lobe resulting in the disruption of its conformation and ability to bind ATP. This intramolecular interaction is mediated by hydrophobic contacts between hydrophobic residues that span helix αR2 and the interacting interface on the N-terminal lobe. Residues from helix αR2 which are implicated in this include Ala 306, Phe 307, Ala 309, Ala 312, Val 313, and Met 316, whilst those which contribute to the N-terminal lobe hydrophobic interface include Leu 26, Gly 27, Phe 31, Leu 36, Leu 45, Leu 97, Ser 99, and Gly 101 (Goldberg, Nairn et al. 1996).

Due to the interactions between helix αR2 and the N-terminal ATP binding lobe, conformational changes occur within the ATP binding loop (also known as the p-

loop) associated with the N-terminal lobe. One notable localised contribution to this is attributed to a conserved Phe 31 within the respective loop which engages in van der vaal interactions with Phe 307 from helix αR2. Structural comparison of this loop region in autoinhibited CaMK1A with active PKA indicates that the side chain of Phe 31 occludes the position normally occupied by the ribose moiety of bound ATP as exemplified in PKA. Phe 54, the equivalent residue in PKA functions to cap the β and γ phosphate groups of bound ATP from solvent exposure but due to the inactive conformation of the CaMK1A N-terminal lobe induced by helix αR2, Phe 31 is unable to facilitate this. Moreover, Phe 54 is part of a group of hydrophobic residues suggested to play a role in maintaining the active state of PKA that includes an active activation loop (Goldberg, Nairn et al. 1996). The AID thus mediates autoinhibition on CaMK1A and presumeably other CaMK1s by inducing the ATP binding site into an inactive open conformation and occludes substrate binding via the pseudosubstrate mechanism.
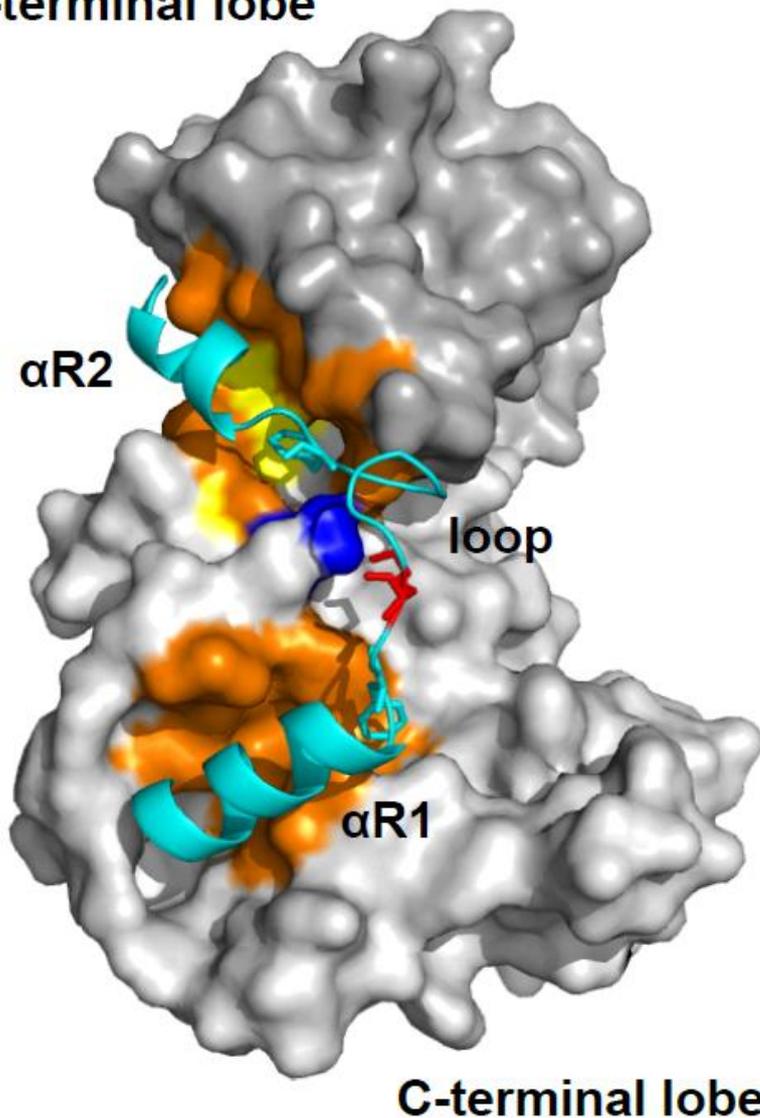
Figure 1.13. Surface and cartoon representation of the AID pseudosubstrate model in rCaMK1A.

The AID (cyan) mediates autoinhibition of rat CaMK1A by occupying the substrate binding site in the C-terminal lobe (light gray) via the αR1 helix and by interacting with the N-terminal ATP binding lobe (dark gray) via the αR2 helix. The pseudosubstrate interactions are predominantly hydrophobic (orange). An electrostatic interaction in the AID loop is also highlighted between (blue) and red (red). The figure was adapted from the rat CaMK1A structure determined by Goldberg and co-workers (Goldberg, Nairn et al. 1996).

CaMK1s and Ca^{2+}/CaM in the cell cycle

Cell proliferation is associated with an increase in intracellular Ca^{2+} concentration and Ca^{2+}/CaM is implicated in the cell cycle via the activation of downstream CaMKs (Kahl and Means 2003). The involvement of CaMK1s in cell cycle progression has been demonstrated in various studies. For example, treatment of fibroblasts expressing CaMK1 and CaMK2 with KN-93, a specific inhibitor of CaMK1, CaMK2, and CaMK4 caused cells to arrest in the late G1 phase, preventing activation of the cell cycle regulator cyclin D1/Cdk4 and phosphorylation of pRB. Both are necessary for cell cycle progression from G1 to S (synthesis) phase. A CaMK1 deficient cell line confirmed that cell cycle arrest was due to the inhibition and inactivation of CaMK1 and not CaMK2. Moreover CaMK2 deficient cells did not arrest (Kahl and Means 2003). Human or mouse cells cultured in media containing low Ca^{2+} concentrations cease cell division and accumulate in the G1 phase. CaM inhibitors W-7 and W-13 inhibit proliferation and colony formation in breast cancer cells and other cell lines.

CaMK1s in disease

Deregulation of the cell cycle can result in uncontrolled cell proliferation and therefore cancer. Rodriguez-Mora et al using the tumorgenic breast cancer cell line MCF-7 tested the effects of CaMK inhibition on its cell cycle progression and proliferation. Results from this study indicated that CaMKK and CaMK1 are required for cell cycle progression in the MCF-7 cell line. In this case KN-93 inhibited MCF-7 cell proliferation by arresting cells in the G1 phase but did not induce apoptosis. A similar result was acquired using CaMK1 siRNA and CaMKK siRNA. Interestingly expression levels of cyclin D1 exhibited reduced expression in MCF-7 following treatment with KN-93 suggesting that CaMK1 inhibition also causes an inhibition in cyclin D1 expression. The effect of CaMK1 inhibition on cell proliferation in the non-turmogenic cell line MCF-10 showed a reduced inhibitory effect on cell proliferation in comparison to the tumorgenic MCF-7. This difference was accounted for in terms of the differential expression of CaMK1 between the cell lines, with expression of CaMK1 being greater in MCF-7 than MCF-10 (Rodriguez-Mora, LaHair et al. 2005; Rodriguez-Mora, Lahair et al. 2006).

The role of CaMKs in cancer also has been highlighted in endometrial carcinomas in which over expression of CaMK1, CaMK2, and CaMK4 was found to correlate with the malignant potential of endometrial carcinomas. Their results demonstrated that endometrial carcinomas treated with KN-93 induced cell cycle arrest and down regulated expression of cyclin D1 as exemplified in previous reports by Rodriguez-Mora et al with MCF-7. Moreover it also induced apoptosis which correlated with the down regulation of the anti-apoptotic bcl-2 and up regulation of the pro-apoptotic caspase-3. Sub cellular localisation of CaMKI in endometrial carcinomas is predominantly cytosolic with some presence in the nucleus.

Interestingly CaMK1D has been implicated in breast cancer (Bergamaschi, Kim et al. 2008) in which it was found to be up regulated from a microarray of breast cancer tumors defined as a basal-like subtype. This cancer type exhibited an epithelial-mesenchymal transition phenotype which is characterised by increased cell proliferation, loss of cell-cell adhesion, and increased migration and invasion indicative of metastasis. Engineered CaMK1D expression in the immortalised but non-turmorgenic cell line MCF-10A demonstrated increased cell growth versus the vector-alone control cells. In addition, CaMK1D increased the S phase fraction therefore suggesting increased cell cycle progression. Evidence of CaMK1D promoting metastasis was suggested from decreased expression levels of E-cadherin and increased levels of Vimentin in the MCF-10A cell line. Imunofluorescence studies illustrated a positive correlation between the expression of CaMK1D and Vimentin in the cell line. Furthermore, a wound healing assay revealed that CaMK1D expression enhanced cell migration.

Calcium signalling

Ca^{2+} is one of the most important and widely used secondary messengers within the cell being implicated in many cellular processes such as cell cycle regulation, cell proliferation, apoptosis, gene expression, cell differentiation, muscle contraction and hormone signalling. However, Ca^{2+} is also highly toxic to the cell if the intracellular concentration is high and exposure is abnormally prolonged. Regulation of

intracellular Ca^{2+} concentrations is therefore essential for normal cellular function (Berridge, Bootman et al. 1998). The signalling effects of Ca^{2+} are mediated through controlled increases in its cytosolic concentration in a temporal and spatial manner. Cells utilise Ca^{2+} derived from the extracellular environment and from intracellular stores such as the endoplasmic reticulum, sarcoplasmic reticulum, mitochondria and perinuclear space (Berridge, Bootman et al. 1998). Typically the cytosolic Ca^{2+} concentration is kept low around 100 nM during the basal state of the cell to prevent constitutively aberrant signalling and to ensure that a Ca^{2+} concentration gradient is maintained across the cytosol and the intracellular and extracellular Ca^{2+} stores. ATPase transporters actively pump cytosolic Ca^{2+} into the extracellular and intracellular stores to maintain a Ca^{2+} concentration gradient. Exchangers and buffers also help maintain a low cytosolic Ca^{2+} concentration (Figure 1.14). Upon cell stimulation this gradient permits a rapid influx of Ca^{2+} into the cytosol whereby the cytosolic Ca^{2+} concentration can increase up to $1 - 2$ µM (Clapham 2007).

In excitable cell types like neurons and muscle, Ca^{2+} can enter the cell and into the cytosol from the extracellular store through voltage operated channels (VOCs or voltage gated channels) and receptor operated channels (ROCs or ligand gated channels). In non excitable cells, Ca^{2+} entry from the extracellular store is facilitated via store operated channels (SOCs), which are opened following the depletion of intracellular Ca^{2+} stores. Entry of Ca^{2+} from intracellular stores into the cytosol is mediated by activation of the IP_3 and Ryanodine receptors. The IP_3 receptor is activated by inositol 1,4,5 triphosphate (IP_3) which is produced by Phospholipase C (PLC) as a cleavage product from phosphatidylinositol 4,5 bisphosphate (PIP_2). Activation of PLC itself is mediated by the activation of upstream G-protein coupled receptors (GPCRs) or receptor tyrosine kinases (RTKs) by external stimuli such as growth factors or GPCR agonists (Figure 1.14).

Activation of the Ryanodine receptor is mediated by an increase in the cytosolic Ca^{2+} concentration in a process known as Ca^{2+} induced Ca^{2+} release (CICR). Subsequently more Ca^{2+} is released via the Ryanodine receptor from the intracellular store into the cytosol which initiates a positive feedback response with

the Ryanodine receptor. Activation of the Ryanodine receptor is thus coupled to the activation of VOCs, ROCs, SOCs, and the IP3 receptor because these channels and receptors can mediate an increase in the cytosolic Ca^{2+} concentration.

Influxes of Ca^{2+} into the cytosol occurs in transient aliquots described as sparks, waves, or puffs, depending on the intracellular area covered. Ca^{2+} waves can also occur in an intercellular manner as means of coordinating the signalling effect globally. Moreover cells can initiate repetitive Ca^{2+} signalling termed oscillations and by modulation of the amplitude and frequency of the Ca^{2+} oscillations it allows cells to encode diverse signalling information and therefore activate a variety of signalling responses. It has been reported for example that the expression efficiency and specificity of some genes are dependent on an optimal intracellular Ca^{2+} frequency (Dolmetsch, Xu et al. 1998).
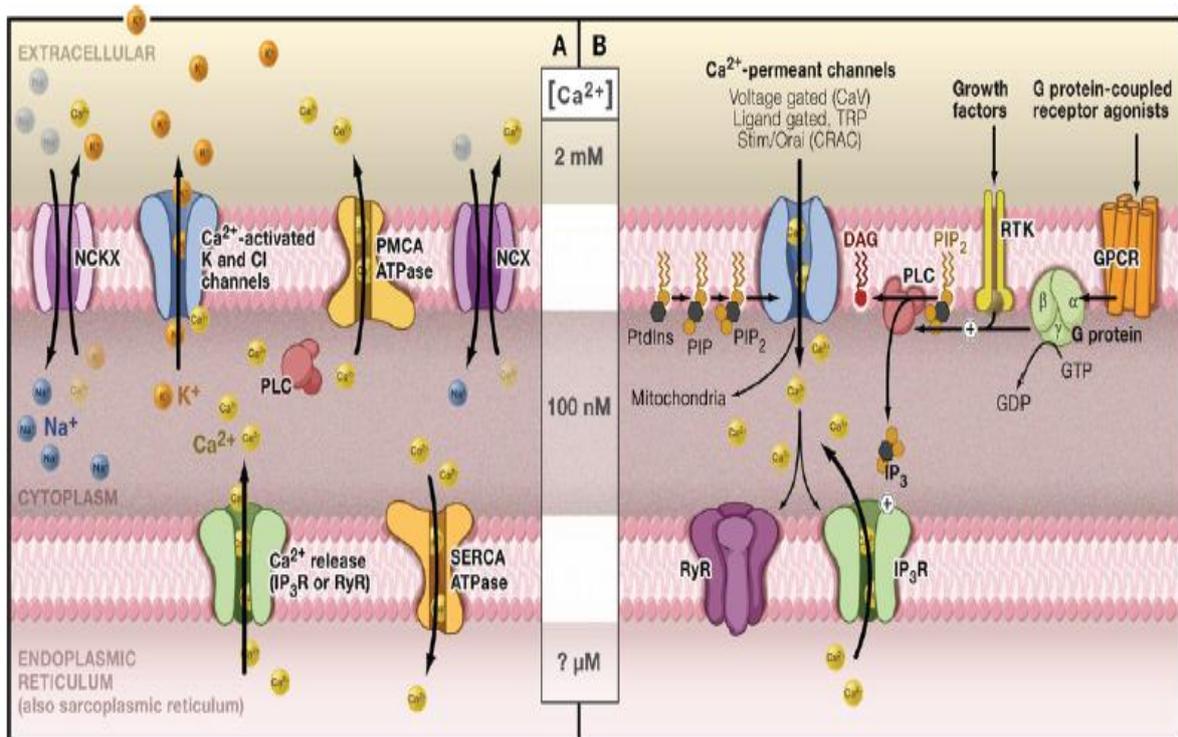
Figure 1.14. Calcium signalling and concentration gradients.

Upstream signalling events of GPCRs and RTKs can initiate downstream Ca^{2+} signalling. Transporters and exchangers help to maintain the concentration gradient (Clapham 2007).

Calmodulin

Ca^{2+} secondary messenger effects are mediated through Ca^{2+} binding proteins that can either directly or indirectly initiate downstream signalling effects and cellular processes. Protein kinase C (PKC) for example is a direct mediator of Ca^{2+} signalling by binding to Ca^{2+} via its C2 domain. Indirect mediators of Ca^{2+} signalling include proteins that constitute the E-F hand motif family, a structural motif that enables Ca^{2+} binding. Generally there are two main types of the E-F hand motif proteins. One type binds to Ca^{2+} but do not undergo any significant conformational changes and are considered to function as Ca^{2+} buffers or transporters. The second type bind Ca^{2+} and experience conformational changes and are known as Ca^{2+} sensors. These sensor proteins couple Ca^{2+} signalling to many biological processes via modulation of downstream substrate proteins.

One of the main decoders of Ca^{2+} signalling is calmodulin (CaM), a highly conserved and ubiquitous Ca^{2+} sensor protein with a molecular weight of approximately 17k Da. In a Ca^{2+} dependent manner is an allosteric activator of many proteins that possess the CBD. In terms of structure, CaM comprises of an N-terminal and C-terminal globular domain connected by a flexible helical linker with an overall fold that ressembles a dumbell (Figure 1.15). Each globular domain contains two E-F hand motifs that feature a helix-loop-helix motif and are capable of binding Ca^{2+}. In total CaM can bind up to four Ca^{2+} ions with one Ca^{2+} associated with each of the four E-F hand motifs (Chin and Means 2000). Six to seven negatively charged oxygen atoms of the carbonyl and carboxyl groups within the motif coordinate Ca^{2+}.

The N-terminal Ca^{2+} binding sites are denoted as sites I and II, whilst the C-terminal sites are III and IV. Although the N-terminal and C-terminal globular domains are homologous, the C-terminal pair of E-F hand motifs have 10-fold higher affinity for Ca^{2+} (Hook and Means 2001). As a Ca^{2+} sensor protein, the affinity of CaM for Ca^{2+} is within the cytosolic concentration range (10^{-6}M to 10^{-7}M) thus enabling it to respond to the relevant changes in cytosolic Ca^{2+} concentrations. Ca^{2+} binding is a cooperative event. It has been suggested that Ca^{2+} might bind preferentially to site III

first. However, have also shown that Ca^{2+} binds to sites III and IV simultaneously (Hook and Means 2001).

Upon transient increases in the intracellular Ca^{2+} concentration, Ca^{2+} binds to one of the CaM E-F hand motifs and causes the helices of the E-F hand motif to become pulled closer together. This Ca^{2+} induced conformational change initiates a cooperative binding event that is transduced through a β-sheet connected to the neighbouring E-F hand motif. Consequently there is an increase in its Ca^{2+} binding affinity (Hook and Means 2001). Subsequent binding of another Ca^{2+} ion at the neighbouring E-F hand motif causes both E-F hand motifs to pull further apart and induce the exposure of a hydrophobic pocket within the globular domain. This enables CaM to bind to the CBD of its substrate targets. Furthermore, the presence of a substrate target peptide can enhance the cooperative binding event whilst the presence of substrate proteins has also been reported to increase its affinity for Ca^{2+} by up to 10-fold (Peersen, Madsen et al. 1997).

Global conformational changes also occur in which the helical content of CaM is increased. Most notably this involves the formation of an eight turn α helix denoted as helix IV (Chin and Means 2000; Hook and Means 2001) from the previously flexible linker connecting the globular domains of the CaM apo state. This helical linker facilitates the binding of CaM to many diverse substrates harbouring the CBD via its conformational plasticity. In the Ca^{2+} bound state, the globular domains adopt open conformations which results in the exposure of hydrophobic groups that form a hydrophobic crevice. Four conserved Met residues are contained within these hydrophobic crevices and contribute towards important interactions with substrates. Both the N-terminal and C-terminal crevices form a hydrophobic tunnel in which the CDB of substrates traverse. Generally the CBD region of substrates exhibits an amphipathic helix conformation that traverses the hydrophobic tunnel (Hook and Means 2001). Each of the globular domains can also adopt a different conformation in the absence of Ca^{2+} with the N-terminal domain forming a closed conformation whilst the C-terminal domain takes on a semi open conformation which results in the partial solvent exposure of a hydrophobic patch. It has been suggested that this

enables CaM to interact with some substrate targets whilst in its apo state at resting Ca^{2+} concentrations (Swindells and Ikura 1996). CaM substrate targets are capable of exhibiting various binding modes with CaM.

CaM binding proteins can be categorised into six groups according to their mode of regulation by CaM in the presence and absence of Ca^{2+} (Chin and Means 2000). One group are capable of binding to CaM irreversibly regardless of the Ca^{2+} concentration and are typified by phoshporylase kinase. A second group describes proteins that bind to CaM in the absence of Ca^{2+} but dissociate in the presence of Ca^{2+}. Such proteins are exemplified by neuromodulin and neurogranin which are considered to function as intracellular reservoirs for CaM at the basal Ca^{2+} concentration until transient increases in the Ca^{2+} concentration causes the intracellular reservoir proteins to liberate Ca^{2+} bound CaM. A third group represents proteins that can interact weakly with apo CaM or CaM that is only partially bound to Ca^{2+} to form an inactive complex with CaM. However, upon an increasing Ca^{2+} concentration, the low affinity complex becomes a high affinity active complex as CaM becomes saturated with Ca^{2+}. Some proteins that bind to Ca^{2+} saturated CaM form an inactive complex and this represents the forth group of CaM regulation. Proteins of this group include for example, GPCRs, and the IP3 receptor type 1. Conversely there are proteins that interact with Ca^{2+} saturated CaM and subsequently form an active high affinity complex as demonstrated by the CaMKs. This describes the most conventional CaM mode of regulation and is categorised as the fifth group. Lastly the sixth group of CaM regulated proteins describes those such as CaMKs which exhibit the same mode of regulation as the previous fifth group but also includes further indirect activation of the protein via phosphorylation by an upstream CaMK which is also activated by Ca^{2+} saturated CaM. In the case of CaMKII, phosphorylation is mediated by autophosphorylation upon Ca^{2+}/CaM binding.
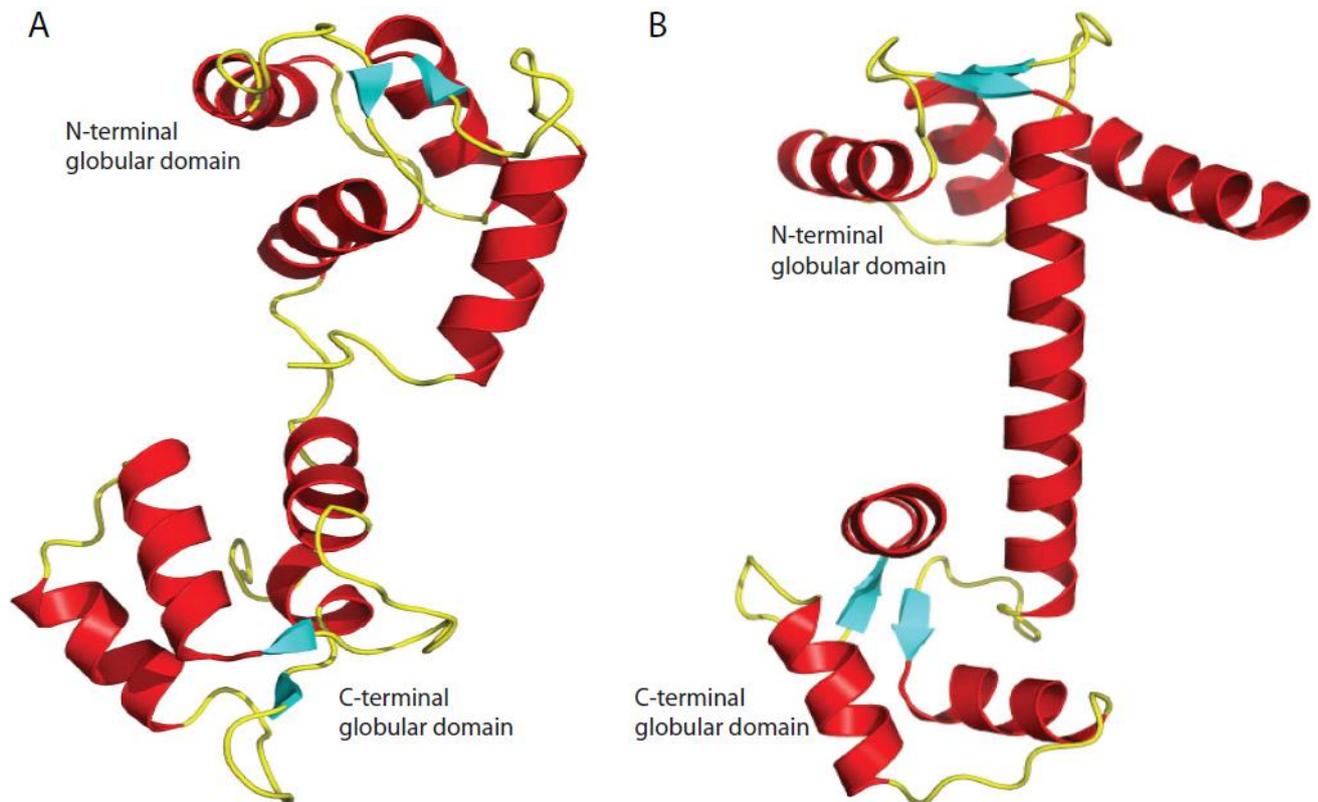
Figure 1.15. The structure of apo CaM and Ca^{2+} bound CaM.

The cartoon structure of CaM is represented by α helices, loops, and β strands shown in red, yellow, and cyan respectively. On the left side showing apo CaM the N-terminal globular domain adopts a closed conformation with the E-F hand motifs packed close together whilst the C-terminal CaM globular domain takes on a semi open conformation. On the right hand side showing Ca^{2+}/CaM both globular domains adopt the open conformationand coupled with the exposure of their respective hydrophobic pockets and the formation of a helical linker.

Isothermal Titration Calorimetry

Binding constants and thermodynamic parameters

Isothermal titration calorimetry (ITC) is a technique that enables determination of the binding affinity (association constant, K_a), enthalpy change (ΔH), entropy change (ΔS), and stoichiometry (N) associated with a chemical reaction at a constant temperature. Typically this involves an interaction between ligand L and macromolecule M to form complex ML. Both ΔH and ΔS are two thermodynamic parameters which can be used to define K_a and are related to the Gibbs free energy of binding (ΔG) according to equations [1], [2], and [3], shown below and using these terms the K_A can be deduced (Wiseman, Williston et al. 1989; Leavitt and Freire 2001).

[1] $\Delta G = -RT\ln K_a = \Delta H - T\Delta S,$

[2] $K_a = e^{(-\Delta G/RT)},$

[3] $K_a = e^{(-\Delta H - T\Delta S/RT)},$

(R= gas constant, T = temperature)

ΔH reflects the strength of non-covalent interactions between L and M relative to those existing with the solvent and describes the heat changes associated with processes that are coupled with the interaction. These processes include binding, protonation, ionisation, and conformational changes for example. ΔS reflects the changes in solvation entropy and changes in conformational entropy associated with the interaction. ΔG describes whether a reaction is a spontaneous (when ΔG is negative, reaction occurs without the addition of energy into the system from surroundings) or non spontaneous (when ΔG is positive, system requires addition of energy from surroundings in order for reaction to occur) process and can indicate the affinity of binding (Leavitt and Freire 2001).

K_a describes the binding affinity for a binding interaction between L and M. Whereas the inverse of K_a, the dissociation constant (K_d) describes the strength of binding between L and M in terms of the tendency for the ML complex to separate and is defined as the concentration of L at which half of the M bindings sites are occupied by L to form ML. When L binds reversibly to a macromolecule M with a single binding site as shown by equation [4], the association and dissociation constants can be described by equations [5] and [6] respectively. Current ITC instruments can determine K_d values that range from 0.01-100 µM (Ladbury and Chowdhry 1996; Ladbury 2007).

[4] $[M] + [L] \leftrightarrow [ML]$,
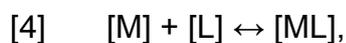
[5] $K_a = [ML] / [M] [L]$,

[6] $K_d = [M] [L] / [ML]$,

Principle of the ITC experiment

The ITC instrument is comprised of a sample cell and reference cell contained in an adiabatic jacket (Figure 1.15). Both cells are electronically linked to each other and to the jacket via thermopile circuits which detect temperature differences between the cells and with respect to the jacket. A feedback system ensures that any detected temperature difference between the cells is minimised and that they are maintained at the experiment temperature. The feedback system in ITC operates on a heat compensation principle in which the directly observed feedback signal in an experiment is the heat energy per unit time ($\mu cal\ s^{-1}$) which must be added or removed from the sample cell in order to maintain zero temperature difference between the sample cell and reference cell at the designated experiment temperature. The temperature of the jacket is typically cooled 5-10 K below the experiment temperature therefore the cells are heated in order to maintain the cells at the experiment temperature. Prior to the experiment this serves as the baseline signal. During an ITC experiment a specified volume of ligand L (i.e. Ca^{2+}/CaM) is injected from the syringe into the sample cell containing the macromolecule M (i.e.

CaMK1D) and is constantly mixed (Pierce, Raman et al. 1999; Ladbury 2007). Each injection results in either a release or absorption of heat energy (q_i) proportional to the amount of L ($V\Delta L_i$, where V is the cell volume and ΔL_i the increase in concentration of bound ligand after i^{th} injection) that binds to M for that injection and the ΔH for the interaction as described below in equation [7] (Leavitt and Freire 2001).

[7] $q_i = V\Delta L_i \Delta H$

In response to each injection the heat change is measured with respect to the reference cell temperature and the feedback signal indicates the power ($\mu cal\ s^{-1}$) that is added (if the heat change is endothermic) or removed (if the heat change is exothermic) from the sample cell in order to maintain the zero temperature difference between the sample and reference cell at the designated experiment temperature. The area under each peak that the feedback signal produces correlates with the heat change associated for that particular injection. During the first few injections, L is completely bound to M thus producing large heat changes but subsequent injections result in fewer available M binding sites with smaller heat changes. Eventually M binding sites are saturated and the heat changes that follow are caused by dilution and mixing effects. If titrations are repeated at different temperatures it is possible to obtain the heat capacity (ΔC_p) for the interaction which describes the temperature dependence of ΔH (Pierce, Raman et al. 1999; Ladbury 2007).

The c value

When designing an ITC experiment an important parameter that must be considered is the c value which has no units and is related to the product of stoichiometry N, association constant K_a, and the total macromolecule concentration $[M]_T$ as shown below in equation [8].

[8] $c = NK_a[M]_T$

In order for an accurate determination of the K_a, the c value should be between 1 and 1000 in order to produce an ideal sigmoid shaped binding isotherm (Wiseman, Williston et al. 1989; Pierce, Raman et al. 1999). However, if the c value is too large because the concentration of M_T binding sites is much greater than $1/K_a$ then all of L will be bound to M until saturation which results in a binding isotherm rectangular in shape with a sharp slope that has few points. Conversely if the c value is too small because the concentration of M_T binding sites is much lower than $1/K_a$ then full saturation of M is difficult to achieve and consequently the binding isotherm appears shallow. In both cases where the c value is not optimal determination of the K_a therefore becomes difficult, but by adjusting the M_T concentration, the c value can also be adjusted.

In the case of weak interactions where K_a is small this will require a higher concentration of M_T and 10 to 20-fold excess concentration of L with respect to M_T in order to achieve saturation of M_T but this will depend on the availability and solubility of M_T and L at their respective higher concentrations. For strong interactions where K_a is large and the saturation of M_T is easier to achieve (i.e. CaMK1D and Ca^{2+}/CaM), this requires a lower concentration of M_T but will depend on how large the heat change is at lower concentrations of M_T because if the measured heat changes are smaller than the sensitivity limits of the ITC instrument then the real heat changes associated with the interaction cannot be distinguished from unreal heat changes due to dilution and mechanical effects of stirring.
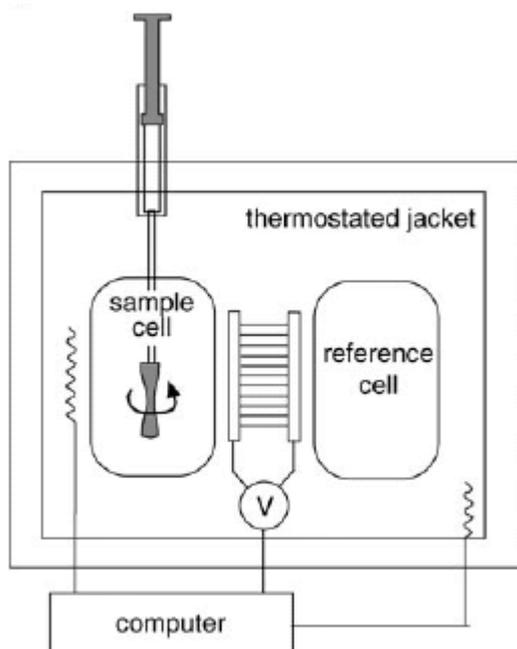
Figure 1.15. ITC instrument.

The ITC instrument setup is made up of several components. A sample cell containing the macromolecule M, and a reference cell containing buffer or water. Both cells are housed inside an adiabatic jacket that is maintained at the chosen experiment temperature. The syringe contains the ligand L, and also functions as a stirrer. The cells are cooled 5-10 K below the chosen experiment temperature therefore an input of heat energy is used to maintain the cells at this temperature. Thermopile circuits detect changes in heat energy between the cells and also within the jacket then convert the changes into electrical energy and convey them to the computer in order to operate a feedback system that minimises the changes in heat energy between the sample cell and reference cell and also ensure that the jacket temperature is maintained at the experiment temperature. The output signal from the ITC instrument displays the increase or decrease in power ($\mu cal\ s^{-1}$) that is necessary to maintain the zero temperature difference between the sample and reference cell at the designated experiment temperature (Velazquez-Campoy, Ohtaka et al. 2004).
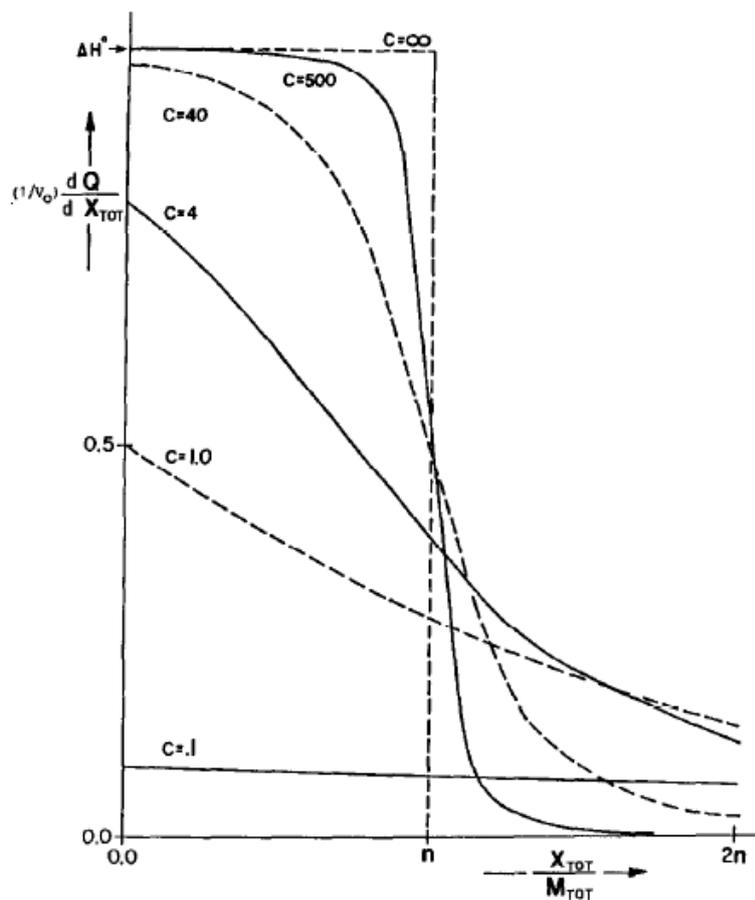
Figure 1.16. Binding isotherms for different C values.

The binding isotherms for different c values are shown to illustrate the how the shape of the curve is dependent on the c value which is related to the product of K_a, $[M]_T$ and N. To determine the K_a, the c value should be between 1 and 1000 by adjusting, $[M]_T$. The y axis and x axis correspond to $(1/V_0)dq/dx$ (kcal mol^{-1}) and $[L_T]/[M_T]$. V_0 denotes the volume of the sample cell (Wiseman, Williston et al. 1989).

Nuclear Magenetic Resonance

Nuclear magnetic resonance (NMR) spectroscopy is a technique that can determine the solution structure of proteins at atomic resolution using isotope labelled proteins. In addition NMR can provide information about interactions, conformational changes and dynamics (Wuthrich 2003). NMR is complementary to protein crystallography which is another technique for high resolution structure determination. However protein crystallography requires proteins that are receptive to crystallisation which can present a major limiting step. In contrast NMR requires isotope labelled proteins in solution. An advantage with protein crystallography is that any protein size can be studied whereas NMR is often limited to proteins no greater than 30 kDa due to resonance line shape broadening and overlap. Protein structures that are determined by protein crystallography only represent a snap shot of a particular conformation that is amenable to crystallisation and may not be reflective of the native protein conformation. This can be due to the influence of crystallisation artefacts such as crystal contacts which can exert bias towards certain protein conformations. Protein structures determined by NMR represent an ensemble of the conformations that the protein can sample in solution and there are no artificial contacts, consequently the NMR protein structure is more likely to resemble the native conformation. To appreciate the application of protein NMR and its advantages and disadvantages, the relevant NMR theory will be discussed in the subsequent sections.

Nuclear spin

Atomic nuclei possess an intrinsic property termed spin that can give rise to angular momentum (P). Spin is characterised by the nuclear spin quantum number (I), also termed spin I, which describes the shape of the magnetic field and can adopt either zero, integral (i.e. 1, 2, 3 etc) or half integral (i.e.1/2, 3/2, 5/2 etc) values depending on the proton and neutron content of the nucleus. Only integral and half integral values result in nuclei with a net spin and an associated magnetic field. These types of nuclei are termed spin active because they are amenable to NMR. Spin active nuclei are associated with a magnetic field described as the magnetic moment (μ) because nuclei have charge and in combination with the motion of net spin this

generates a local magnetic field. Both the μ and P of a spin active nucleus are related to each other by the magnetogyric ratio (γ), a constant that describes the ratio between μ and P for a particular nucleus and is an indicator of its magnetic strength. In addition μ can be parallel (when γ is positive) or anti-parallel (when γ is negative) to P. The relationship between these terms is: $\mu = \gamma P$.

Spin 1/2 nuclei

P is a vector quantity and therefore has both magnitude and direction. Consequently the P of a nucleus characterised by spin I has direction defined in terms of 2I+1 number of possible orientations which it can adopt with respect to an arbitrary chosen axis. By convention this is defined as the z-axis, thus the z-component of P is denoted as P_z and is quantized as: $P_z = \hbar m$, where $\hbar$ is planks constant in terms of Joules radians^{-1}.Furthermore the direction of the 2I+1 allowed orientations are described by the magnetic quantum number (m) which has 2I+1 values that are between +I and -I. For example nuclei of interest in protein NMR include ^{1}H, ^{15}N and ^{13}C, which are spin 1/2 nuclei and therefore restricted to two orientations (2*1/2+1 = 2) described by m as spin +1/2 and spin -1/2. Permitted directions of the P_z component for spin 1/2 nuclei can thus be represented as: $P_z = +1/2\hbar$, $P_z = -1/2\hbar$, or $P_z = \pm 1/2\hbar$

α and β states

In the absence of an external magnetic field (B_0), the P of spin I nucleus has no preference for any of its allowed orientations. Moreover, these orientations are energetically equivalent but when B_0 is applied this degeneracy is removed because the associated μ of a spin I nucleus enables it to interact with B_0 in 2I+1 ways and possess different energy states. Spin 1/2 nuclei for example can align in two allowed orientations that have distinct energy states depending on the alignment of spin 1/2 nuclei with respect to B_0. Alignment of spin 1/2 nuclei can be parallel (denoted +1/2) or anti-parallel (denoted -1/2) to B_0 which correlates to a lower energy state (denoted α) and higher energy state (denoted β) respectively. Spectroscopy effectively

involves measuring $\Delta E = h\nu$ (h is planks constant, ν is frequency) associated with the transition between distinct energy states. Therefore in NMR the transitions between α and β states of spin 1/2 nuclei can be induced by applying an appropriate ν of radiation that corresponds to the precession ν of spin 1/2 nuclei also known as the Larmor ν. Since ΔE and ν are related by h, Larmor ν therefore correlates to the ΔE of the α and β transition, whilst Larmor ν itself is determined by the nuclei γ and B_0, therefore: $\Delta E = h\nu = h\gamma B_0 / 2\pi = \hbar\gamma B_0$

Boltzmann distribution and sensitivity

Low frequency radiation is used in NMR to induce transitions because the ΔE between α and β energy states is small (since $\Delta E = h\nu$). At thermal equilibrium the population difference between the α and β states is described by the Boltzmann equation: $N\alpha / N\beta = e^{(\Delta E / RT)}$ which indicates that ΔE for the transition is small due to the small population difference between α and β states. Given that the net absorption of ΔE correlates to the intensity of the transitions, the sensitivity of NMR is thus relatively low because the ΔE of the transitions studied is also low. However due to the relationship: $\Delta E = \hbar\gamma B_0$, increasing the B_0 strength will increase the transition ΔE and improve sensitivity (Figure 1.18).
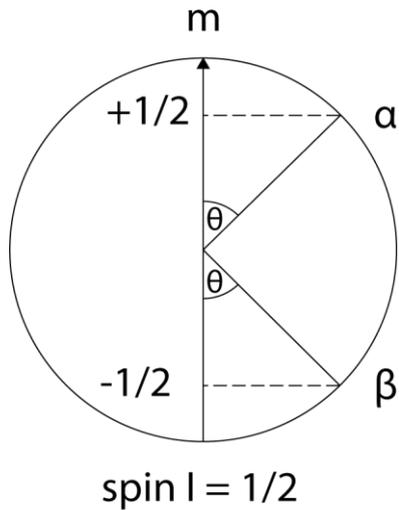
Figure 1.17. Space quantization of spin 1/2 nuclei.

The magnetic quantum number (m) describes the direction of the 2I+1 allowed orientations for the angular momentum (P) of a spin I nucleus. This schematic represents a space quantization of m for a spin 1/2 nucleus showing the two possible directions m = +1/2 and m = -1/2.
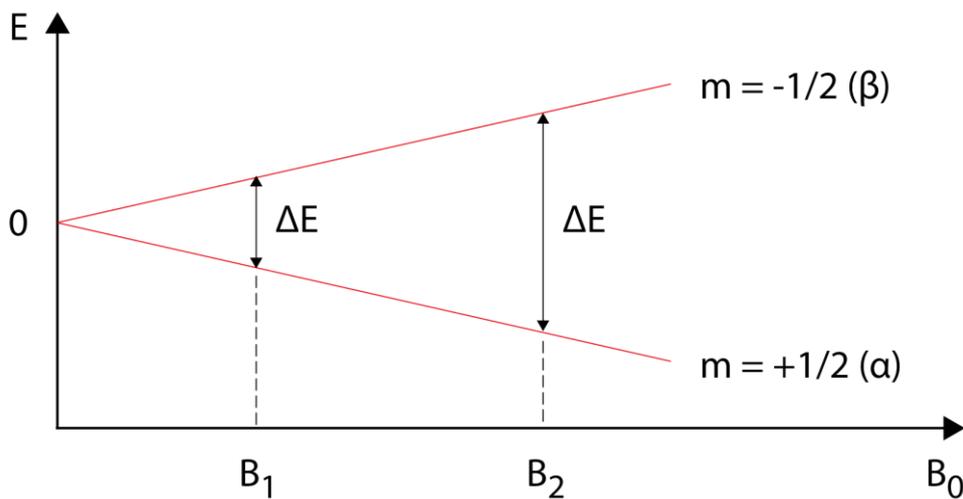
Figure 1.18. Energy state of allowed orientations for spin 1/2 nuclei and dependence on B_0.

Spin 1/2 nuclei can align in a high (α) and low (β) energy state in the presence of B_0. The sensitivity of NMR depends on the transition ΔE. Due to the relationship ΔE= $\hbar\gamma B_0$, the transition ΔE and thus NMR sensitivity can increase with increasing B_0.

The vector model of NMR

At thermal equilibrium and in the presence of B_0 the Boltzmann distribution describes a slight excess of the low energy α state which results in a net magnetization vector (M) along the z-axis (parallel to B_0) termed longitudinal magnetization. When a second magnetic field B_1 is applied perpendicular to B_0 via a transmitter coil, this forces M into the xy plane termed transverse magnetization which precesses at the Larmor v. Subsequently this induces into the receiver coil an oscillating voltage that decays over time to produce the free induction decay (FID) which is Fourier transformed into an NMR spectrum. Mean while a concomitant relaxation processes reduce transverse M and increase longitudinal M until it is parallel to the z-axis again.

Longitudinal Relaxation

There are two relaxation processes known to occur in NMR termed longitudinal relaxation (denoted T1) and transverse relaxation (denoted T2) with both contributing to the decay of the FID. Firstly longitudinal relaxation is the process that restores the equilibrium population difference between the high and low energy spin states. In terms of the vector model this corresponds to a restoration of M to the +z-axis. Longitudinal relaxation occurs because spin active nuclei exhibit vibrational and rotational motion which changes their associated magnetic field and their respective interactions with the magnetic field of the surroundings termed the lattice. These magnetic interactions are mediated directly through space defined as dipolar interactions. Magnetic fields associated with the lattice contain components that are identical in Larmor v to the spin active nuclei and can thus interact with high energy state spin active nuclei. Subsequently the energy associated with such spin active nuclei is subsequently dissipated into the lattice causing an increase in the vibrational and rotational energy of the lattice. Hence longitudinal relaxation is also termed spin-lattice relaxation.

Transverse relaxation

T2 relaxation describes the loss of transverse M predominantly due to a loss of phase coherence in the xy plane. This arises because spin active nuclei when initially in the xy plane experience slightly different local magnetic fields due to fluctuations in local magnetic fields and a non-homogenous static magnetic field. As a result these nuclei exhibit different Lamor ν causing a loss of phase coherence and therefore no resultant M in the xy plane. Accordingly the broad distribution of Larmor ν correlates to a spread of frequencies in the NMR spectrum that are defined by a line shape with line width described in terms of: $\Delta V_{1/2} = 1 / \pi T_2$. Another mechanism that contributes to transverse relaxation is the transfer of energy between interacting spin active nuclei that have identical Larmor ν but different m states. Such nuclei exchange m states and thus undergo spin flipping i.e. spin 1/2 nuclei interchange between β state and α state. Effectively no net change in the population of energy states occurs although the average life time of the spin active nuclei in the excited state is reduced which can cause line shape broadening.

Transverse Relaxation Optimised Spectroscopy and large proteins

A major limitation of NMR is protein size because large proteins give rise to a greater number of resonances that results in resonance overlap. Furthermore resonance signals broaden out due to slow tumbling of proteins in solution and fast transverse relaxation rates (T2). Tumbling is described by the rotational correlation time (τ_c) which is proportional to 1/T2. Therefore large proteins have large τ_c and short T2 and vice versa. At high magnetic field strengths the two main types of relaxation that occur are to due to chemical shift anisotropy (CSA) and dipole dipole (DD) interactions (Pervushin, Riek et al. 1997; Pervushin 2000; Riek, Pervushin et al. 2000; Fernandez and Wider 2003).

The CSA interaction arises because nuclei can have different local magnetic fields depending on their orientation with respect to the external magnetic field hence they are anisotropic. CSA mediated relaxation can be described by a CSA rate constant

which increases with the square of the magnetic field. DD interactions originate from nuclei being able to generate an instantaneous magnetic dipolar field that is proportional to µ. The effectiveness of DD mediated relaxation is dependent on µ associated with the nucleus and the inverse sixth power of the distance between the nuclei ($1/r^6$) (Pervushin 2000).

In proteins the main source of DD mediated relaxation is from ^{1}H. Significantly relaxation via DD can be reduced with the introduction of ^{2}H at non-labile positions within a large protein. The TROSY effect (Figure 1.19) can reduce the rate of transverse relaxation rate (T2) (describes the dephasing of transverse components of the precessing magnetization) by exploiting interference between CSA and DD to effectively suppress the transverse relaxation. In a TROSY experiment the slowest relaxing component is independent of T2 and is therefore selected to produce a well resolved peak (Pervushin 2000).
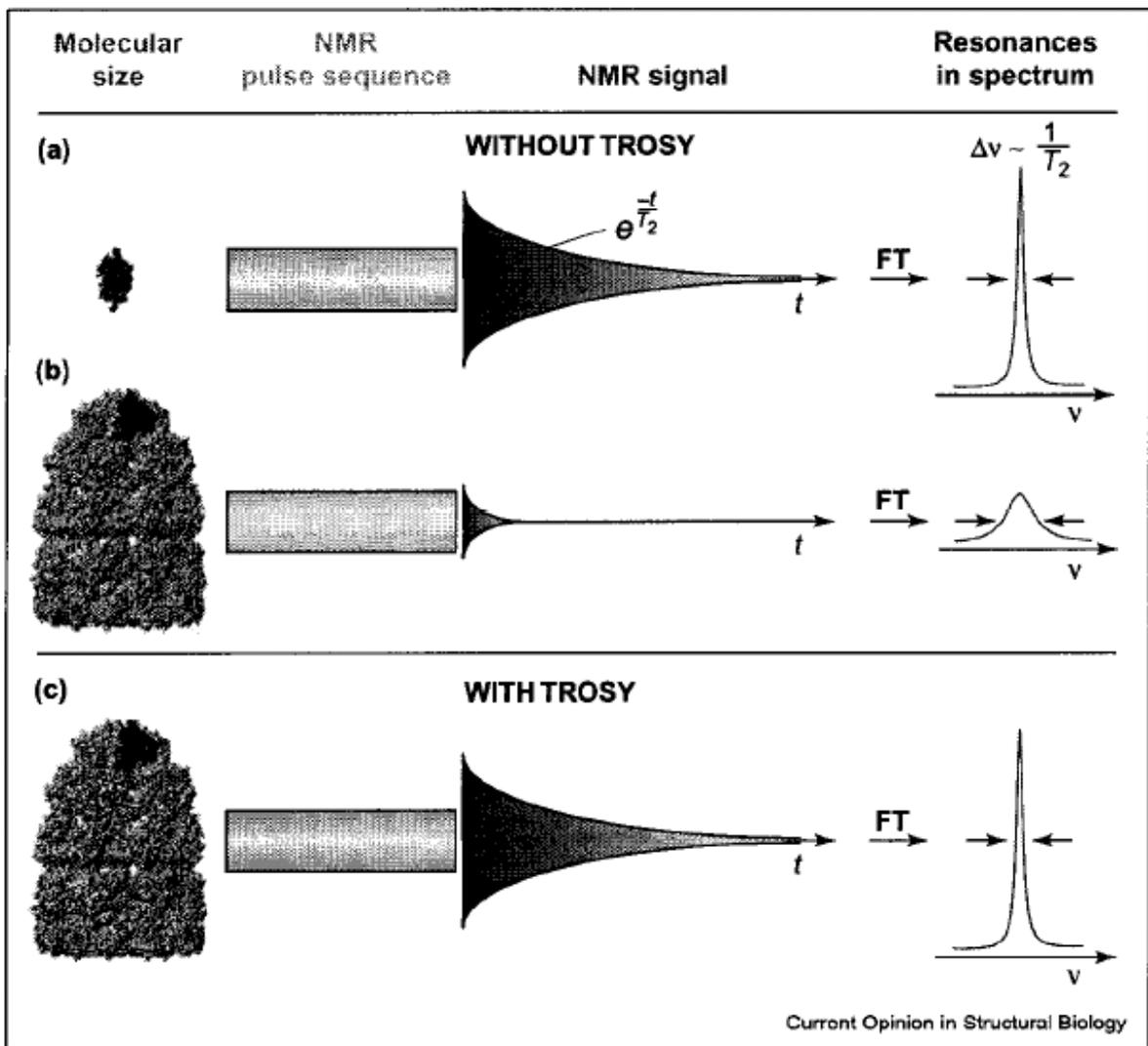
Figure 1.19. TROSY effect.

The FID decays less rapidly when TROSY is applied and corresponds to sharper resonance line shapes and an improvement in resolution.

Chemical shift

It is known that the larmor ν for a given nucleus is proportional to the applied external magnetic field B_0 but for nuclei within molecules such as proteins the situation is slightly different because these nuclei are surrounded by moving electrons. A consequence of this is that the local magnetic field around the nucleus is altered due to an opposing magnetic field produced by the surrounding electron. Basically the external magnetic field experienced by the nucleus is reduced which in turn changes its larmor ν. This effect is described as shielding and its extent is depedent on the density and distribution of the electron which surrounds the nucleus. The chemical shift (δ) therefore describes the dependence of NMR energy levels ($\Delta E = h\nu$) on the electronic environment. Identical nuclei in different chemical environments can thus possess different larmor ν. For any resonance, δ can be defined as:

$$\delta = \nu - \nu_{Ref} / \nu_{Ref}$$

Chemical shift perturbation

Chemical shift perturbations (CSPs) are a property used in NMR to indicate changes in the local environment of nuclei. CSPs associated with 15N nuclei are greater than 1H therefore weighting is required to address this imbalance and described as normalisation of the CSP. Thus standard formula used to account for this is: $\Delta\delta$ total $= ((\Delta\delta_H)^2 + (\Delta\delta_N)^2)^{0.5}$

HSQC

Heteronuclear single quantum coherence (HSQC) is a 2D NMR experiment that resolves protein ^1H peaks via the ^1H and ^{15}N dimensions. It is commonly used to assess the suitability of a protein for NMR backbone assignment. The transverse relaxation optimised spectroscopy (TROSY) HSQC is the same experiment as a standard HSQC but is designed for large molecular weight proteins that exhibit broad lineshapes and overlap. TROSY does not utilise decoupling, selecting only the

narrowest lineshape resonance and is often combined with deuterated proteins to improve signal intensity and lineshape leading to an improvement in resolution. These experiments can also be used to test for ligand or protein binding against a ^{1}H, ^{15}N labelled protein along with characterisation of binding affinities. The amide ^{1}H is sensitive to changes in its local chemical environment therefore the HSCQ is a sensitive method for the detection of ligand and protein-protein interactions.

Backbone assignment strategy

Sequential assignment of the backbone amide resonances in a protein can be achieved by acquiring a set of triple resonance experiments using a ^{15}N, and ^{13}C uniformly labelled protein. Proteins of large molecular weight (i.e. $\geq$ 25 kDa) may require deuteration or perdeuteration. All the subsequent experiments to be described provide carbon resonance information which are correlated to the backbone ^{1}HN (amide proton) and ^{15}N (amide nitrogen) resonances evident in a HSQC spectrum. Provided the primary sequence is known a backbone assignment is accomplished by establishing the connectivity between residues where possible via their ^{13}CO (carbonyl carbon), ^{13}CA (α carbon) and ^{13}CB (β carbon) resonances. Certain residue types have unique ^{13}CA and ^{13}CB resonances thus the identification of residues in the primary sequence for building a sequential assignment relies on this aspect.

HN(CA)CO and HNCO

A backbone assignment process begins with the HNCO experiment which correlates the backbone amide resonance of residue i to the ^{13}CO (carbonyl carbon) resonance of the preceding residue termed residue i-1 (Figure 1.20). Each backbone amide resonance therefore produces one inter (residue i-1) ^{13}CO resonance in the HNCO spectrum. In terms of signal to noise ratio, the HNCO is the most sensitive experiment which helps to confirm the number of real backbone amide resonances observed in the HSQC particularly in overlapped regions. However the HNCO alone cannot determine the connectivity of backbone amide resonances via ^{13}CO

resonances and therefore requires the HN(CA)CO (brackets around the nuclei denotes that the magnetization is transferred via the nuclei but its resonance is not observed). This complementary experiment correlates the backbone amide resonance of residue i to its intra (residue i) ^{13}CO and to the inter (residue i-1) ^{13}CO. Thus for each backbone amide resonance correlated to the HN(CA)CO spectrum two ^{13}CO resonances are produced. Superimposing HNCO and HN(CA)CO spectra therefore helps to confirm the intra ^{13}CO and inter ^{13}CO resonance of residue i and i-1.

HNCA and HN(CO)CA

Another pair of experiments required for sequential backbone assignment includes the HNCA and HN(CO)CA, with the former experiment correlating the backbone amide resonance of residue i to its intra (residue i) ^{13}CA and to the inter (residue i-1) ^{13}CA. Subsequently two ^{13}CA resonances are produced for each backbone amide resonance correlated to the HNCA. However the ^{13}CA resonance information provided by the HNCA needs to be complemented by the HN(CO)CA because this experiment correlates the backbone amide resonance of residue i to the inter (residue i-1) ^{13}CA resonance. In this case one inter (residue i-1) ^{13}CA resonance is produced for each backbone amide resonance correlated to the HN(CO)CA. Combining the HNCA and HN(CO)CA spectra therefore confirms the intra ^{13}CA and inter ^{13}CA resonance of residue i and i-1.

HNCACB and HN(COCA)CB

In order to increase confidence in identifying residue types and establishment of unambiguous sequential assignments it is necessary to acquire ^{13}CB resonance information. Hence the HNCACB is acquired because this experiment correlates the backbone amide resonance of residue i to the intra (residue i) and inter (residue i-1) ^{13}CA and ^{13}CB resonances (Figure 1.20). Potentially four, three or two peaks maybe produced for each backbone amide resonance correlated to the HNCACB depending on whether both residues i and -1 are Gly or if one of them is Gly because only Gly

residues lack a CB side chain. Exploiting the HNCACB requires combining it with experiments that help to distinguish between the intra (residue i) and inter (residue i-1) ^{13}CA and ^{13}CB resonances. Identifying the inter (residue i-1) ^{13}CB in a HNCACB spectrum depends on the HN(COCA)CB because this experiment correlates the backbone amide resonance of residue i to the inter (residue i-1) ^{13}CB resonance. Therefore the HN(COCA)CB produces one inter (residue i-1) ^{13}CB resonance for each backbone amide resonance correlated to the HN(COCA)CB assuming that residue i-1 is not a Gly. A combination of the HNCACB, HNCA and HN(CO)CA will also help confirm the intra (residue i) and inter (residue i-1) ^{13}CA resonances.
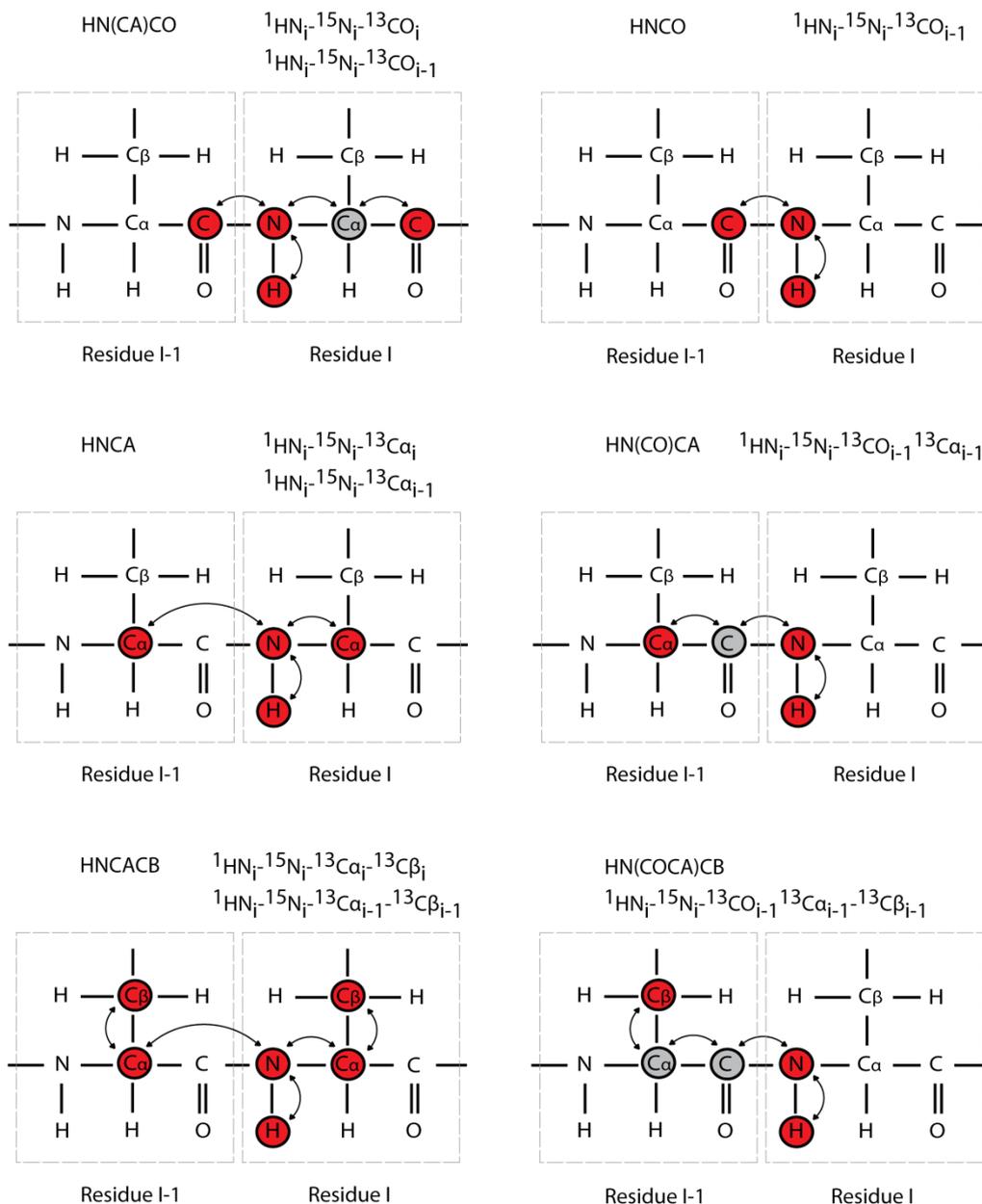
Figure 1.20. Triple resonance experiments HN(CA)CO, HNCO, HNCA, HN(CO)CA, HNCACB and HN(COCA)CB.

Magnetization transfer pathways of the experiments are shown denoted by an arrow. Nuclei highlighted in red indicate that resonances for these are observable in the experiment. Nuclei highlighted in gray participate in he magnetization transfer but are not observable in the experiment.

Chapter II Materials and Methods

All chemicals, reagents, and equipment purchased or used came from Alta Biocience, Beckman, Biorad, Calbiochem, Expedeon, Fisher scientific, GE healthcare, Invitrogen, Merck, Roche, Stratagene, and Sigma-Aldrich.

cDNA clones

All cDNA clones shown in table1 were provided by our collaborators the Structural Genomics Consortium (SGC). Both pLIC-SGC1 and pNIC-Bsa4 expression vectors encode an N-terminal His$_6$-tag and an adjacent TEV protease cleavage site for His$_6$-tag removal. The pGEX-6p-2 encodes an N-terminal GST-tag, followed by a prescission^{TM} protease (GE healthcare) cleavage site to allow removal of the GST-tag.

cDNA clone	Default Cell line	Antibiotic resistance	Expression vector	Affinity Tag
TOPK3-322	BL21 (DE3)	50 µg/ml Kanamycin	pNIC28-Bsa4	His$_6$
TOPK10-322	BL21 (DE3)	50 µg/ml Kanamycin	pNIC28-Bsa4	His$_6$
Pim1	BL21 (DE3)	100 µg/ml Ampicillin	pLIC-SGC1	His$_6$
CSNK1G3	BL21 (DE3)	50 µg/ml Kanamycin	pNIC28-Bsa4	His$_6$
CAMK1D	Rosetta (DE3)	50 µg/ml Kanamycin 34 µg/ml chloramphenicol	pNIC28-Bsa4	His$_6$
CaMK1G	BL21 (DE3) R3-p2RARE	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
VRK2A	Rosetta (DE3)	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
PAK4	BL21 (DE3)	100 µg/ml Ampicillin	pGEX-6p-2	GST
PAK5	BL21 (DE3)	50 µg/ml Kanamycin	pNIC28-Bsa4	His$_6$
PAK6	BL21 (DE3)	50 µg/ml Kanamycin	pNIC28-Bsa4	His$_6$
Calmodulin	BL21 (DE3)	50 µg/ml Kanamycin	pNIC28-Bsa4	His$_6$

Table1. SGC cDNA clone details. The pNIC28-Bsa4 and pLIC-SGC1 are customised pET vectors created by the SGC. All subsequent media used in the protocols described were supplemented with the relevant antibiotics shown in this table. See appendix for vector maps.

cDNA clone	Optimised Cell line	Antibiotic resistance	Expression vector	Affinity Tag
CaMK1D	BL21 (DE3) RIPL codon plus	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
CaMK1D1-298	BL21 (DE3) RIPL codon plus	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
CAMK1D1-280	BL21 (DE3) RIPL codon plus	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
CaMK1D_T180E	BL21 (DE3) RIPL codon plus	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
CaMK1D_T180E1-298	BL21 (DE3) RIPL codon plus	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$
CaMK1D_T180E1-280	BL21 (DE3) RIPL codon plus	50 µg/ml Kanamycin 34 µg/ml Chloramphenicol	pNIC28-Bsa4	His$_6$

Table2. Expression construct details of CaMK1D and CaMK1D mutants after cell line optimisation. Listed are details of the CaMK1D constructs that were used in this thesis. All subsequent media used in the protocols described were supplemented with the relevant antibiotics shown in this table.

Vector Maps
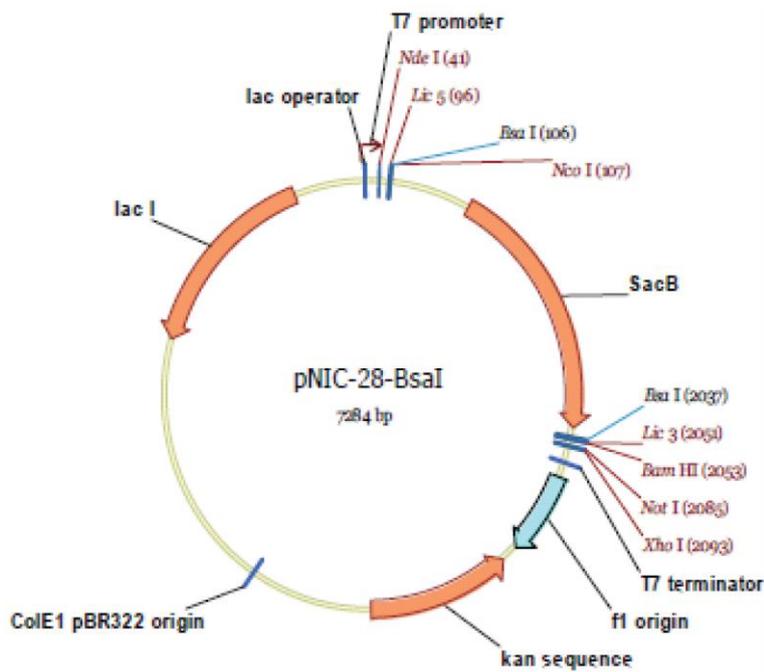
Figure.2.0. pNIC-28-Bsa4 vector.
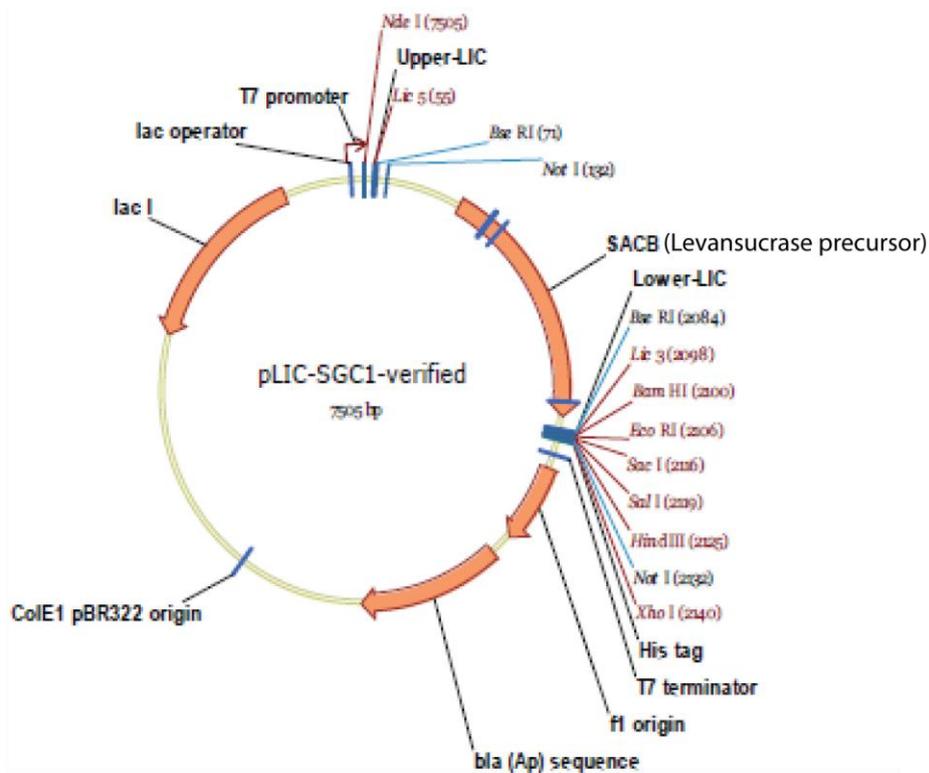
Figure.2.1. pLIC-SGC1 vector.
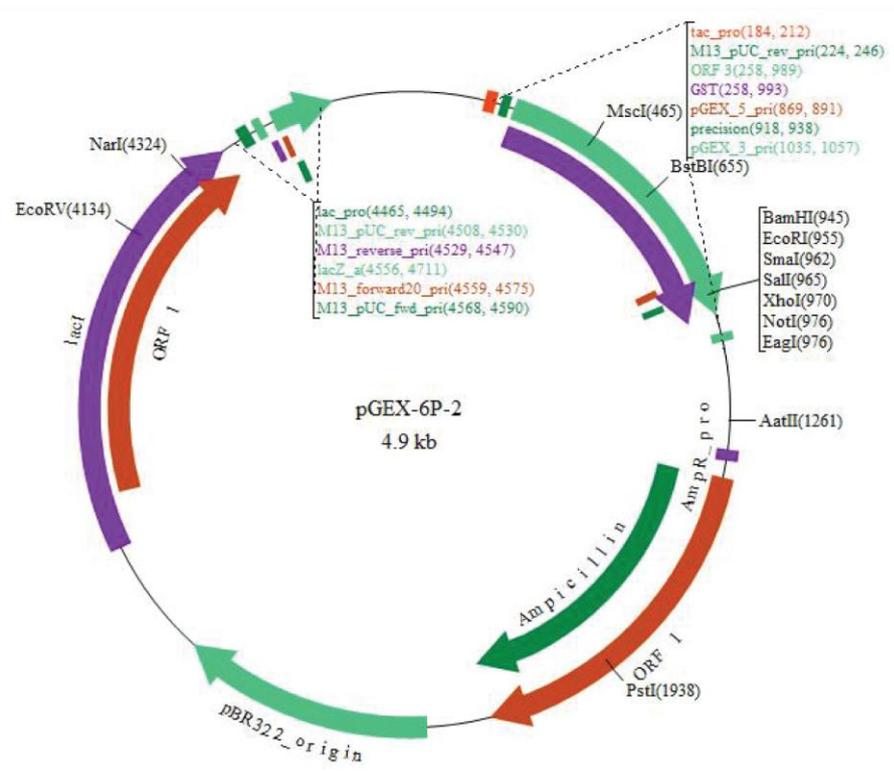
Figure 2.2. pGEX-6P-2 vector.

Transformations

Rosetta (DE3) and BL21 (DE3)

The concentration of cDNA stocks used in transformation typically ranged from 20-200 ng µL^{-1}. The cDNA plasmids were transformed into their respective default *E. coli* cell lines by adding 1 µL of cDNA plasmid to 20 µL of competent cells (Novagen) pre-thawed on ice and incubating them on ice for 5 minutes. The cells were then subjected to a 30 second heat shock at 42 °C and transferred back on ice for 2 minutes. Subsequently 80 µL SOC or LB media equilibrated at room temperature was added to the cells which were then incubated at 37 °C with shaking at 200 RPM for 1 hour. Finally 50 µL of cells were plated out on an LB agar plate supplemented with the appropriate antibiotic resistance, then inverted and incubated at 37 °C overnight.

BL21 (DE3) RIPL codon plus

cDNA plasmids of CaMK1D and its mutants plasmids (cDNA stock concentrations typically ranged from 20-200 ng µL^{-1}) were into BL21 (DE3) RIPL codon plus (Stratagene) by adding 1 µL of 10-fold dilute β-mercaptoethanol mix (stratagene) to 20 µL of competent cells on ice for 10 minutes. Subsequently 1 µL of cDNA plasmid was added to the cells and incubated on ice for 30 minutes. Cells were then heat shocked for 20 seconds at 42 °C and transferred back on ice for 2 minutes followed by the addition of 180 µL of SOC or LB pre-heated at 42 °C and incubation at 37 °C, 200 RPM, for 1 hour. Finally 50 µL of cells were plated out on an LB agar plate supplemented with the appropriate antibiotic resistance, then inverted and incubated at 37 °C overnight.

Protein expression

Expression tests in LB

Small scale expression tests were conducted by inoculating 5 mL of LB culture containing the appropriate antiobiotic(s) with a single cDNA plasmid transformed colony then incubated at 37 °C, 200 RPM, overnight. 0.5 mL of this start up culture was then used to inoculate 100 ml of LB culture the next morning. All the kinase proteins were grown at 37 °C then cooled to 18 °C when $O.D_{600nm}$ was 0.3, whilst CaM was grown at 37 °C. When cultures reached $O.D_{600nm}$ equal to 0.6-0.7, a 1 mL sample was taken as a pre-induction sample then 0.5 mM IPTG was added to the main culture and left to grow overnight. 1 mL post induction samples were taken for expression test comparison i.e. after 4, 8, 16 post induction hours. Normalised pre and post induction samples from whole cell lysates were prepared for SDS PAGE (BioRad, Tris HCl, 15% gel: 15% resolving gel, 4% stacking gel) and stained with comassie blue or instant blue[TM] (Expedeon)

Solubility tests

100 µL of post-induction sample were sonicated on ice at 20 second intervals. Sonicated samples were then centrifuged at 18000xg for 3 minutes. The supernatant was transferred to a new eppendorf. An equal volume of SDS sample buffer was added to the supernatant (soluble fraction) whilst the pellet (insoluble fraction) was made up to the same final volume in SDS sample buffer and resuspended.

CaMK1D Cell line and deuterated M9 minimal media expression tests

Expression tests of CaMK1D in BL21 (DE3), BL21 DE3 RIPL codon plus, and Rosetta (DE3), in deuterated M9 minimal media (see appendix) was conducted as follows. Single transformed colonies of BL21 (DE3), BL21 DE3 RIPL codon plus, and Rosetta (DE3) were selected for inoculation of 2 ml of LB and grown at 37 °C, 200 RPM for 6-8 hours. 0.5 ml of LB start up was then used to inoculate 10 ml of M9 minimal media containing 70% (v/v) $^{2}H_2O$ then grown at 37 °C, 200 RPM overnight.

The next morning 2 ml of the M9 minimal media start up was then used to inoculate 100 ml of M9 minimal media containing 70 % and 99 % (v/v) ^2H_2O in 250 ml flasks then grown at 37 °C, 200 RPM until the O.D$_{600nm}$ was 0.2, at which point the temperature was lowered to 18 °C. When cultures reached an O.D$_{600nm}$ of 0.4, a 1 ml sample was taken as a pre induction sample then 1mM IPTG was added to the main culture and left to grow overnight at 18 °C, 200 RPM. 1 ml post induction samples were taken the following morning.

Expression of unlabelled kinases and CaM in LB

This protocol describes the expression of protein kinases and CaM. A single colony transformed with the cDNA plasmid of interest was used to inoculate 20 ml of LB and incubated overnight at 37 °C, 200 RPM. In the following morning the 20 ml LB start up culture was used to inoculate 1 L of LB in a 5 L flask and incubated at 37°C, 180 RPM. For kinases at an O.D$_{600nm}$ of 0.4, the culture was cooled to 18 °C and induced with 1 mM IPTG at an O.D$_{600nm}$ of 0.6 and grown overnight with shaking at 180 RPM. For CaM, the temperature was maintained at 37 °C and at an O.D$_{600nm}$ of 0.6 cells were induced with 1 mM IPTG and grown overnight at 37 °C, 180 RPM.

Expression of isotope and deuterium labelled proteins

Preliminary ^1H, ^{15}N labelled kinase samples, ^2H CaM and other deuterated proteins were expressed in their default cell lines. All ^2H, ^{15}N CaMK1D constructs, and ^2H, ^{15}N, ^{13}C CaMK1D were expressed in either Rosetta or BL21 (DE3) RIPL codon plus. These proteins were expressed as described below.

2 ml of LB was inoculated with a single colony transformed with the cDNA plasmid of interest and grown at 37 °C, 200 RPM for 6-8 hours. After this, 50 ml of M9 minimal media start up was inoculated with 0.5 ml (1ml for CaM) of the LB start up and incubated at 37 °C, 200 RPM overnight. In the case of ^2H, ^{15}N CaMK1D constructs, ^2H CaM, ^2H, ^{15}N, ^{13}C CaMK1D and other deuterated proteins, 0.5 ml was used to

inoculate 50 ml M9 minimal media containing 70 % (v/v) ^2H_2O in a and grown overnight at 37 °C, 200 RPM.

Following overnight growth of non-deuterated cultures the 50 ml M9 start up was then used to inoculate 1 L of M9 minimal media in a 5 L flask in the morning and incubated at 37 °C, 180 RPM until the $O.D_{600nm}$ was 0.3, at which point the temperature was lowered to 18 °C. At an $O.D_{600nm}$ equal to 0.6, the 1 L culture was induced with 1 mM ITPG and grown overnight.

After overnight growth of deuterated cultures the 50 ml M9 start up was used to inoculate 1 L (as a final volume) of M9 minimal media containing 99 % (v/v) ^2H_2O for in a 5 L baffled flask in the morning and incubated at 37 °C, 120 RPM until the $O.D_{600nm}$ was 0.2. At this point the temperature was cooled to 18 °C and at an $O.D_{600nm}$ of 0.4, the cells induced with 1 mM IPTG and grown overnight. In the case of deuterated CaM, cells were maintained at 37 °C and induced at an $O.D_{600nm}$ of 0.4 and grown overnight. For the inoculation step a slight dilution effect of the final ^2H_2O content can result when adding the M9 start up containing 70 % (v/v) ^2H_2O. An alternative method to ensure 99 % (v/v) ^2H_2O or perdeuterated growths was to spin down the cells at 4000 RPM (Beckman centrifuge) and resuspend in 1 L (as a final volume) of M9 minimal media containing 99 % (v/v) ^2H_2O. When high levels deuterium incorporation were not desired the amount of deuterium oxide in the final 1 L culture was adjusted accordingly i.e. to 70 % (v/v) ^2H_2O. ^2H CaM was expressed in the same manner as CaMK1D in 99% deuterium oxide but without any ^{15}N, ^{13}C labelling and induction was conducted at 37°C.

Primary Ni^{2+} NTA-affinity protein purification

After overnight expression, cells were centrifuged at 7000xg (Beckman centrifuge) for 15 minutes and resuspended on ice in equilibration buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 5 mM imidazole, 0.02 % NaN$_3$) supplemented with an EDTA-free complete protease inhibitor cocktail (Roche). The cells were lysed

on ice by using either the French press or the Emulsiflex^{TM} cell disruptor with the resulting cell lysate centrifuged at 75000g for 45 minutes (Beckman centrifuge). The supernatant was filtered through a 0.45 µM filter and His$_6$-tagged protein was purified by Ni^{2+}NTA-affinity chromatography (GE Healthcare). This involved the filtered cell lysate being applied to the Ni^{2+}NTA-affinity column via a peristaltic pump, followed by a wash with 5 column volumes of wash buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole, 0.02 % NaN$_3$) to remove non-specifically bound proteins. His$_6$-tagged protein i.e. CaMK1D was eluted with elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 250 mM imidazole, 0.02 % NaN$_3$) and His$_6$-tagged CaMK1D containing fractions were pooled together as determined from SDS PAGE.

His$_6$-tag cleavage and dialysis

Pooled His$_6$-tagged CaMK1D (~20 mg/ml, ~ 520 µM) was transferred to a dialysis tube to which 0.5-1 mg/ml TEV protease (expressed with a His$_6$-tag and purified inhouse from a cDNA plasmid provided by the SGC) was added to remove the His$_6$-tag, and dialyse out the imidazole at 4 °C overnight in dialysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM TCEP, 0.02 % NaN$_3$).

Secondary Ni^{2+} NTA-affinity purification

After overnight dialysis, CaMK1D was again purified on the Ni^{2+}NTA-affinity column by loading the sample on with a peristaltic pump. The flow through was collected then 5 column volumes of equilibration buffer and wash buffer was passed through respectively and each collected separately. Finally 5 column volumes of elution buffer was passed through and collected. Fractions containing CaMK1D were determined by SDS PAGE and pooled together. TEV protease co-elutes in the CaMK1D fractions but is removed in the following size exclusion step.

Buffer compatibility for CaMK1D•Ca^{2+}/CaM- HEPES buffer

For the binding studies and activity assays involving CaMK1D, CaMK1D_T180E, CaMK1D_W306S, CaMK1D_M319A and the C-terminal truncation constructs the running buffer for the superdex S75 size exclusion was 50 mM HEPES pH7.5, 150 NaCl, 0.5 mM TCEP, 0.02 % NaN_3. HEPES was used instead of Na phosphate because CaMK1D, CaMK1D_T180E, CaMK1D_W306S, and CaMK1D_M319A all require Ca^{2+}/CaM for activation but Ca^{2+} when combined with Na phosphate (default kinase buffer) results in precipitate of Ca^{2+} phosphate. Although the C-terminal truncation constructs have Ca^{2+}/CaM independent activity, the same HEPES buffer system was used for these to maintain consistency with the other CaMK1D constructs for binding and activity studies.

Size exclusion chromatography

Post his_6-tag cleavage and second Ni2+ NTa affinity purification proteins were pooled and concentrated to 1 mL in a 10 K MWCO concentrator (Vivaspin or Amicon Ultra-15 centrifugal device (Millipore)). This sample was then centrifuged at 18000xg for 10 minutes to remove aggregate and precipitate. The resulting supernatant was then extracted into a syringe and loaded on to the superdex S75 (GE Healthcare) size exclusion column in 50 mM Na phosphate or HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02% NaN_3. The size exclusion elution profile was monitored at 280 nM and the elution profile indicated that the protein was monomeric. CaM was purified by size exclusion using 300 mM NaCl. The absorbance of CaM is also a lot lower because CaM does not contain any Trp residues. Fractions containing CaMK1D according to SDS PAGE were pooled and the sample was concentrated and exchanged into 50 mM Trizma pH 9.5, 150 mM NaCl, 0.5 mM TCEP and left at room temperature (25 °C) for 24 hours in order to facilitate the exchange of amide ^{2}H with ^{1}H. After complete exchange the sample was exchanged back into buffer containing 50 mM Na phosphate or HEPES pH 7.0, 150 mM, NaCl, 0.5 mM TCEP, 0.02 % NaN_3 using a 10 K MWCO concentrator (Vivaspin or Amicon Ultra-15 centrifugal device (Millipore). The ^{2}H, ^{15}N, ^{13}C labelled CaMK1D sample used for triple resonance backbone experiments was buffer exchanged into the Na phosphate buffer previously described but with 75 mM NaCl. All NMR samples were prepared in 600 µl final volume containing 10 % ^2H_2O.

Anion exchange

CaM was purified in the same manner as described for CaMK1D but with an additional anion exchange (Resource Q column) purification step after size exclusion. Pooled CaM fractions from the Superdex-75 size exclusion purification were concentrated to 1 ml in a 5 K MWCO concentrator (Vivaspin or Amicon Ultra-15 centrifugal device (Millipore) then diluted 10-fold with resource buffer A (50 mM Hepes pH 8, 0.5mM TCEP, 0.02 % NaN_3) and concentrated to 5 ml. Alternatively a superloop can also be used with CaM diluted 10-fold in resource buffer A. CaM was then loaded onto the ResourceQ via a 5 ml loop and eluted over a linear gradient with resource buffer B (50 mM Hepes pH 8.8, 1 M NaCl, 0.5mM TCEP, 0.02 % NaN_3) from 0 % buffer B to 100 % buffer B concentration. CaM containing fractions were pooled as determined from SDS PAGE. For deuterated CaM the amide ^{2}H exchange with ^{1}H procedure was not necessary. TOPK was also purified by anion exchange as a final step with the buffer pH adjusted accordingly to the predicted pI of TOPK

Beer Lambert law

Except for CaM, all the concentrations of all proteins were estimated via UV absorbance at 280nM and calculated using the beer-lambert law.

$A = c\varepsilon l$

Where A = absorbance, c = concentration, ε = extinction coefficient, l = path length

The concentration of CaM was determined via the Bradford assay because CaM does not contain any Trp residues necessary for estimation of its concentration with <10% error by UV absorbance at 280nM.

Bradford assay

Purified CaM was diluted 20-100 fold (final concentration of 10-100 μM) with 50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % NaN$_3$ (CaM storage buffer), in a final volume of 20 μL. A blank sample was prepared using 980 μl of Bradford reagent (Biorad) and 20 μl of CaM storage buffer (50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % NaN$_3$) in a cuvette. The diluted CaM sample was then added to 980 μL of Bradford reagent. After 5 minutes a UV absorbance reading at 595 nM was measured and the concentration of CaM was calculated.

Thermofluor^{TM} (Thermal shift assay)

Stock protein was diluted to 20-100 μM with 50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP and mixed with SYPRO orange dye (Invitrogen) to a final concentration of 10x. 2 uL Samples were then dispensed to a 96-well plate at a final protein concentration of 2-10 μM including the relevant buffer or soluble ligand of interest to a final volume of 20 μl. Thermal shift assay (Thermofluor^{TM}) experiments was conducted over 20-90 C using an Mx3005P RT-PCR machine (stratagene).

Site directed mutagenesis

In order to create an activation loop phosphorylation mimic in CaMK1D, a T180E mutation was engineered into the native phosphorylation site using the Quikchange lighting mutagenesis kit (stratagene). This mutant was dubbed CaMK1D_T180E. In order to generate C-terminal truncated CaMK1D: CaMK1D1-298 and CaMK1D1-280, stop codons were introduced after residue Arg 298 and Ala 280. Site directed mutagenesis was conducted using the Quikchange lightning kit (stratagene) to generate the mutated cDNA plasmids. Mutagensis reactions were confirmed by DNA sequencing. All primers were ordered from Invitrogen or Alta biosciences.

Oligonucleotide	Primer length	GC content	Tm	Mismatch	Mutation
Forward primer 5'-GGC AAA GGA GAT GTG ATG TCC **GAG** GCC TGT GGA ACT CCA GGC TAT GTC-3'	48 bp	56%	84.5°C	6.25%	Thr 180 > Glu
Reverse primer 5'-GAC ATA GCC TGG AGT TCC ACA GGC **CTC** GGA CAT CAC ATC TCC TTT GCC-3'	48 bp	56%	84.5°C	6.25%	Thr 180 > Glu
Foward primer 5'-CAC GAG TCC GTC AGC GCC CAG ATC CGG **TGA** AAC TTT GCC AAG AGC-3'	45 bp	56%	84.4°C	6.7%	Lys 299 > stop codon
Reverse primer 5'-GCT CTT GGC AAA GTT **TCA** CCG GAT CTG GGC GCT GAC GGA CTC GTG-3'	45 bp	56%	84.4°C	6.7%	Lys 299 > stop codon
Foward primer 5'-GCA GCT CGG CAC CCA TGG ATC GCT **TGA** GAC ACA GCC CTC AAC-3'	42 bp	56%	83.7°C	7.14%	Gly 281 > stop codon
Reverse primer 5'-GTT GAG GGC TGT GTC **TCA** AGC GAT CCA TGG GTG CCG AGC TGC-3'	42 bp	56%	83.7°C	7.14%	Gly 281 > stop codon
Forward pimer 5'-AGA TCC GGA AAA ACT TTG CCA AGA GCA **AAT** CAA GAC AAG CAT TTA ATG CC-3'	50 bp	40%	78.4°C	4%	Trp 306 > Ser
Reverse primer 5'-GG CAT TAA ATG CTT G TC TTG **ATT** TGC TCT TGG CAA AGT TTT TCC GGA TCT-3'	50 bp	40%	78.4°C	4%	Trp 306 > Ser
Reverse primer 5'-GCC ACG GCC GTC GTG AGA CAT **GCT** AGA AAA CTA CAC CTC G-3'	40 bp	58%	78.7°C	7.5%	Met 319 > Ala
Reverse primer 5'-C GAG GTG TAG TTT TCT **AGC** ATG TCT CAC GAC GGC CGT GGC-3'	40 bp	58%	78.7°C	7.5%	Met 319 > Ala

Table 3. Primer designed for site directed mutagenesis. The primer sequence highlight in bold and underlined indicates the engineered mutation.

Isothermal titration calorimetry

CaMK1D and CaM samples for ITC were prepared by dialysing into 1 L of ITC buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 TCEP, 10 mM $CaCl_2$) overnight at 4 °C using D-tube dialyzers (Novagen) with 12-14 and 6-8 kDa MWCO for CaMK1D and CaM respectively. Following dialysis, CaMK1D was diluted to a final concentration of 10 µM in a final volume of 2.2 mL (required volume for the cell sample to fill the 1.4 mL cell), whereas CaM was diluted to 150 µM in a final volume of 0.5 mL (required volume for syringe sample). Both samples were diluted using the same ITC buffer that they were dialysed into. These concentrations were used to produce a c value of 333.33, which corresponds to an appropriate binding isotherm for a tight binding interaction and enable determination of the Kd. ITC experiments were conducted using a VP-ITC MicroCalorimeter instrument at 25 °C. Each experiment consisted of an initial 2 L injection followed by 29 injections of 10 L injected during 20 seconds and a spacing of 280 seconds between injections. Data were processed and analysed using the program Origin.

Small angle x-ray scattering

Protein samples for SAXS were prepared by dialysing CaMK1D and CaM into SAXS buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 10 mM $CaCl_2$) overnight at 4 °C using D-tube dialyzers (Novagen) with 12-14 and 6-8 kDa MWCO for CaMK1D and CaM respectively. Following dialysis, samples were prepared to their appropriate final concentration and volume using the same SAXS buffer used in dialysis. SAXS data were collected at the X33 beam line using a MAR345 image plate detector (MarResearch, Norderstedt, Germany) located at the EMBL Hamburg outstation on a bending magnet (sector D) on the storage ring DORIS III of the Deustches Elektronen Synchotron (DESY). A single photon counting pixel detector was used as a detector. A sample detector distance of 2675 mm was used, covering the range of 0.06-6 nm (where $q = 4\pi\sin\theta / \lambda$, θ is the scattering angle and λ is the wavelength of the x-ray). The x-ray λ used was 0.15 nm. Scattering curves were collected at room temperature using 22, 11, 5.5 mg ml^{-1} for CaMK1D and 10, 5, 2.5 mg ml^{-1} for CaM.

SAXS data anaylsis was performed using the PRIMUS software (P. V. Konarev 2003). Values for the radius of gyration (R_g) and maximum particle size were deduced using GNOM (Svergun 1992). Ten low resolution *ab initio* bead models were generated with DAMMIF (Svergun 2009). Spatial discrepancy between the models was minimized with SUPCOMB (Svergun 2001), then averaged using DAMAVER (Svergun 2003) and filtered by DAMFILT to create a final bead model. All programs mentioned were available in the ATSAS suite. Manual fitting of the bead models and crystal structures and models were performed using Pymol (Schrodinger 2010).

Kinase radioactivity assay

E. coli expressed TOPK/PBK and Pim1 were assessed for activity using a radioactivity assay based on a protocol described by Park and co-workers (Park, Lin et al. 2006). The assay buffer was 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM MgCl2, 10 mM NaF, 1 mM Na_3VO_4, 1 mM EDTA, 1 mM DTT, and 50 µM ATP which was supplemented with 5 µCi of [^{32}P-γ]ATP (GE healthcare). A commercial source of active human full length TOPK/PBK (Invitrogen) expressed from the Sf9 insect cell line was used as a positive control. Histone H3 was used as a generic substrate. Activity was confirmed via transfer of the assay reaction to SDS PAGE and then analysed by autoradiography.

NADH coupled activity assay

Samples for assaying kinase activity with the NADH coupled assay were prepared in a final 600 µL volume of assay buffer (50 mM Hepes pH7.5, 150 mM NaCl, 0.5 mM TCEP, 10 mM $CaCl_2$, 20 mM $MgCl_2$), 400 µM NADH, 2 mM PEP, 2 mM ATP, 2U mL^{-1} LDH, 1.5U mL^{-1} PK, 50-100 µM (Sigma Aldrich) substrate peptide (camtide2, Altabioscience) and 500 nM Calmodulin. 50 nM of CaMK1D were used. Assay reactions were conducted at room temperature and monitored by measuring the change in NADH (ε = 6220 M-1) absorbance at 340 nM over five minutes. Specific activity was determined from initial reaction rates.

Nuclear Magnetic Resonance Experiments

Triple resonance

All triple resonance experiments were performed at 25 °C (298 K) on the Varian Inova 800 and 900 MHz NMR spectrometers equipped with triple resonance cryogenic probes with Z-axis pulse field gradients. Backbone assignments were accomplised from TROSY versions of ^{1}H,^{15}N-HSQC, HNCO, HN(CA)CO, HN(CO)CA, HNCA, HNCACB and HN(COCA)CB (Gardner and Kay 1998). Spectra were processed with NMRPipe (Delaglio, Grzesiek et al. 1995) and analysed using CCPNMR analysis (Vranken, Boucher et al. 2005). ^{2}H, ^{15}N, ^{13}C, labelled CaMK1D was buffer exchanged into in 50 mM Na phosphate pH 7.0, 75 mM NaCl, 0.5 mM TCEP, 0.02 % NaN$_3$ and 10 % ^{2}H$_2$O.

1D ^{1}H and ePHOGSY

One dimensional ^{1}H spectra were acquired on a Varian 800 MHz at 25 °C (298K) using the standard water sequence from the Varian biopack. Parameters such as the 90° pulse and T.O.F (transmitter offset) value were calibrated for each ^{1}H NMR experiment. The ePHOGSY experiments were conducted by Dr Christian Ludwid and Dr Sundaresan Rajesh using a final protein concentration of 10 µM TOPK in buffer containing 25 mM HEPES pH 7.5, 150 NaCl, 10 mM MgCl$_2$. A final AMPPNP concentration of 1 mM was titrated into 10 µM TOPK to test for binding. Spectra were processed and analysed using NMRLab (Gunther, Ludwig et al. 2000).

2D ^{1}H, ^{15}N HSQC and ^{1}H, ^{15}N TROSY-HSQC

Two dimensional ^{1}H, ^{15}N HSQC and ^{1}H, ^{15}N TROSY-HSQC experiments were acquired on a Varian 800 MHz at 25 °C using the gNhsqc sequence from the Varian biopack. The 90° pulse, T.O.F value, and PwN were calibrated for each ^{1}H, ^{15}N HSQC and TROSY-HSQC experiment. All CaMK1D titrations conducted by ^{1}H, ^{15}N TROSY-HSQC experiments were performed in 50 mM HEPES pH 7.5, 150 mM

NaCl, 0.5 mM TCEP, 0.02 % NaN$_3$ and 10 % 2H2O. Spectra were acquired with 16 scans and 128 increments.

Crystallography and NMR system (CNS)

The crystal structure of CaMK1D (2JC6) was used as the starting point for structural refinement. For the initial structure residues defined in the crystal structure were kept rigid (residues 11-58, 64-164, and 185-313). Loops which were not defined by the crystal structure were created and given random coil orientation. Restraints derived from TALOS (Shen, Delaglio et al. 2009), were in-putted as ψ and φ dihedral angle constraints. A modified simulated annealing protocol from CNS (Brunger, Adams et al. 1998; Brunger 2007) was used to refine the structure generating ten structures. This consisted of two slow cooling steps, the first consisting of 1000 steps cooling from 5000 K and the second consisting of 3000 steps cooling from 2000 K, followed by a final energy minimisation step. The simulated annealing protocol was repeated four times using the structure from the previous run as the starting structure for the next.

Chapter III Results

Evaluation of TOPK for NMR studies

TOPK purified as a monomer by size exclusion

The TOPK SGC cDNA clone that was used to evaluate its potential for NMR studies was 99 % identical to wild type TOPK dubbed TOPK full length. The only difference between the clones is the deletion of residue E2 (Figure 3.0). This construct was designed by the SGC with the E2 deletion to aid crystallization by either alleviating electrostatic repulsions that could hinder ordered protein-protein interactions during crystallization or to reduce the entropic cost of forming ordered intermolecular contacts. In terms of structure and function, this deletion is unlikely to have any impact as it is located at the N-terminus and not within the KD. ^{1}H, ^{15}N and ^{2}H, ^{15}N labelled TOPK were expressed and purified to 95 % purity based on SDS PAGE. Purification step involving SEC revealed two peaks. SDS PAGE confirmed that the larger second elution peak corresponded to TOPK and was estimated to be monomeric by SEC based on comparison to known protein standards (Figure 3.2). followed by a sharp peak in anion exchange (Figure 3.3). Exchange of amide ^{2}H for ^{1}H was facilitated by buffer exchanging into TRIZMA buffer at pH 9.0.
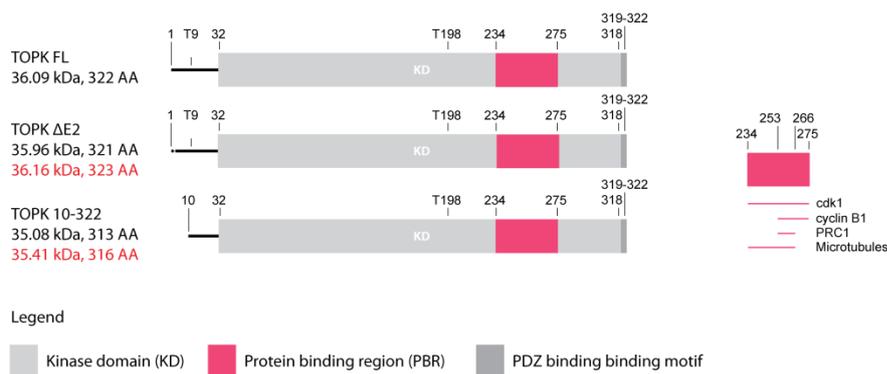
Figure 3.0. TOPK domain architechture.

Full length TOPK consists of an unstructured N-terminus featuring a regulatory phosphorylation site, T9, KD that includes C-terminal T/S-X-V PDZ interaction motif, cdk1 interaction, cyclin B1, PRC1, and microtubule interaction sites. An unstructured N-terminus Two TOPK cDNA clones were engineered by the SGC dubbed TOPK and TOPK10-322 length is shown for comparison with TOPK and $TOPK_{10-322}$. Both TOPK and $TOPK_{10-322}$ constructs encodes the ePK catalytic domain and the C-terminal T/S-X-V PDZ interacting motif represented by the white and yellow boxes respectively. Our TOPK construct is almost identical to TOPK full length except for the deletion of E2. However it does retain T9, which is a phosphorylation site reported to modulate the activity of TOPK. In $TOPK_{10-322}$, T9 is absent due to N-terminal deletion of the first 9 residues. Phosphorylation of Thr 198 is also required for activity. Within the ePK catalytic domain the gray box highlights the substrate binding region with known substrate targets mapped below. The total number of residues shown in brackets includes residues resulting from the TEV protease cleavage site i.e. residues Gly and His shown figure. The gray boxes indicate cdk1, cyclin B1, PRC1,and microtubule interaction sites.

```
            10         20         30         40         50
M-GISNFKTP SKLSEKKKSV LCSTPTINIP ASPFMQKLGF GTGVNVYLMK

            60         70         80         90        100
RSPRGLSHSP WAVKKINPIC NDHYRSVYQK RLMDEAKILK SLHHPNIVGY

           110        120        130        140        150
RAFTEANDGS LCLAMEYGGE KSLNDLIEER YKASQDPFPA AIILKVALNM

           160        170        180        190        200
ARGLKYLHQE KKLLHGDIKS SNVVIKGDFE TIKICDVGVS LPLDENMTVT

           210        220        230        240        250
DPEACYIGTE PWKPKEAVEE NGVITDKADI FAFGLTLWEM MTLSIPHINL

           260        270        280        290        300
SNDDDDEDKT FDESDFDDEA YYAALGTRPP INMEELDESY QKVIELFSVC

           310        320
TNEDPKDRPS AAHIVEALET DV
```

Predicted pI = 5.0

Figure 3.1. TOPK primary sequence.

Active TOPK requires phosphorylation of residues T9 and T198 which are highlighted in red. An unusual aspartic acid rich region located within the substrate binding region is underlined. Residue E2 was deleted by the SGC to aid crystallisation.
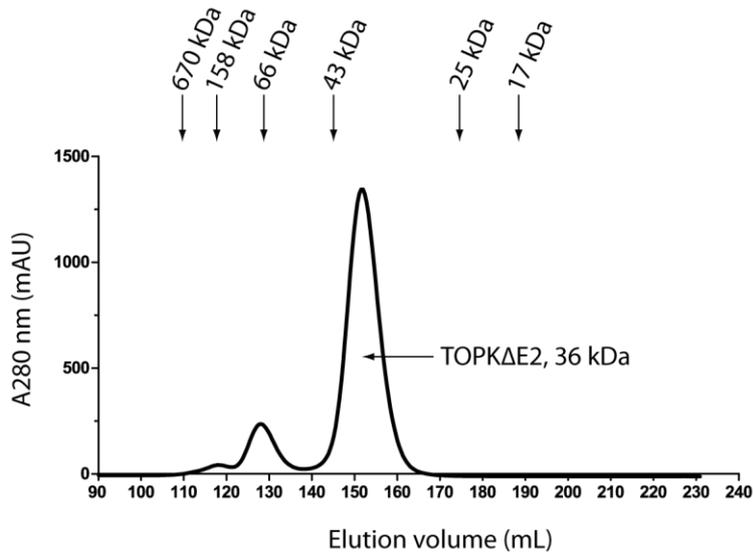
Figure 3.2. TOPK SEC elution profile.

On a superdex 75 26/60 column, TOPK elutes as the second large elution peak between 140-165 mL as confirmed by SDS PAGE. By comparison with protein standards, the peak suggests a monomeric species corresponding to the 36 kDa TOPK
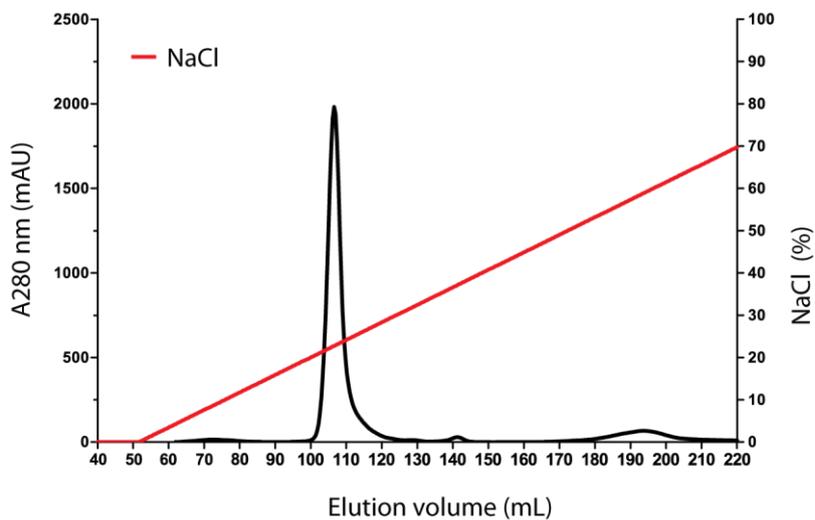
Figure 3.3. TOPK anion exchange elution profile.

On a resource Q column, TOPK was eluted over a 0-100% NaCl gradient. A single elution peak at 25% (250 mM) NaCl corresponds to TOPK.

^{1}H, ^{15}N HSQC of TOPK shows only 70 % of the expected backbone amide resonance peaks

A ^{1}H, ^{15}N HSQC spectrum of ^{1}H, ^{15}N TOPK was acquired to assess its backbone structure and backbone amide resonance peak count (Figure 3.4). The spectrum reveals that *E. coli* expressed TOPK is folded due to the observation of downfield amide resonance peakdispersion within the amide region (8.5-11.5 ppm ^{1}H dimension). Furthermore amide resonance peaks with strong intensities are evident suggesting a stable backbone, although a small subset of resonance peaks with weak intensity are also present, suggesting that the backbone also has some intrinsic dynamic behaviour. The backbone amide resonance peak count was approximately 70 % of the expected total (215/303). At least 90 % is expected for a straight forward amide directed backbone assignment. Peak overlap also was evident in the centre of the spectrum.

^{2}H labelling did not increase the amide resonance peak count

Due to the lower than expected peak count and overlap it was necessary to improve the spectrum before pursuing a feasible backbone assignment. ttempts to improve the ^{1}H, ^{15}N HSQC spectrum were first made by acquiring a ^{1}H, ^{15}H TROSY-HSQC (Figure 3.5) on a ^{2}H, ^{15}N TOPK sample (expressed using M9 minimal media containing 99% ^{2}H$_2$O) which improved the resolution and intensity of peaks but did not increase the amide peak count. The lack of additional amide peaks with ^{2}H labelled TOPK suggests that the missing peaks (broadened beyond detection) was not due to increased T2 relaxation associated with large molecular weight proteins. Instead we suspect that broadening was due to intermediate exchange phenomena. A comparison of the TOPK ^{1}H, ^{15}N HSQC and ^{1}H, ^{15}N TROSY HSQC spectra by superimposition (Figure 3.6) indicates that the peak counts are similar but highlights the improved line shape produced by the TROSY spectrum.
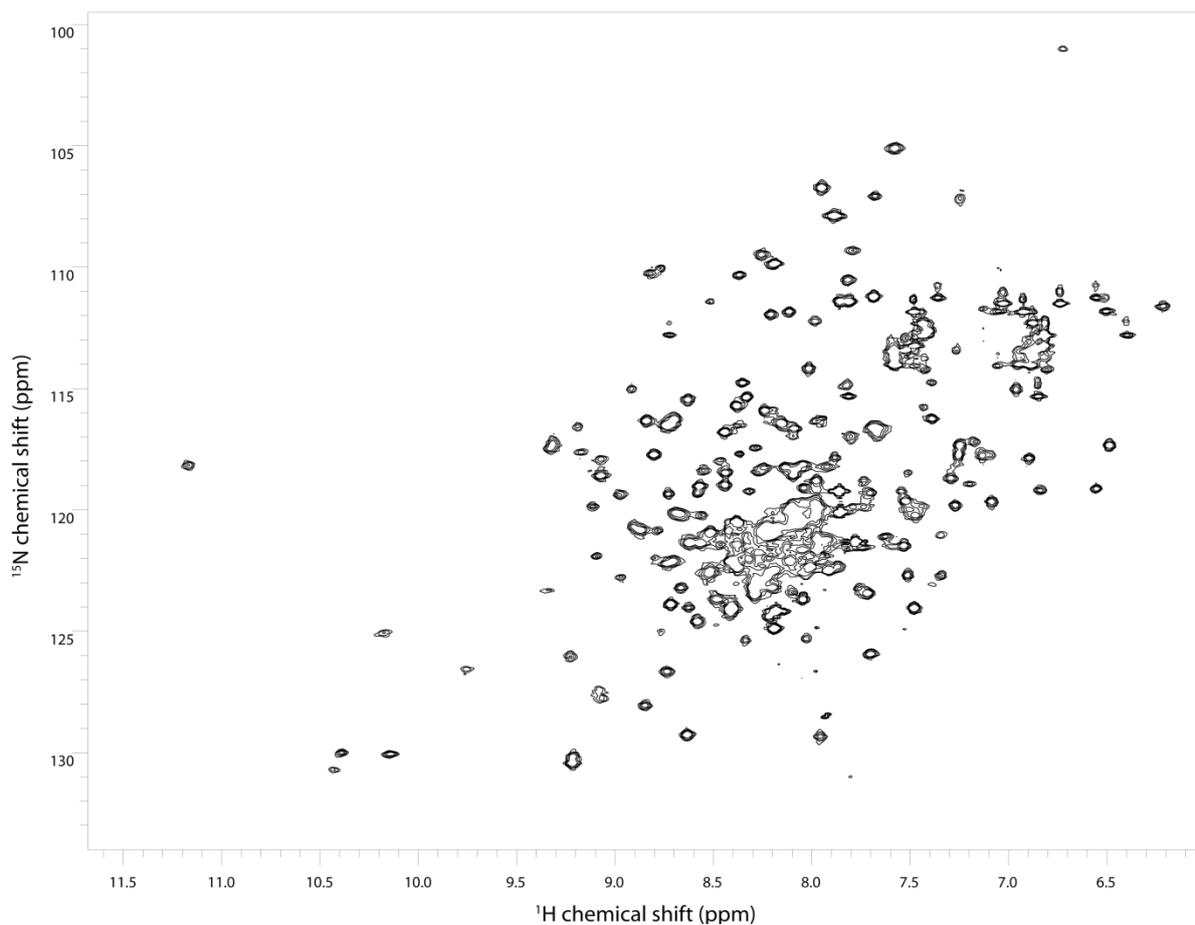
Figure 3.4. TOPK ^{1}H, ^{15}N HSQC.

Backbone amide resonance peak dispersion in the ^{1}H, ^{15}N HSQC spectrum illustrates that expression of ^{15}N TOPK in *E. coli* is folded. Approximately 215 out of 303 expected amide peaks were observable. Considerable amide resonance overlap is apparent in the centre of the spectrum. Three Trp side chain amide 1Hε resonances appear to be present (10.0-10.5 ppm 1H dimension, 129-131 ppm 15N dimension). The spectrum was acquired with 200 µM 1H 15N TOPK in 50 mM HEPES pH 7.5, 100 mM NaCl, 2.0 mM TCEP, 100 µM AEBSF, 10 % 2H2O, on a Varian 800 MHZ at 30 °C.
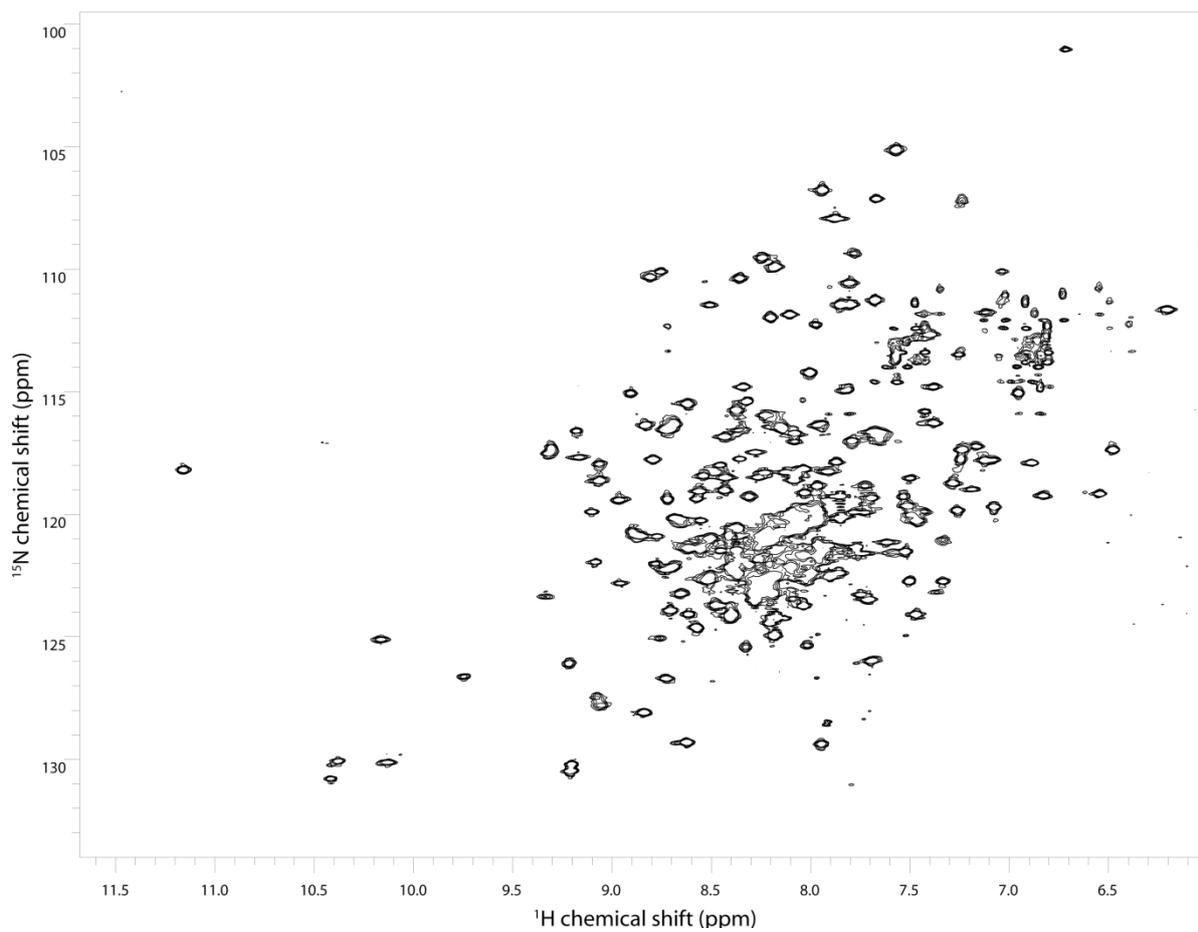
Figure 3.5. ^{2}H, ^{15}N TOPK TROSY-HSQC.

Amide resonance dispersion in the spectrum reveals that ^{2}H, ^{15}N TOPK is folded. Due to the TROSY effect the line shapes appear narrower when compared to the TOPK HSQC (Figure 3.4). However, ^{2}H labelling does not increase the peak count suggesting that the missing peaks are not due to increased T2 relaxation associated with the large molecular weight of TOPK but because of intrinsic intermediate exchange broadening. The spectrum was acquired with 200 µM ^{2}H, ^{15}N TOPK in 50 mM HEPES pH 7.5, 100 mM NaCl, 2.0 mM TCEP, 100 µM AEBSF, 10 % ^{2}H$_2$O on a Varian 800 MHZ at 30 °C.
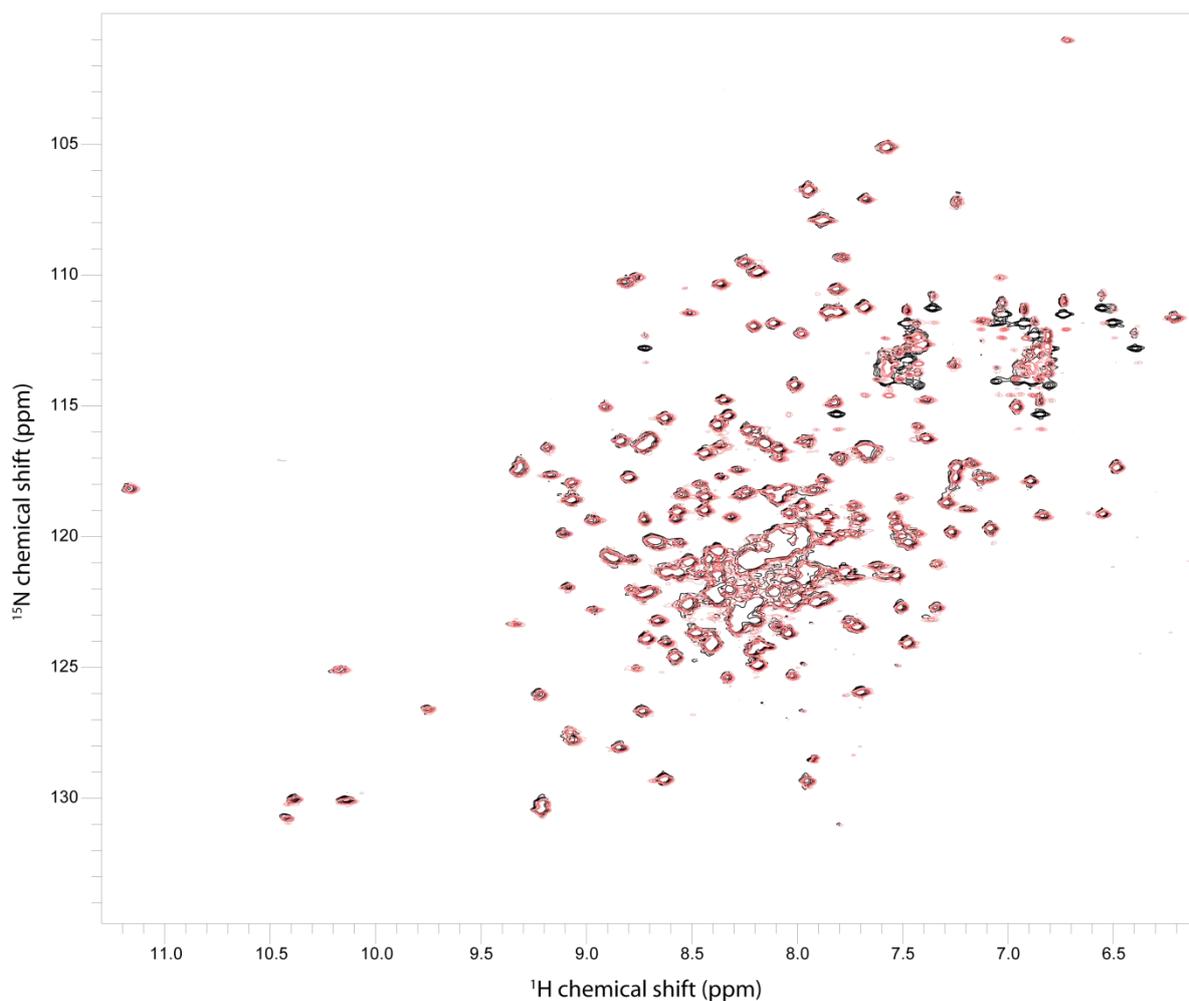
Figure 3.6. TOPK ^{1}H, ^{15}N HSQC and TROSY-HSQC.

Superimposition of the ^{1}H, ^{15}N HSQC (black) of ^{1}H, ^{15}N TOPK and ^{1}H, ^{15}N TROSY-HSQC (red) of ^{2}H ^{15}N TOPK indicates that they are nearly identical. Differences are evident in the Asn and Gln side chain amide region (6.2-8.0ppm ^{1}H dimension, 110-115 ppm ^{15}N dimension) in which the intensity of these amide resonances appear weaker in the ^{1}H, ^{15}N TROSY-HSQC compared to the ^{1}H, ^{15}N HSQC. Moreover the TROSY-HSQC spectrum of ^{2}H, ^{15}N TOPK exhibits narrower lineshapes and stronger signal intensity resulting in greater resolution compared to the HSQC spectrum. Both spectra were acquired as previously described for the TOPK ^{1}H, ^{15}N HSQC and ^{1}H, ^{15}N TROSY-HSQC using 200 µM TOPK in 50 mM HEPES pH 7.5, 100 mM NaCl, 2.0 mM TCEP, 100 µM AEBSF, 10 % ^{2}H$_2$O at 30 °C.

Thermofluor^{TM} suggests that TOPK exhibits low-moderate thermal stability in buffers that range from pH 6-8

From our initial ^{1}H, ^{15}N HSQC spectrum of TOPK it became apparent that the amide peak count was insufficient for an informative backbone assignment. In order to improve on the TOPK amide peak count we screened different buffer solution conditions that improved TOPK stability using thermofluor^{TM}, a thermal denaturing assay that can estimate the melting temperature (Tm) of a protein and therefore indicate protein stability. Sigmoidal melting curves are produced by thermofluor^{TM} for a protein that undergoes a transition from a folded to unfolded state. The Tm can be estimated from the midpoint of the sigmoidal thermal melting curve. Our buffer screen contained buffer solution conditions that varied in buffer type, pH, and stabilizers.

Improving the stability of TOPK and other protein kinases was of interest because we wanted to validate whether an increase in stability could also conceivably translate into significantly improved NMR spectra. Since most protein kinases are highly dynamic entities we believe that the stability and dynamics of a kinase are inherently related thus a kinase with high stability should possess limited dynamics and vice versa. By increasing the stability of a kinase we aim to reduce its mobility and alter the dynamics of its backbone amide resonances. Consequently residues that are stabilized become more rigid which translates to backbone amide resonances with stronger signal intensity. Amide resonances that have broadened out due to dynamics on an intermediate exchange timescale should therefore produce an observable resonance signal assuming their dynamic properties can be changed by stabilization.

Assessing the Tm of TOPK in different buffer conditions was therefore necessary to find an optimal condition that increased protein stability whilst improving its NMR spectra. Figure 3.7 presents the melting curves for TOPK in the different buffers and pH that range from 6.0-8.0. These results indicate that TOPK has Tm values that range from 42.0-51.0 °C, although the higher Tm values do not appear to correlate

with a higher pH. For example our default buffer of 50 mM HEPES pH 7.5, produced a Tm of 47.5 °C, whereas, 50 mM Na cocadylate pH 6.5, mM Na phosphate pH 7.0 and 50 mM Tris pH 8.0 produced a Tm of 48.2, 51.0 and 50.3 °C respectively. Based on our experience with thermofluor^{TM} and other kinases, we define a ΔTm $\geq$ 2.0 °C as being a significant change in stability. Furthermore we consider the range of TOPK Tm values to fall within a range that is typical of low-moderate protein stability (<45 °C is low, 45-55 °C is moderate stability, >55 °C is high stability). This approximate correlation between Tm and actual protein stability was established by comparing thermofluor^{TM} data with microdrop screening, another method for validation of protein stability by visually monitoring precipitation over time (Lepre 1998).

Thus the ΔTm of Tris pH 8.0 and Na phosphate pH 7.5 with respect to HEPES pH 7.5 were considered to be significant with ΔTm of +2.8 and +3.0 °C respectively, whereas the ΔTm between HEPES pH 7.5 and Na cocadylate pH 6.5 was +0.7 °C suggesting an insignificant change in stability. We expected the stability of TOPK to increase at higher pH because the predicted isoelectric point (pI) of TOPK is 5.0, hence there was an increase in Tm for Tris pH 8.0 but with Na phosphate pH 7.0 we suspect that the Tm increase maybe due to the presence of phosphate. This suggests a buffer type effect that can impart stability despite the pH being slightly lower than our default pH. It is possible that the mechanism behind this is due to the phosphate ions binding to the active site cleft and stabilizing it.

Thermofluor^{TM} suggests that NaCl increases TOPK stability

In the context of NMR, a high pH is not ideal for observing the maximum number of amide peaks in a ^{1}H, ^{15}N HSQC because the exchange rate of the amide ^{1}H with the solvent is base catalysed. At basic pH, the average lifetime of the protonated amide will be shorter, having greater propensity to be deprotonated, which results in amides resonances with weaker intensity. Hence the pH we selected for studying kinases was typically within the 6.5-7.5 range depending on the stability of the kinase and its predicted pI. Stabilizers in our buffer screen were therefore important in increasing

kinase stability. For example, 150 mM NaCl was often associated with an increase in the Tm by 2 °C or more (data not shown).

The amide peak count for TOPK did not improve at pH 7.0

Considering our thermofluor[TM] results for the buffer, pH and stabilizer effects on the Tm of TOPK, we decided to lower the default buffer pH and validate any potential improvements in the TOPK ^{1}H, ^{15}N HSQC. Thus TOPK was prepared in 50 mM HEPES pH 7.0, 150 mM NaCl, 2.0 mM TCEP, 100 µM AEBSF and 10 % ^{2}H$_2$O. Another buffer solution was prepared as described but with Na phosphate pH 7.0 subsituted for HEPES pH 7.0 because thermofluor[TM] results implied an increase in Tm with Na phosphate pH 7.0. Subsequent ^{1}H, ^{15}N HSQC spectra in the new buffer solutions showed no significant improvements in the peak count (data not shown). Despite changing buffer conditions via buffer type, pH and stabilizers to increase TOPK stability in terms of Tm this did not correlate with a significant improvement in NMR spectra.
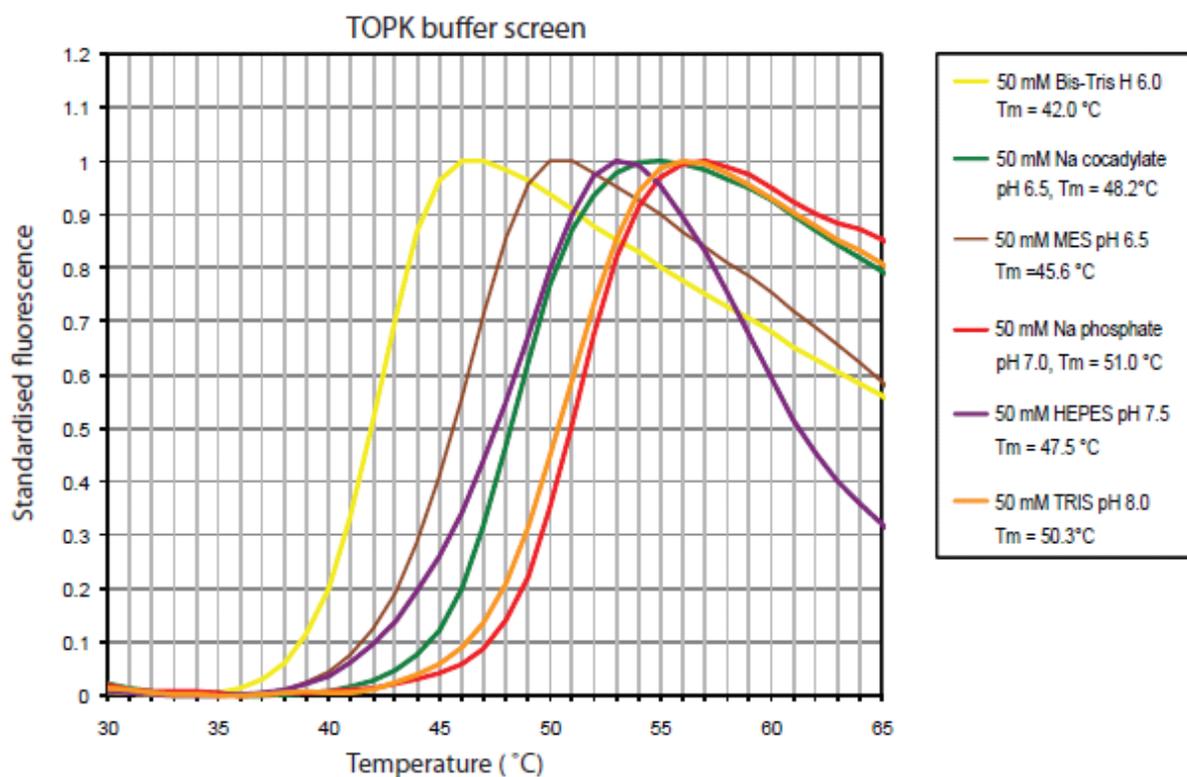
Figure 3.7. TOPK thermofluor.pH screen

TOPK thermal stability was assessed in various buffer types and pH by thermofluor[TM]. In the buffers that were tested, the highest Tm that TOPK displayed was 51.0 °C in 50 mM Na phosphate pH 7.0. Conversely Bis-Tris pH 6.0 produced the lowest Tm most likely because the pH is near the TOPK predicted pI of 5.0. Results could not be obtained for 50 mM MOPS pH 7.0 because the buffer was not available.

Bacterially expressed TOPK is inactive and the active site is occluded

Although the TOPK spectra confirmed that it was folded, the functional state of the protein was unknown. Therefore TOPK activity was assessed using a kinase radioactivity assay (Park, Lin et al. 2006). Phosphorylation of the substrate Histone H3 was not observed for TOPK thus indicating inactivity (Figure). However activity was evident with our positive control, TOPK-Invitrogen. This suggests that our bacterially expressed TOPK is inactive whereas TOPK-Invitrogen expressed from an insect cell line exhibited kinase activity. We presume that bacterially expressed TOPK is lacking the necessary post translational modifications necessary for TOPK activation. Presumably TOPK-Invitrogen is active because in insect cells it is able to receive the necessary post translational modifications from activating kinases that phosphorylate Thr 9 and Thr 198 (Invitrogen 2007). Active TOPK has been produced by Abe et al using cell lysates as a bulk activation method. Cdk1 is known to phoshorylate TOPK Thr 9 but the kinase which phosphorylates Thr 198 remains undetermined (Abe, Matsumoto et al. 2000; Matsumoto, Abe et al. 2004; Abe, Takeuchi et al. 2007). Sequence analysis of TOPK reveals that Thr 198 is located within the activation loop suggesting that this residue is implicated in up-regulating kinase activity in a phosphorylation dependent mechanism analogous to PKA and the tyrosine kinase domain of the insulin receptor (InRK) aswell as many other kinases (Adams 2003).

Furthermore, we hypothesised that TOPK inactivity might correlate with inhibition of the active site cleft. Therefore binding of Mg^{2+}-AMPPNP to the active site was assessed by the 1D NMR ePHOGSY experiment. Results indicates that TOPK does not exhibit any binding to Mg^{2+}-AMPPNP (Figure 3.8 A, B, C) because summation of the two spectra (^{1}H spectrum and presaturated ^{1}H spectrum) show no overall difference, thus suggesting that TOPK active site cleft is occluded. Our collaborators the SGC (data not shown) performed high throuput ligand screening against TOPK using thermofluor^{TM} with their library of kinase inhibitors but found no hits reinforcing the idea that the ATP binding site was somehow occluded. PDZ, a known substrate of TOPK also failed to indicate any obvious interactions (data not shown) suggesting that the substrate binding site was also inhibited.
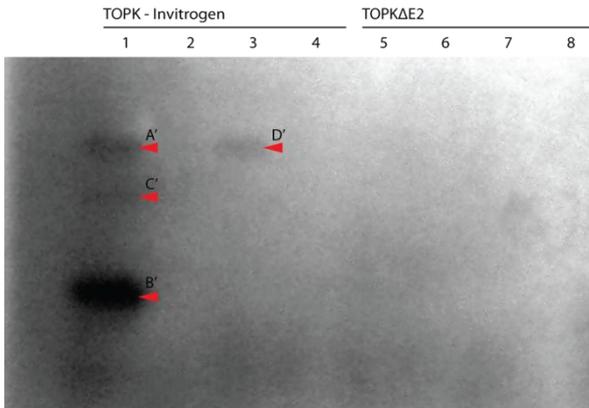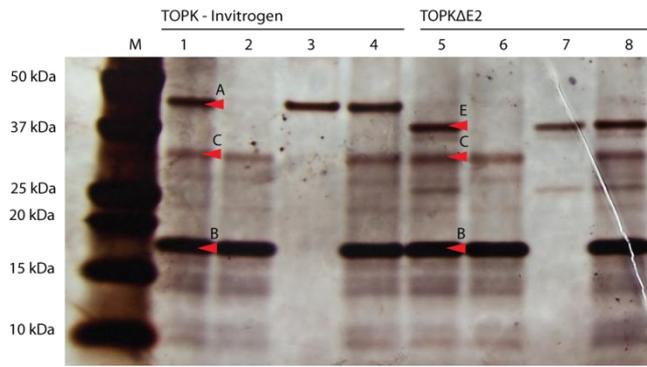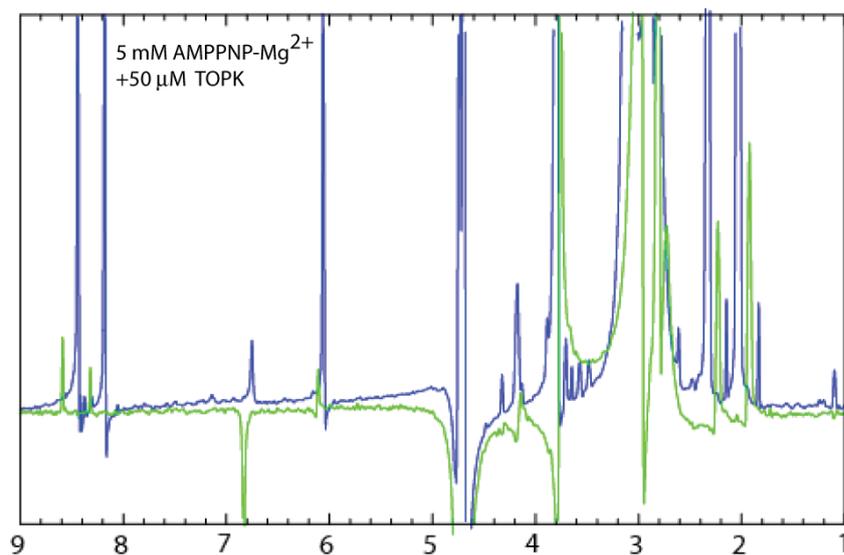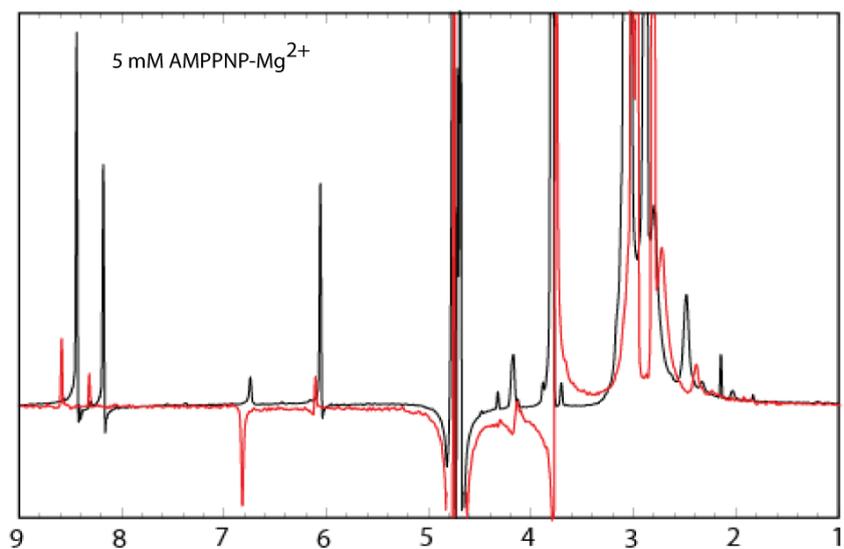
Figure 3.8 A and B. ePHOGSY individual spectra of Mg^{2+}-AMPPNP in the presence and absence of TOPK.

The upper spectrum illustrates the control experiment: ^{1}H spectrum of Mg^{2+}-AMPPNP (black) and its pre-saturated equivalent (red) .The lower spectrum is the actual binding experiment: ^{1}H spectrum of TOPK+ Mg^{2+}-AMPPNP (blue) and its pre-saturated equivalent (green).
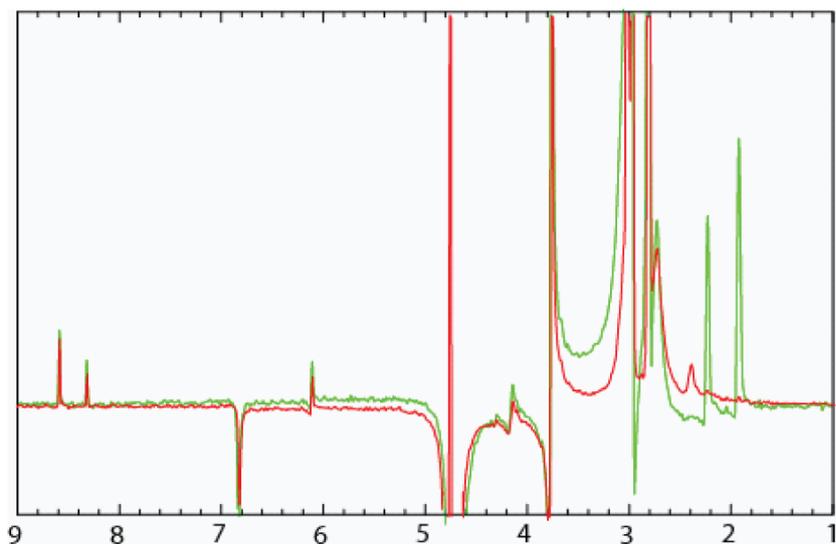
Figure 3.8 C. ePHOGSY difference spectra of Mg^{2+}-AMPPNP, and TOPK + Mg^{2+}-AMPPNP.

The summation of the Mg^{2+}-AMPPNP (red) and TOPK + Mg^{2+}-AMPPNP (green) ePHOGSY spectra illustrates that there is no significant difference between them therefore suggesting that TOPK did not bind AMPPNP. If binding were to occur then this would be indicated by negative AMPPNP resonances in the spectrum after summation of the apo ligand and bound ligand spectra. The ^{1}H resonances occurring near 2 ppm are due to buffer contamination in the TOPK + Mg^{2+}-AMPPNP sample.

Evalution of TOPK10-322 for NMR studies

A putative TOPK N-terminal AID

The mechanism behind TOPK inactivity is unknown but we suspect that the TOPK N-terminus has a putative autoinhibitory role. Therefore nine N-terminal residues of TOPK were deleted (TOPKΔ1-9) and we termed this construct TOPK10-322. Our deletion includes Thr 9, a phosphorylation site believed to up-regulate TOPK activity. It is known that once Thr 9 is phosphorylated by cdk1, the affinity of TOPK for cdk1/cyclin B1 increases (Matsumoto, Abe et al. 2004). We hypothesise that phosphorylation on Thr 9 and binding of cdk1cyclin B1 relieves autoinhibiton imparted by the N-terminus through local and possibly allosteric conformational changes resulting in TOPK activation. Moreover we anticipate that conformational changes associated with activation could improve on the TOPK ^{1}H, ^{15}N HSQC spectrum. This is because conformational changes correlate with changes in dynamics therefore if the backbone amides peaks under intermediate exchange in inactive TOPK experience conformational changes upon activation then this may shift them towards fast or slow exchange, phenomena which are observable.

TOPK10-322 purified to 95 % purity and is monomeric

TOPK10-322 was expressed and purified to 95 % purity based on SDS PAGE. Deletion of TOPKΔ1-9 had negligible effects on TOPK10-322 expression and purification when compared to TOPK. Size exclusion purification of TOPK10-322 produced a single monomeric peak corresponding to a 36 kDa protein (Figure 3.9). In the subsequent anion exchange purification, TOPK10-322 eluted as a sharp peak (Figure 3.10). Exchange of amide ^{2}H for ^{1}H was facilitated by buffer exchanging into TRIZMA buffer at pH 9.0.
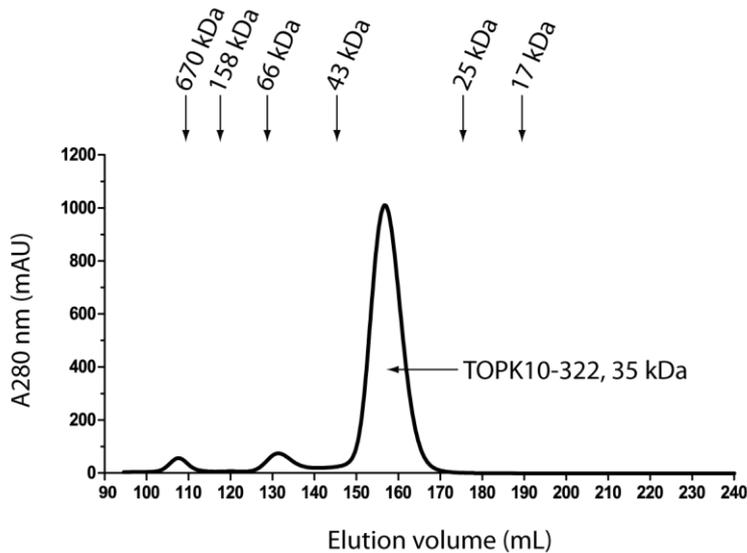
Figure 3.9. TOPK10-322 Size exclusion elution profile.

TOPK10-322 is 35 kDa and appears monomeric by size exclusion purification (S75 column) because it elutes between 150–165 ml. This is a position that corresponds to a 35 kDa species based on comparison with the elution of protein standards.
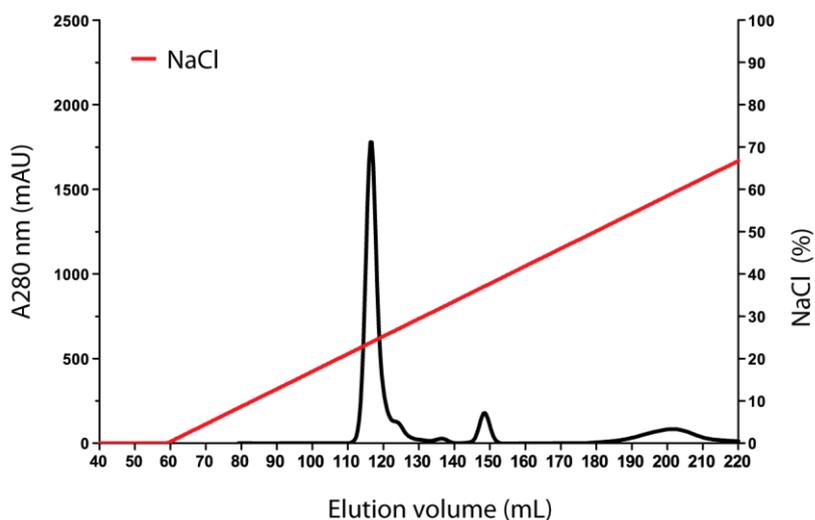
Figure 3.10. TOPK10-322 anion exchange elution profile.

TOPK10-322 elutes as a single sharp peak by anion exchange purification (Resource Q column) at approximately 25% of 1M NaCl (250 mM) with 95 % purity by SDS PAGE.

TOPK10-322 is folded but only an estimated 66 % of expected backbone amide resonances are observable in its ^{1}H, ^{15}N HSQC

A ^{1}H, ^{15}N HSQC spectrum of ^{1}H, ^{15}N labelled TOPK10-322 (Figure 3.11) revealed amide resonance dispersion indicative of a folded protein. Most of the amide resonances showed moderate signal intensity. A minority of weaker amide resonances were also evident suggesting that they corresponded to the dynamic regions of TOPK10-322. In the centre of the spectrum resonance overlap is apparent. The peak count was estimated to be 194 out of 295 expected backbone amide peaks which equates to approximately 66 % of the total. This is a lower than the total backbone amide peak count in TOPK which had 70 % of its expected amide resonances visible. Thus TOPK10-322 did not improve on the TOPK peak count and indicates that TOPKΔ1-9 has made 4 % more peaks enter intermediate exchange and become unobservable. Furthermore it suggests that in TOPK the backbone amide peaks in intermediate exchange are not due to the N-terminal segment. It is possible that relief of autoinhibition imparted by TOPKΔ1-9 and subsequent activation has caused parts of TOPK to become more dynamic on an intermediate exchange time scale. An alternative possibility is that deletion of TOPKΔ1-9 destabilizes TOPK and therefore increases its dynamic behaviour in the intermediate exchange regime.

^{2}H labelling did not improve the TOPK peak count

A ^{1}H, ^{15}N TROSY-HSQC spectrum of ^{2}H, ^{15}N labelled TOPK10-322 (Figure 3.12) did not exhibit any significant improvement in the peak count suggesting that the unobserved backbone amide resonances was due to intrinsic dynamics on an intermediate exchange timescale. However, notable spectral improvements were evident in the line shapes which were finer and the peak intensities were stronger relative to the HSQC spectrum (Figure 3.13).

The backbone solution conformations of TOPK10-322 and TOPK are similar

We compared the backbone conformation of TOPK10-322 with TOPK by superimposing their spectra (Figure 3.14). This revealed that their global backbone conformations are similar with only a few slight CSPs. Deletion of TOPKΔ1-9 therefore does not significantly alter the backbone conformation of TOPK. Differences are evident in the central region of the superimposed spectra where there appears to be less amide resonance overlap with TOPK10-322 presumably because some peaks have undergone CSP, intermediate exchange, or have simply been removed as they are part of the TOPKΔ1-9 segment.
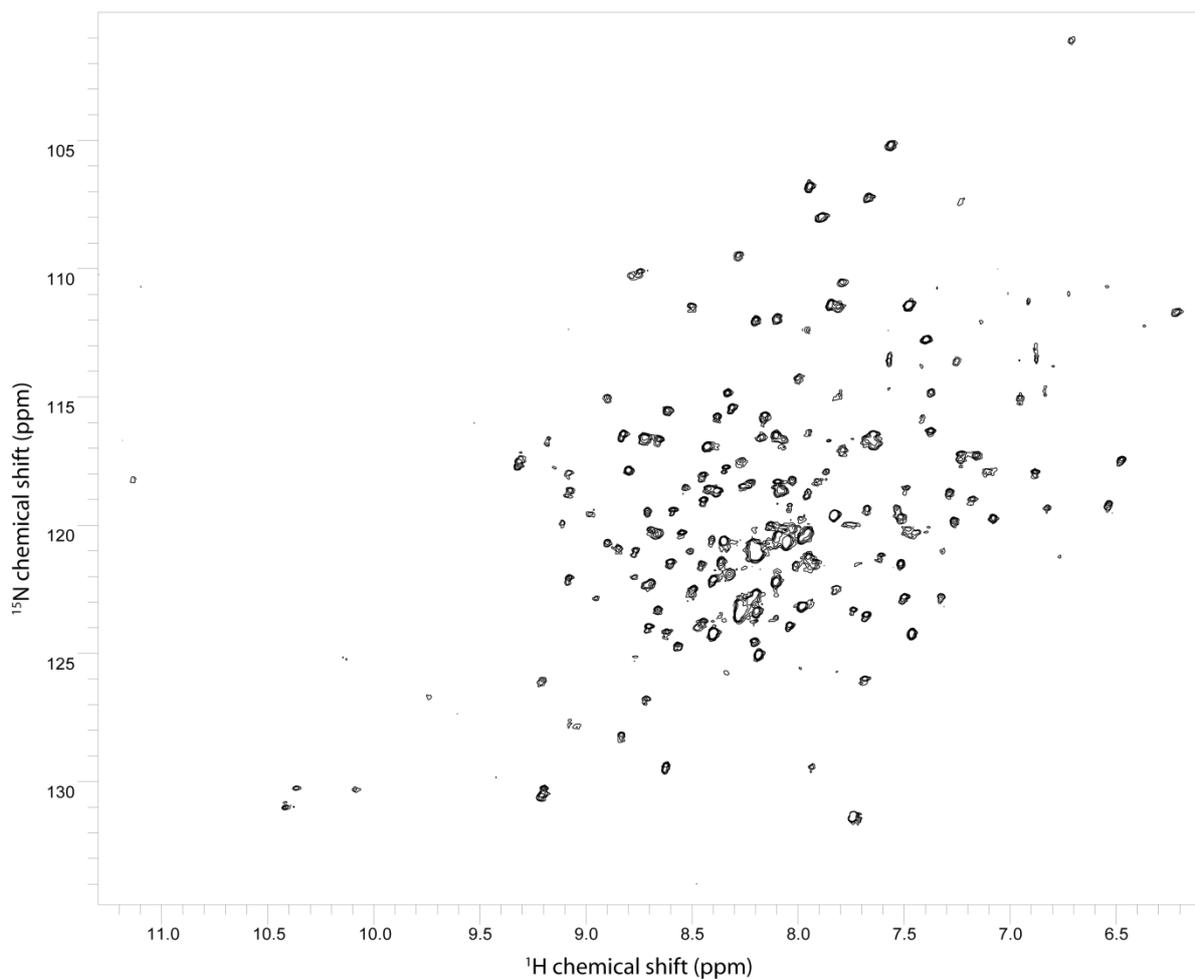
Figure 3.11. TOPK10-322 ^{1}H, ^{15}N HSQC.

Amide resonance dispersion in the ^{1}H, ^{15}N HSQC spectrum of ^{1}H, ^{15}N labelled TOPK10-322 illustrates that the construct is folded. An approximate peak count of backbone amides was estimated to be 211 out of 295 expected peaks (72% of the total). The spectrum was acquired with 200 µM TOPK10-322 in 50 mM HEPES pH 7.5, 100 mM NaCl, 2.0 mM TCEP, 100 µM AEBSF, 10 % ^{2}H$_2$O at 30°C on a Varian 800 MHZ.
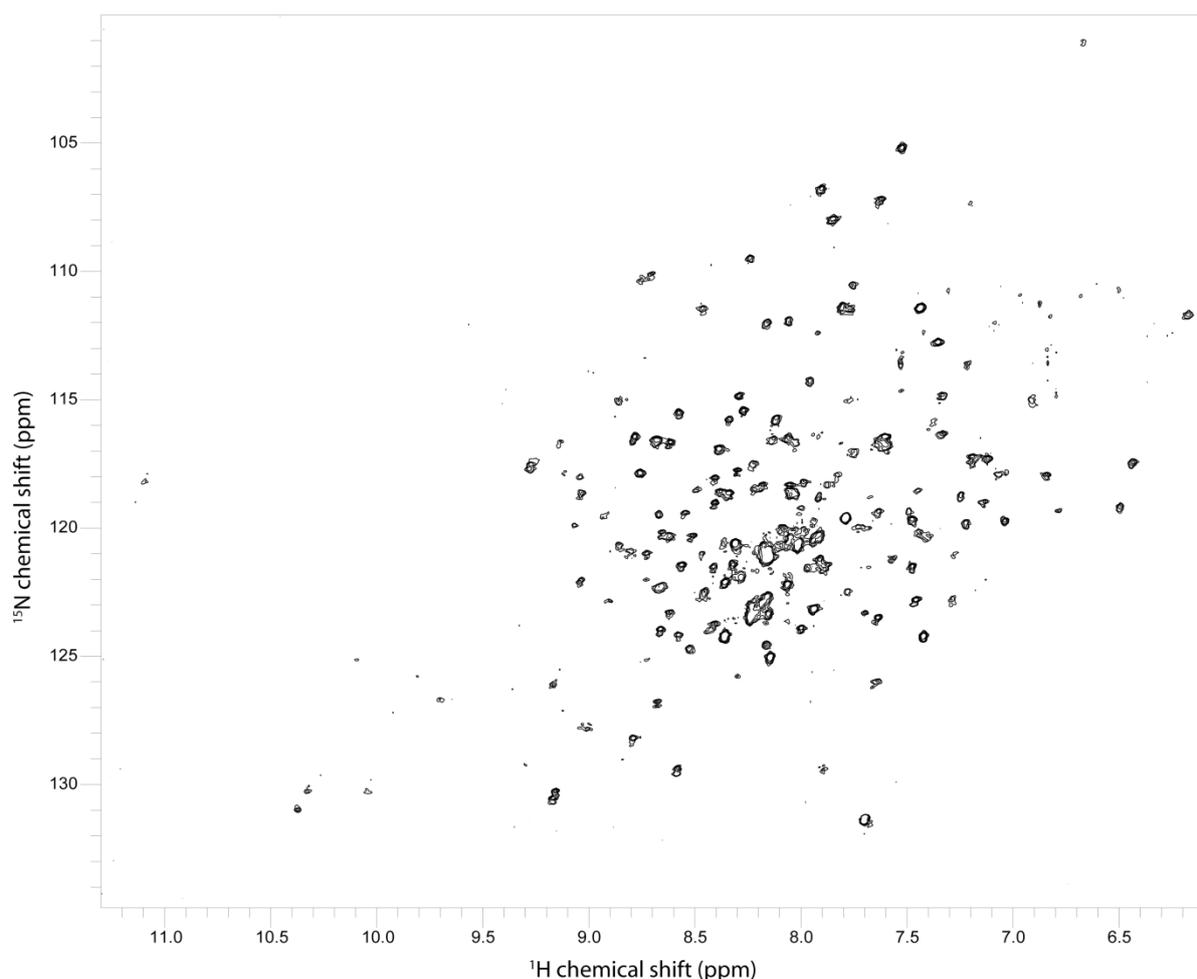
Figure 3.12. TOPK10-322 ^{1}H, ^{15}N TROSY-HSQC.

Only a slight improvement in the line shape is evident in the ^{1}H, ^{15}N TROSY-HSQC of ^{2}H, ^{15}N labelled TOPK10-322 when compared to the ^{1}H, ^{15}N HSQC of ^{1}H, ^{15}N labelled TOPK10-322. No additional backbone amide peaks were evident. The TROSY spectrum was acquired under the same conditions as described for the ^{1}H, ^{15}N HSQC of TOPK10-322 with 200 µM TOPK10-322 in 50 mM HEPES pH 7.5, 100 mM NaCl, 2.0 mM TCEP, 100 µM AEBSF, 10 % ^2H_2O on a Varian 800 MHZ at 30°C.
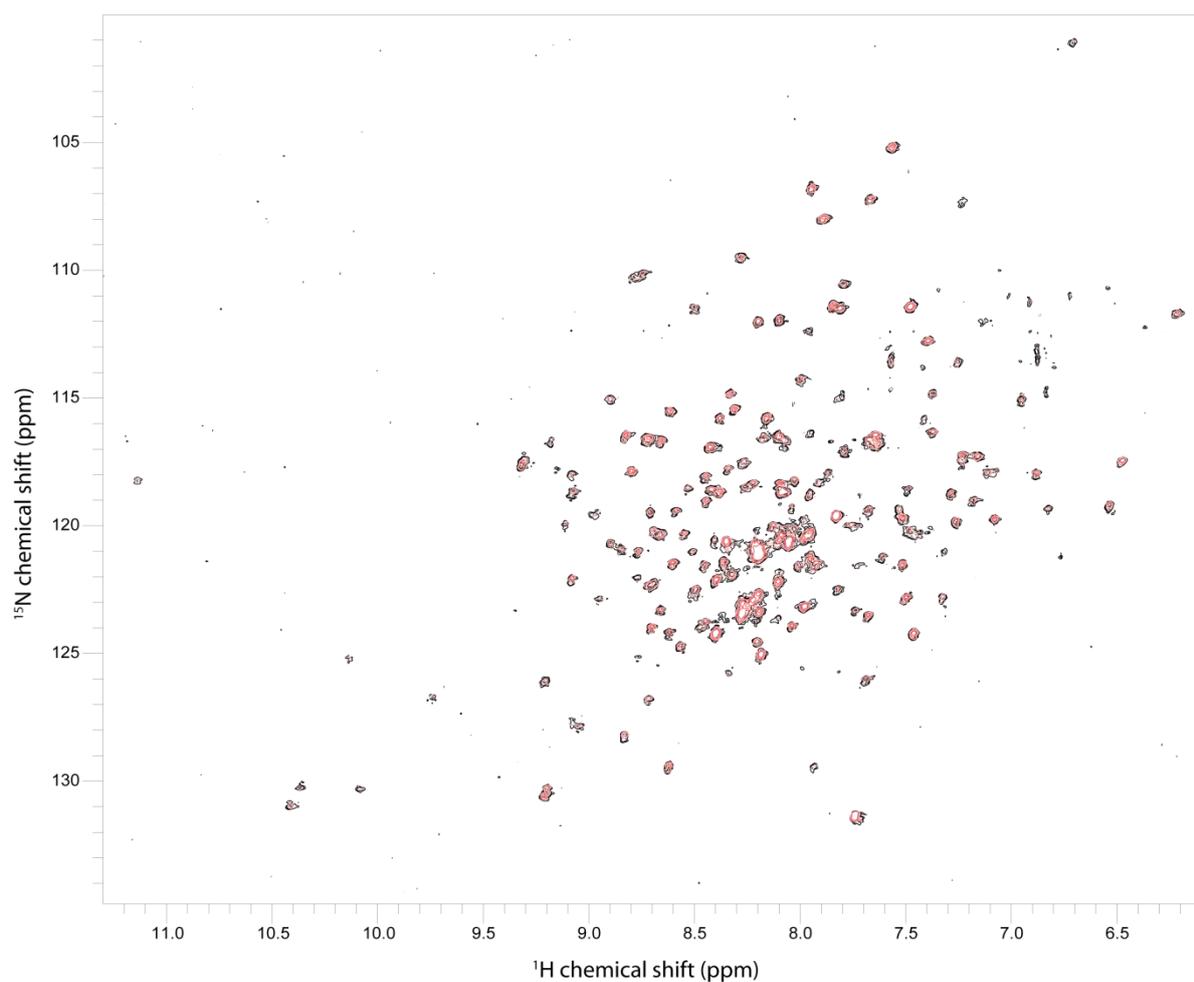
Figure 3.13. TOPK10-322 ^{1}H, ^{15}N HSQC and TROSY-HSQC.

A comparison of the ^{1}H, ^{15}N TOPK10-322 ^{1}H, ^{15}N HSQC (black) and ^{2}H, ^{15}N TOPK10-322 ^{1}H, ^{15}N TROSY-HSQC (red) by superimposition illustrates that they are both identical in peak count. The line shapes of backbone amide resonances in the ^{1}H, ^{15}N TROSY-HSQC are sharper due to the TROSY effect and resonance intensities are slightly stronger due to ^{2}H labelling. Both spectra were acquired with 200 μM TOPK10-322 in 50 mM HEPES pH 7.5, 100 mM NaCl, 2.0 mM TCEP, 100 μM AEBSF, 10 % ^{2}H$_2$O on a Varian 800 MHZ at 30 °C.
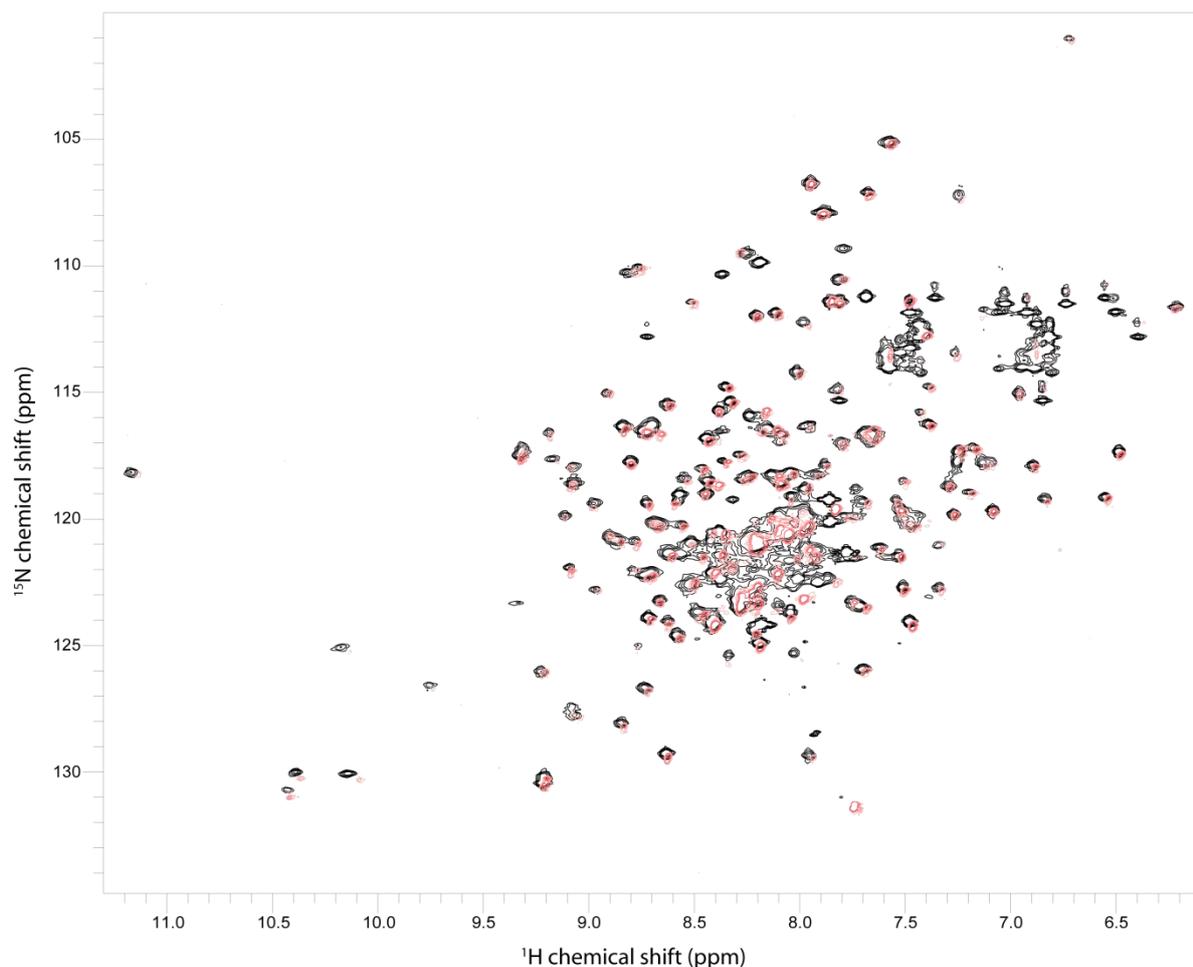
Figure 3.14. TOPK ^{1}H, ^{15}N HSQC and TOPK10-322 ^{1}H, ^{15}N HSQC

Superimposition of the ^{1}H, ^{15}N HSQC spectra for ^{1}H, ^{15}N TOPK (black) and ^{1}H, ^{15}N TOPK10-322 (red) reveals that their backbone conformations are similar. Many of the peaks superimpose with minimal difference in chemical shift, indicating that deletion of TOPKΔ1-9 has not significantly altered the TOPK backbone conformation. However differences are apparent in the central region of the superimposed spectra with the TOPK10-322 displaying less resonance overlap. Each respective spectrum was acquired as previously described.

Thermofluor^{TM} suggests that TOPK10-322 exhibits low-moderate thermal stability in buffers that range from pH 6-8

Buffer screening was conducted on TOPK10-322 by thermofluor^{TM} (Figure 3.15) to find an optimal buffer pH for NMR without compromising stability. As a result of this TOPK10-322 produced Tm values that ranged from 41.6-46.6 °C, which suggested a kinase with low-moderate stability. Results indicated that TOPK10-322 was most stable in 50 mM Tris pH 8.0 with an average Tm of 46.6 °C. Unfortunately its average Tm in HEPES pH 7.5 could not be determined due to non-sigmoidal melting curves. In 50 mM Na phosphate pH 7.0, 50 mM Na cocadylate pH 6.5, 50 mM MES pH 6.5 and 50 mM Bis-Tris pH 6.5, TOPK10-322 produced Tm values corresponding to 46.0, 44.0, 41.6 and 41.7 °C respectively. A comparison of these Tm values for TOPK10-322 with the Tm values of TOPK acquired in the equivalent buffer type and pH reveals that TOPK has greater thermal stability. This suggests that the differences in Tm between TOPK10-322 and TOPK are due to TOPKΔ1-9 causing destabilization. For example, there is a ΔTm of 4 °C between TOPK and TOPK10-322 in 50 mM Na phosphate pH 7.0. Moreover with the exception of MES pH 6.5, there appears to be a correlation between increasing Tm and increasing pH. In contrast TOPK does not exhibit this correlation despite having the same pI as TOPK10-322 because TOPKΔ1-9 does not have any effect on the pI. The lack of correlation in TOPK maybe an artefact associated with sample aggregation, precipitation or degradation. Screening of stabilizers indicated that 150 mM NaCl increased the Tm of TOPK10-322 in a similar extent to TOPK (data not shown).

The amide peak count for TOPK10-322 did not improve at pH 6.5

In order to improve the NMR spectra of TOPK10-322 in terms of its amide peak count we decided to lower the buffer pH from 7.5 to 6.5 based on our evaluation of Tm data for TOPK10-322 whilst considering that TOPK did not improve its peak count that at pH 7.0. A comparison of the ^{1}H, ^{15}N labelled TOPK10-322 ^{1}H, ^{15}N HSQC at pH 7.5 and 6.5 showed that lowering the pH to 6.5 had no significant effect on the peak count therefore there was no change in the number of observable backbone amide resonances. Only a few pH dependent CSPs were evident when lowering the pH (Figure 3.16).
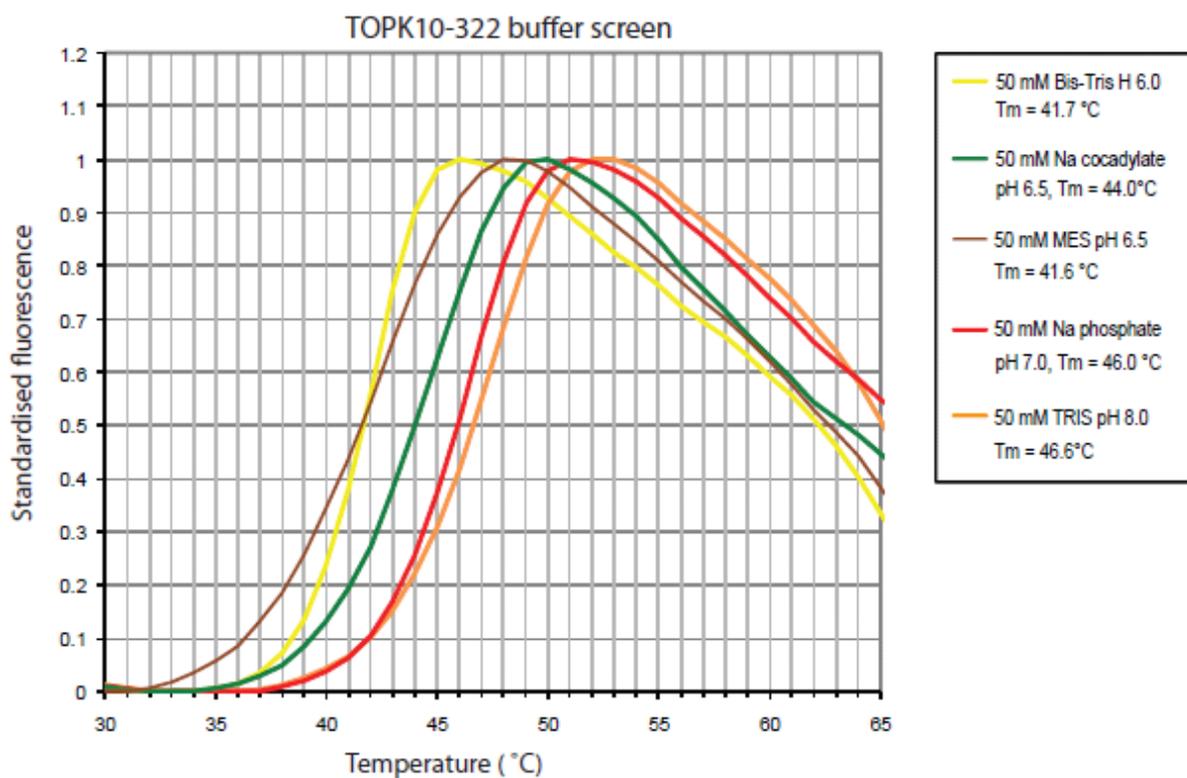
Figure 3.15. TOPK10-322 thermofluor pH screen.

The thermal stability of TOPK10-322 was assessed in different buffer types and pH by thermofluor[TM] to estimate its optimal pH for stability. Of the buffers that were tested 50 mM TRIS pH 8.0 appeared to produce the highest Tm of 46.6 °C. At lower pH the Tm of TOPK is significantly lower, for example 50 mM Bis-Tris pH 6.0 and 50 mM MES pH 6.5 produce Tm values of 41.7 and 41.6 °C respectively. However in 50 mM Na cocadylate pH 6.5, the Tm is 44.0 °C which is higher than the Tm produced by 50 mM MES pH 6.5 suggesting the difference in stability could be a buffer type effect. Results for 50 mM HEPES pH 7.5 was omitted due to non-sigmoidal melting curves. 50 mM MOPS pH 7.0 was not available for testing.
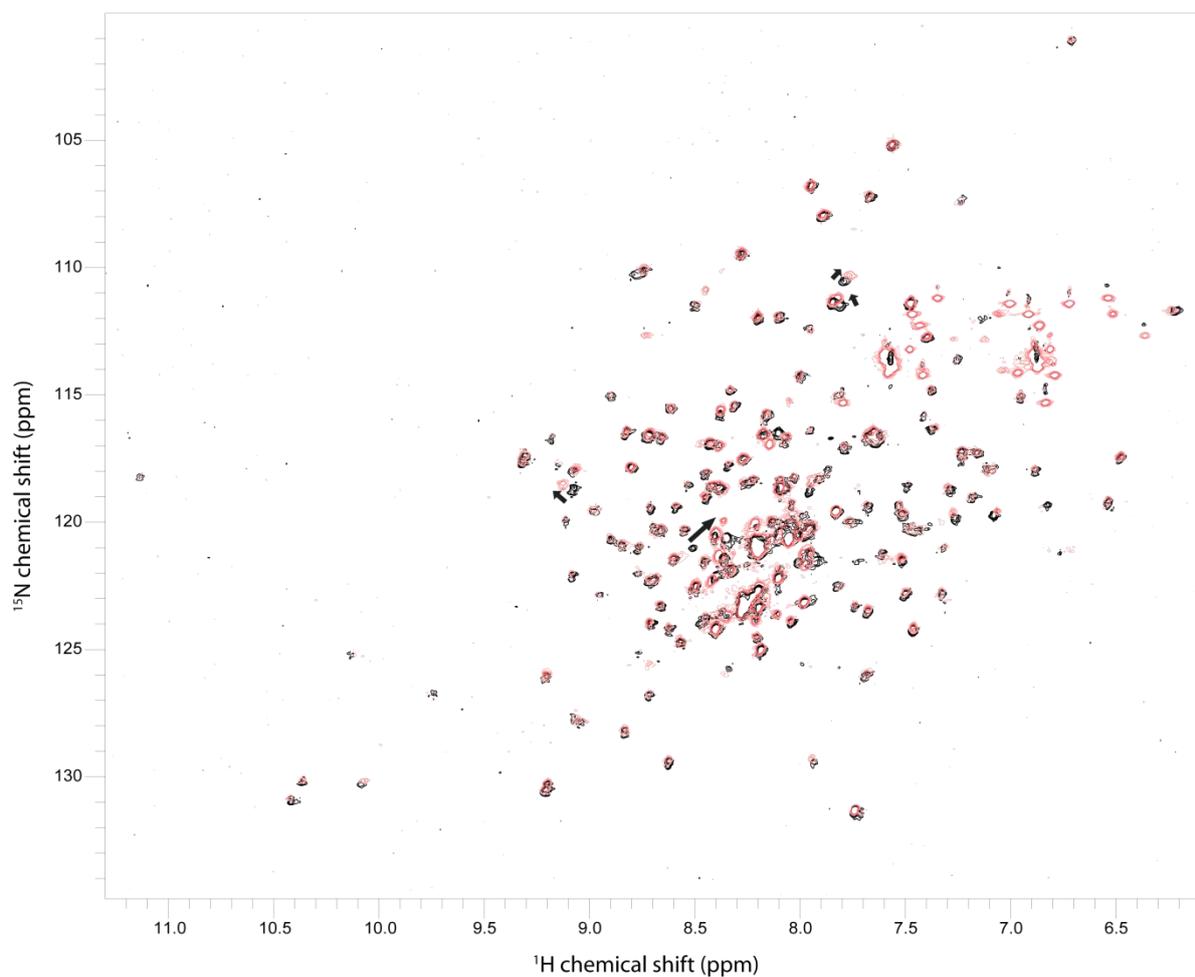
Figure 3.16. TOPK10-322 ^{1}H, ^{15}N HSQC pH 7.5 and 6.5.

In order to increase the peak count of TOPK10-322 the pH of the buffer was lowered from 7.5 (blue) to 6.5 (red). However an overlay of the ^{1}H, ^{15}N HSQC of TOPK10-322 at pH 7.0 and pH 6.5 revealed that additional peaks were not evident at pH 6.5. Only a few pH dependent CSPs were notable as indicated by the arrows. The spectra were acquired with 0.2 mM TOPK10-322 buffer in 50 mM HEPES pH 7.5 (blue), pH 6.5 (red), 100 mM NaCl, 2.0 mM TCEP, 100 μM AEBSF, 10 % ^2H_2O at 30 °C on a Varian 800 MHz.

TOPK10-322 and TOPK_D9_D198 are inactive

Bacterially expressed TOPK10-322 was shown by its ^{1}H, ^{15}N HSQC spectrum to be folded although its functional state was unknown. TOPK10-322 was tested for activity in a kinase radioactivity assay (Park, Lin et al. 2006) which revealed it to be inactive because phosphorylation of the substrate Histone H3 was not observed (Figure). This suggests that TOPKΔ1-9 is not sufficient in activating TOPK. It is possible that residues imparting autoinhibition on TOPK may extend beyond Thr 9 because the beginning of the kinase boundary is defined from Ser 32 based on primary sequence databases (kinbase ; uniprot). Alternatively our hypothesis of an N-terminal autoinhibitory segment could be wrong. Both TOPK and TOPK10-322 lack Thr 198 phosphorylation which could also explain the inactivity of our TOPK constructs. Consequently we produced single and double phosphorylation mimics by mutating Thr 9 and Thr 198 to Asp 9 and Asp 198, respectively to try and active TOPK. However these constructs did not display any activity when tested with the NADH coupled activity assay (data not shown).

Thermofluor^{TM} assay suggests TOPK_T9D_T198D is folded and exhibits low-moderate thermal stability in buffers range from pH 6-8

Kinase activity was not detected for TOPK_T9D_T198D and we were unsure of its folded state. Due to time constraints a ^{1}H, ^{15}N labelled sample of TOPK_T9D_T198D could not be prepared for structural analysis. Instead thermofluor^{TM} was used to quickly gain insight into the stability of the double mutant. Buffer screening by thermofluor^{TM} suggested that TOPK_T9D_T198D was folded because it exhibited sigmoidal melting curves characteristic of a folded protein (Figure 3.17). Tm values for TOPK_T9D_T198D ranged from 42.9-46.6 °C, suggesting low-moderate stability. The highest Tm corresponded to 50 mM TRIS pH 8.0. TOPK_T9D_T198D stability appears to increase with pH although the ΔTm with increasing pH starting from pH 6.5 was less than 2.0 °C. This suggests that the change in stability within the pH range of 6.5-8.0 was not significant. Thermofluor^{TM} results also indicate that TOPK_T9D_T198D was less stable in 50 mM MES pH 6.5 compared to 50 mM MOPS pH 6.5 with a ΔTm = 2.3 °C, which suggests a potential buffer type effect.

TOPK10-322
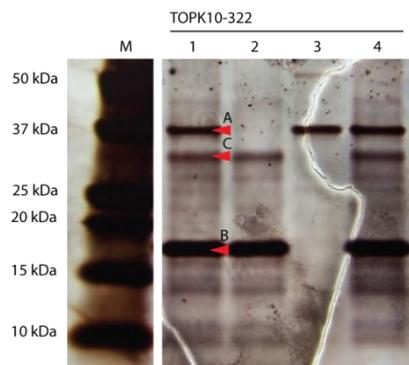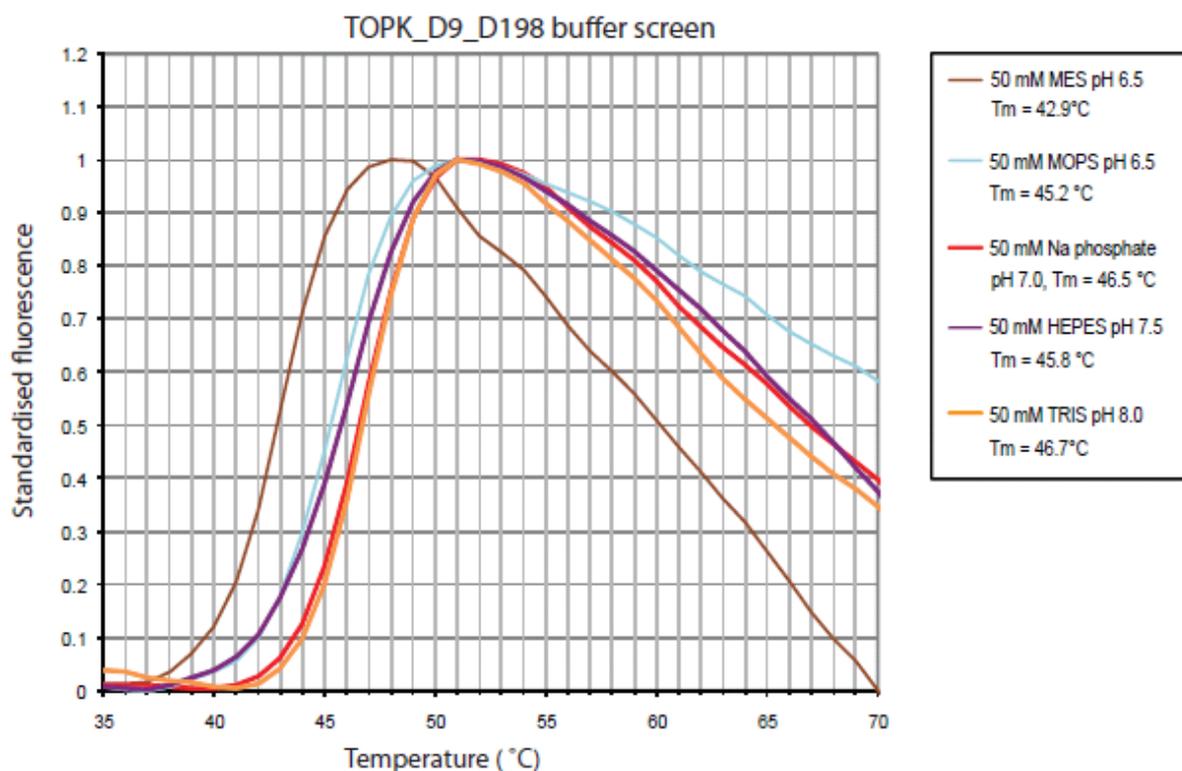
Figure 3.17. TOPK_T9D_T198D thermofluor pH screen.

Characterisation of TOPK_T9D_T198D thermal stability in different buffer types and pH was conducted by thermofluor[TM]. Sigmoidal melting curves typical of a folded protein were produced suggesting that the double mutation has not unfolded the protein. Stability of TOPK_T9D_T198D appears to correlate positively with high pH. Both 50 mM Bis-Tris pH 6.0 and 50 mM Na cocadylate pH 6.5 were omitted as they produced non-sigmoidal melting curves.

Evalutation of Pim1 for NMR studies

Pim1 purified to 90 % purity despite His$_6$-tag cleavage issues

Another kinase target provided by the SGC for NMR analysis was Pim1, a well characterised therapeutic target. Pim1 (Figure 3.18, 3.19) was initially expressed and purified to 90 % purity based on SDS PAGE but at half the expected yield due to difficulty encountered with Pim1 His$_6$-tag cleavage mediated by TEV protease. Only 50 % cleavage was observed by SDS PAGE which meant that 50 % of Pim1 retained its N-terminal His$_6$ tag and could not be further purified. Despite attempts to optimise the cleavage efficiency of TEV protease, it was discovered that due to the constitutive activity of Pim1 (Qian, Wang et al. 2005), the N-terminal His$_6$-tag was phosphorylated by Pim1 (Jacobs, Black et al. 2005). This interfered with the ability of TEV protease to cleave the TEV site adjacent to the His$_6$-tag but the issue was resolved by adding alkaline phosphatase. Purified Pim1 produced a peak with a slight shoulder suggesting the presence of contaminants that co elute at a similar size (Figure 3.22) although SDS PAGE indicates 90% purity.

Pim1 is constitutively active

A radioactive kinase activity assay (Park, Lin et al. 2006) was used to assess the activity of Pim1 and confirmed it to be constitutively active because phosphorylation of the generic substrate Histone H3 was detected in the autoradiograph. However Pim1 did not exhibit autophosphorylation (Figure 3.28). It is believed that because Pim1 is constitutively active, undergoing autophosphorylation during expression but not dephosphorylation (Bullock, Debreczeni et al. 2005). Therefore we postulate that all Pim1 autophosphorylation sites were already phosphorylated which would account for the lack of autophosphorylation with Pim1 in the radioactive assay. Elucidation of the Pim1 phosphorylation state was not confirmed.
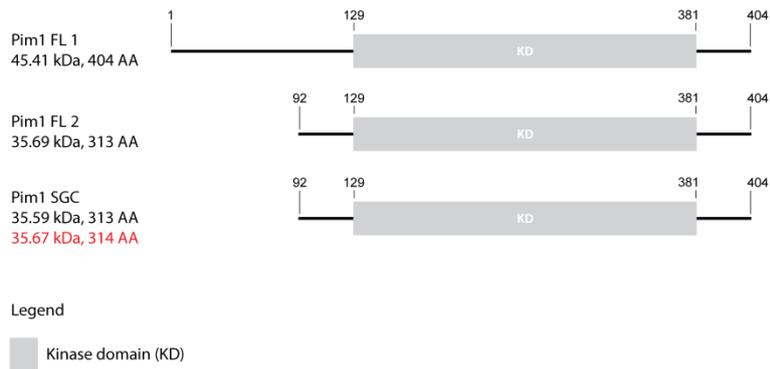
Figure 3.18. Pim1 full length 44kDa and 34kDa variants.

Our Pim1 construct is 99% identical to the Pim1 full length 34kDa shorter variant. Residue Gly 250 in the 34kDa Pim1 variant however is replaced by Arg 250 in our construct due to a single residue polymorphism.

```
          100         110         120         130         140
 MLLSKINSL AHLRAAPCND LHATKLAPGK EKEPLESQYQ VGPLLGSGGF

          150         160         170         180         190
GSVYSGIRVS DNLPVAIKHV EKDRISDWGE LPNGTRVPME VVLLKKVSSG

          200         210         220         230         240
FSGVIRLLDW FERPDSFVLI LERPEPVQDL FDFITERGAL QEELARSFFW

          250         260         270         280         290
QVLEAVRHCH NCGVLHRDIK DENILIDLNR GELKLIDFGS GALLKDTVYT

          300         310         320         330         340
DFDGTRVYSP PEWIRYHRYH GRSAAVWSLG ILLYDMVCGD IPFEHDEEII

          350         360         370         380         390
RGQVFFRQRV SSECQHLIRW CLALRPSDRP TFEEIQNHPW MQDVLLPQET

          400
AEIHLHSLSP GPSK


Predicted pI = 5.7
```

Figure 3.19. Pim1 primary sequence.

The primary sequence encoded for our Pim1 construct and its predicted pI are shown. Pim1 is constitutively active and does not require phosphorylation for activation but it can undergo autophosphorylation.

Figure 3.20. SDS PAGE of His$_6$-Pim1 Ni^{2+} NTA purification.

Elution fractions were pooled from fraction number 3-9.
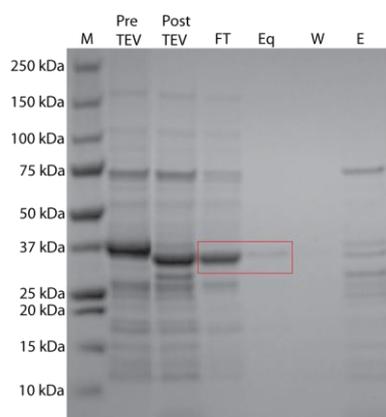
Figure 3.21. SDS PAGE of Pim1 post His$_6$-tag cleavage.

Pim in the FT and Eq fractions were pooled after overnight incubation with TEV followed by Ni^{2+} NTA purification to purify from incomplete His$_6$-tag cleavage, free His$_6$-tag, and other protein contaminants.
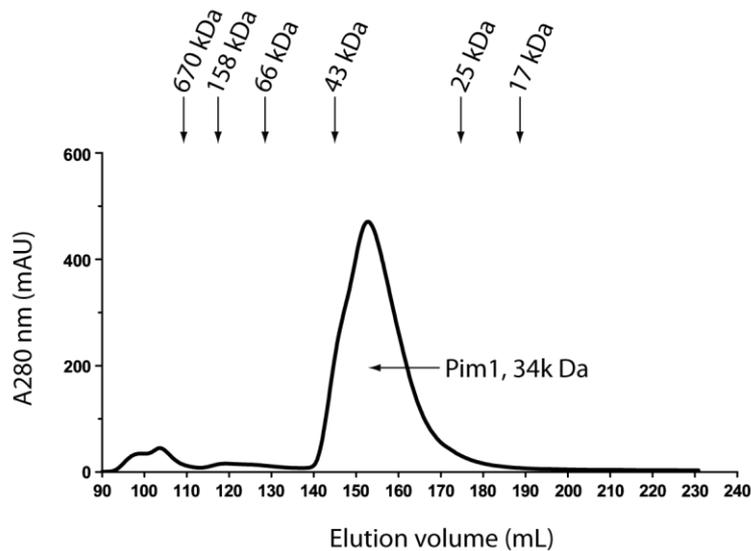
Figure 3.22. Pim1 size exclusion elution profile and SDS PAGE.

Pim1 purified by size exclusion purification (S75 26/60 column) elutes approximately between 145–165 ml. A slight shoulder in the peak is due to a large molecular weight contaminant that elutes at a similar size and shape. Pim1 is 34 kDa and elutes at a position that correlates to a 34 kDa species based on comparison with protein standards eluted by size exclusion. Pim1 therefore appears to be monomeric. Fractions highlighted by the red box were pooled together.

Pim1 is folded but the majority of expected backbone amide resonances are unobservable in its ^{1}H, ^{15}N HSQC

Having deduced that purified Pim1 is constitutively actively we decided to analyse its backbone solution structure. A ^{1}H, ^{15}N HSQC spectrum of ^{1}H, ^{15}N labelled Pim1 (Figure 3.23) reveals that a significant number of backbone amide resonances are missing thus yielding a very low peak count for a protein expected to have 302 backbone amide peaks. Most of the observable backbone amide resonances appear near the centre of the spectrum with limited dispersion. They also display non-uniform resonance intensities suggesting that Pim1 is dynamic in solution. Only four out of six Trp side chain amide Hε resonances are visible but their presence indicates a folded protein. Many of the expected backbone amide resonances that were not observed were likely undergoing intermediate exchange phenomena but also the large number resonances that broaden out suggests possible aggregation resulting in a larger molecular weight species and thus a slower tumbling rate. Another potential issue that might have contributed to the low spectral quality of Pim1 was sample heterogeneity arising from multiple phosphorylated Pim1 species. Precipitation of Pim1 after acquisition of its ^{1}H, ^{15}N HSQC at 25 °C was also evident thus indicating low stability. This would contribute towards the uniform reduction in amide peak intensity. Considering all these difficulties associated with Pim1 we conclude that a backbone assignment of Pim1 is not feasible without further extensive optimisation.
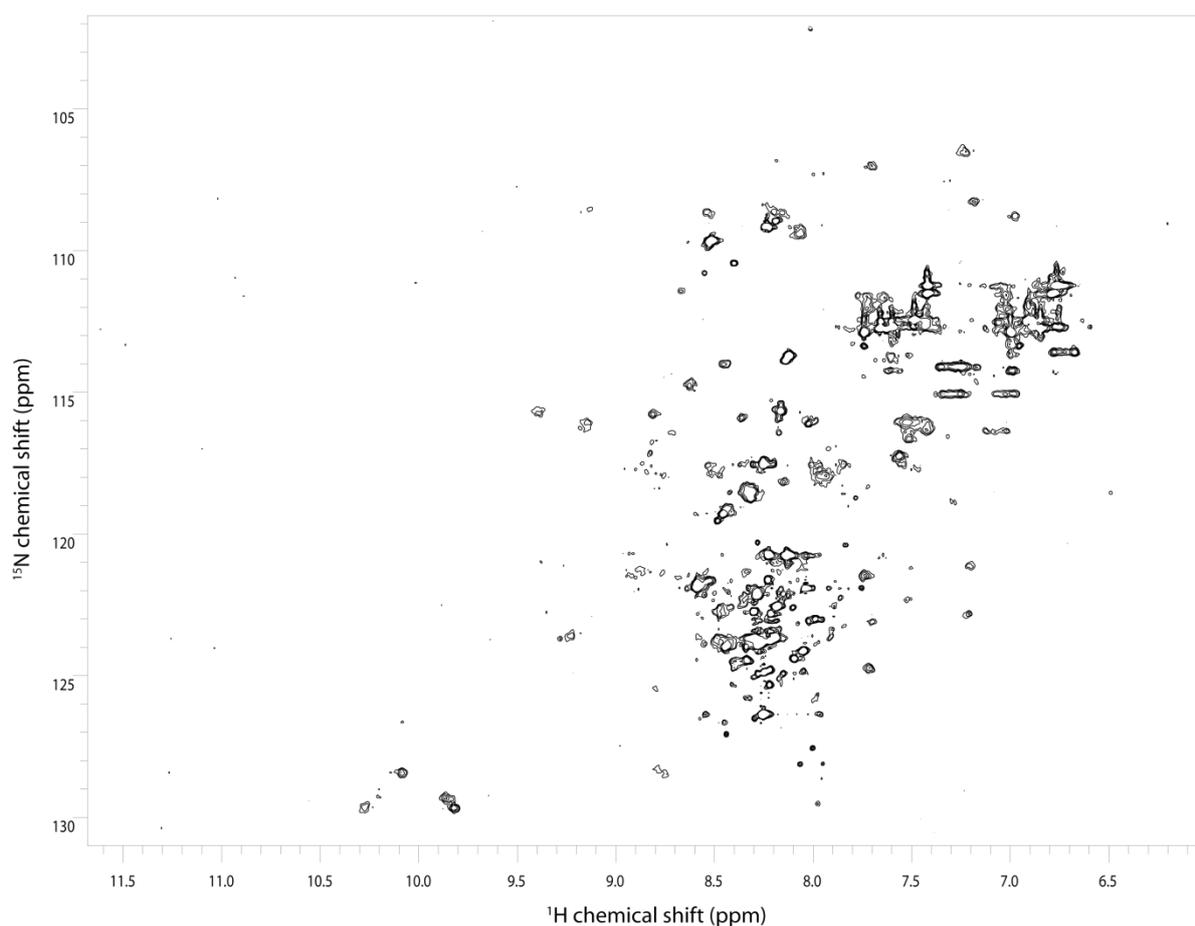
Figure 3.23. Pim1 ^{1}H, ^{15}N HSQC.

Pim1 exhibits a significantly lower than expected amide peak count for a 34k Da protein composed of 313 residues. Backbone amide resonance dispersion appears to be narrow and clustered around the center of the spectrum. Dispersed Trp side chain amide ^{1}Hε resonances are evident in the ^{1}H, ^{15}N HSQC spectrum (9.8-10.3 ppm ^{1}H dimension, 128-130 ppm ^{15}N dimension) suggesting that Pim1 is folded. Most of the amide resonances are unobservable because they are suspected to be in intermediate exchange or broadened out due to intrinsic intermediate dynamics and aggregation respectively. To confirm the folded state of Pim1 an activity assay was conducted and demonstrated that Pim1 is constitutively active indicating that it is folded. The spectrum was acquired with 200µM ^{1}H, ^{15}N Pim1 at 25°C in 50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na azide, 10 % ^{2}H$_2$O on a Varian 800 MHZ.

Pim1 is intrinsically dynamic and is conformationally unstable

In order to further understand the dynamic behaviour of Pim1 we examined the B-factors (also known as the temperature factor) associated with the crystal structure of Pim1 in complex with the bisindolymaleimide I, hydrochloride (BIM-1) inhibitor because B-factors describe the thermal motion about each atom and indicate mobility (Willis 1975; Ringe and Petsko 1986). Figure 3.24 reveals that intermediate and high B-factors (based on the B-factor scale) span the Pim1 structure thus indicating that Pim1 exhibits intermediate and high levels of mobility over many parts of both Pim1 kinase lobes even with Bim-1 bound to the active site cleft. Evidently the B-factors support the notion previously inferred from the Pim1 ^{1}H, ^{15}N HSQC that Pim1 is a dynamic entity but it is likely that the true extent of Pim1 dynamics is greater in apo form because there would be no constraints imposed by the inhibitor on Pim1.

However, a potential caveat associated with the B-factors is the influence of crystal lattice contacts which can mask or reduce regions of mobility. Even so, intermediate and high levels of mobility are still apparent in the Pim1 B-factors. Considering the B-factors, ^{1}H, ^{15}N HSQC and activity data presented for Pim1, we postulate that the dynamic properties of Pim1 are inherently linked to its constitutively active nature. The rationale for this is based on the notion that the active site cleft of constitutively active Pim1 dynamically interconverts between an ensemble of catalytically productive and non-productive conformations. Therefore we expect such conformational changes around the active site cleft to convey mobility across both Pim1 kinase lobes and about the connecting hinge since the hinge and interface between the lobes define the active site cleft.
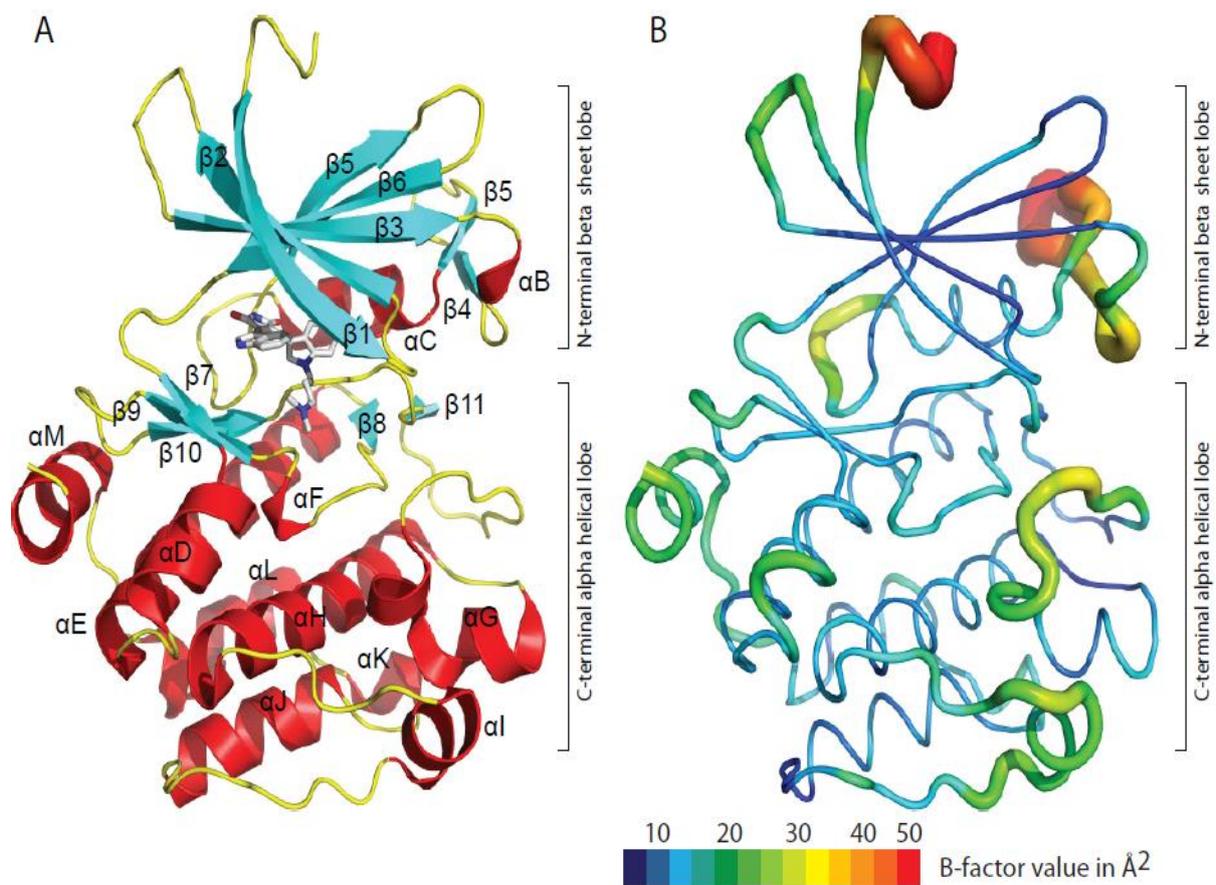
Figure 3.24 Pim1-BIM1 inhibitor crystal structure and B-factors.

Pim1 in complex with the BIM-1 inhibitor (PDB: 1XWS) exhibits a canonical kinase fold comprised of an N-terminal ATP binding lobe and a larger C-terminal substrate binding lobe. Notably the activation loop is well defined. B-factors describe the dynamic and rigid regions of Pim1 which displays high (red and orange), intermediate (green and yellow) and low (blue) B-factors across both the N-terminal and C-terminal lobes.

Thermofluor^{TM} suggests that Pim1 exhibits low-moderate stability in buffers that range from pH 6-8

Optimisation of the default Pim1 NMR buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02% Na Azide, 10 % ^2H_2O) was necessary in order to try and improve its stability and amide peak count. Different buffer conditions were screened using the thermofluor^{TM} assay although evaluation of Pim1 stability was difficult because Pim1 often produced atypical sigmoidal melting curves with a high initial fluorescence suggesting that the protein sample was precipitated, unfolded and aggregated. Extensive efforts to resolve this required high speed centrifugation and re-purification by size exclusion to remove precipitation and aggregation respectively.

Figure 3.25 presents the thermofluor^{TM} results for Pim1 in different buffer types in which the effect of pH on Pim1 stability was assessed. Our data shows that the Tm of Pim1 ranges from 43.5-46.5 °C for the tested buffer types with pH varying from 6.0-8.0 and suggests that Pim1 has low-moderate stability. In addition, the Tm of Pim1 does not vary significantly with pH and there is no apparent correlation between Tm and pH with the lowest Tm produced by 50 mM Na phosphate pH 7.0 and the highest Tm produced by 50 mM MES pH 6.5. Compared to the default buffer type of 50 mM HEPES pH 7.5, which produced a Tm of 44.7 °C, the Tm of Pim1 in 50 mM MES pH 6.5 was 46.5 °C, corresponding to a ΔTm of 1.8 °C. This suggests an increase in stability even though MES pH 6.5 is closer than HEPES pH 7.5 to the Pim1 predicted pI of 5.7. It is possible that this result is an artefact or the predicted pI is not 5.7.

The amide peak count for Pim1 did not improve at pH 6.5

Despite this we substituted 50 mM MES pH 6.5 for 50 mM HEPES pH 7.5 in our default NMR buffer in order to try and improve the amide peak count because thermofluor^{TM} results suggested that there was an insignificant change in the stability of Pim1 at pH 6.5 compared to the default pH 7.5. Subsequently the effect of lowering the pH on Pim1 in terms of its amide peak count and stability was validated

by acquiring a ^{1}H, ^{15}N HSQC at 25 °C. Disappointingly the resulting spectrum (data not shown) did not yield any significant improvement in the amide peak count compared to Pim1 in its default NMR buffer. Likewise no improvement in stability was apparent because precipitation was evident after acquiring the ^{1}H, ^{15}N HSQC spectrum, thus suggesting that an increase in Tm of 1.8 °C at pH 6.5 versus pH 7.5 was not sufficient for a significant change in Pim1 stability or the thermofluor^{TM} data was a false estimate. The lack of correlation between buffer pH and stability made it difficult to predict an optimal pH for NMR studies without significantly compromising protein stability.

Lowering the pH did not improve the amide peak count of Pim1 and given that Pim1 produces a low quality ^{1}H, ^{15}N HSQC and exhibits low stability suggests that the spectrum and stability of Pim1 are intrinsically correlated to some extent. It was therefore important to try and address this as a potential means of improving the spectral quality of Pim1. Further estimates of the Tm for Pim1 in different buffers even in the presence of 150 mM NaCl as a stabiliser proved difficult due to thermofluor^{TM} data yielding non-sigmoidal melting curves. Optimisation of Pim1 NMR spectra in terms of its buffer conditions was therefore not pursued any further.

BIM-1 binds to Pim1 and increases its stability

Our previous efforts to screen different buffer conditions with varying pH and additives to find an optimal condition that improved the NMR spectra and stability of Pim1 proved unsuccessful. Therefore we explored an alternative strategy for optimising the stability of Pim1 involving the use of ligands that promote stabilization of Pim1. The rationale behind this was to suppress the constitutive activity of Pim1 and thus change its dynamic properties so that Pim1 adopts a single stable conformation as opposed to the multiple conformations that we suspect it samples when constitutively active. With a reduction in Pim1 mobility we expect a concomitant increase in amide peak intensity, CSPs and a change in the dynamics of its backbone amide resonances.

Before proceeding with the ligand mediated stabilization strategy by NMR we first assessed Pim1 ligand binding using thermofluor^{TM} by starting with a known tight binding inhibitor termed Bisindolymaleimide I, Hydrochloride (BIM-1), a water soluble reversible ATP competitive inhibitor that binds and inhibits PKC with a Ki of 10 nM and PKA with a Ki of 2 μM. However a Ki for Pim1 binding to BIM-1 is not known but the interaction is presumed to be relatively tight because it was co-crystallised with Pim1 by the SGC (PDB: 1XWS). AMPPNP was a weak binding Pim1 ligand that we also assessed by thermofluor^{TM} for comparison with BIM-1. The resulting ΔTm induced by each ligand was determined to infer the change in Pim1 stability. In the presence of BIM-1 the Tm of Pim1 increases with increasing concentrations of BIM-1 suggesting that the BIM-1 binds to Pim1 in a stabilizing manner (Figure 3.26). For example it is evident that at 1, 10 and 100 μM BIM-1 the ΔTm relative to apo Pim1 increases by 2.3, 4.7, and 8.3 °C respectively suggesting a tight binding and stable interaction. In contrast, addition of AMPPNP up to 1 mM exhibited a ΔTm = -0.6 °C indicating a decrease in Tm but because the ΔTm < 2 °C we considered this to be an insignificant. It also suggests that the interaction is very weak (Figure 43). We believe that there is at least a weak interaction because Pim1 is constitutively active and therefore receptive to nucleotide binding.
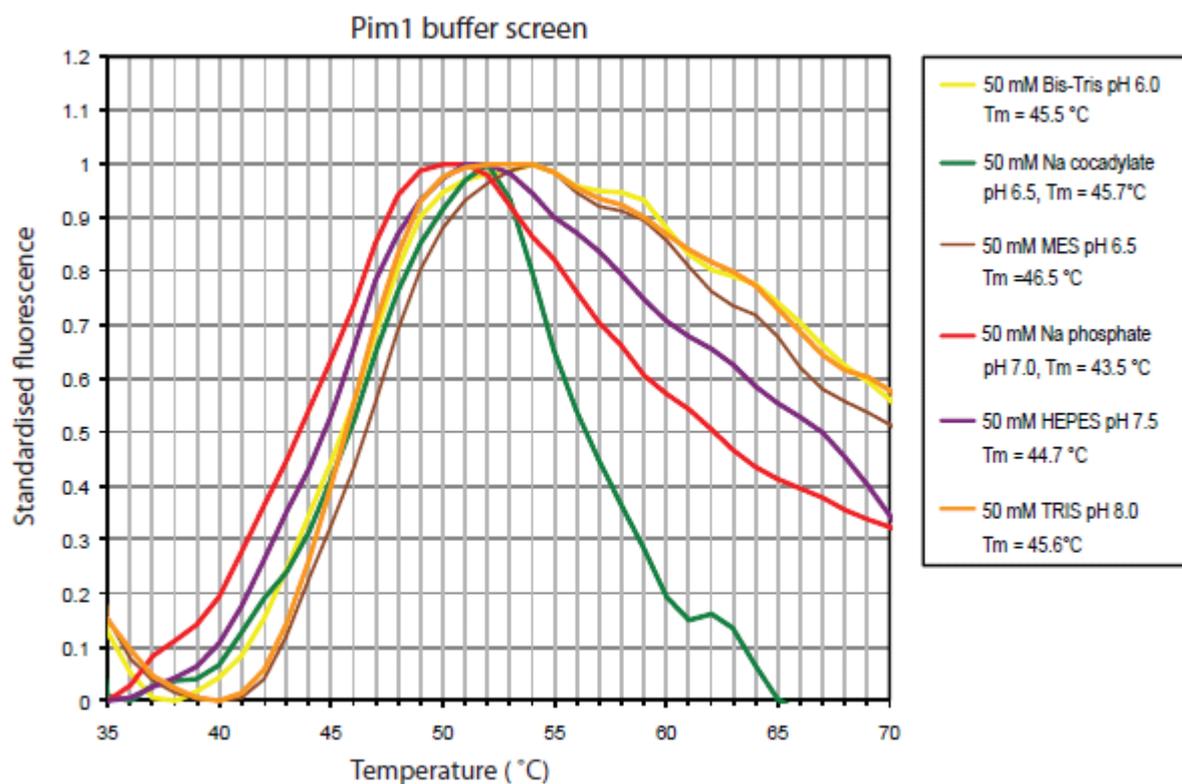
Figure 3.25. Pim1 thermofluor pH screening.

Pim1 thermal stability was characterised in various buffer types and pH. No apparent correlation is evident between the Tm of Pim1 and buffer pH. Our thermofluor[TM] data suggests that Pim1 has a maximum Tm of 45.6 °C when the buffer type and pH is 50 mM MES pH 6.5. An average Tm for 50 mM MOPS pH 7.0 was not determined due non-sigmoidal melting curves.
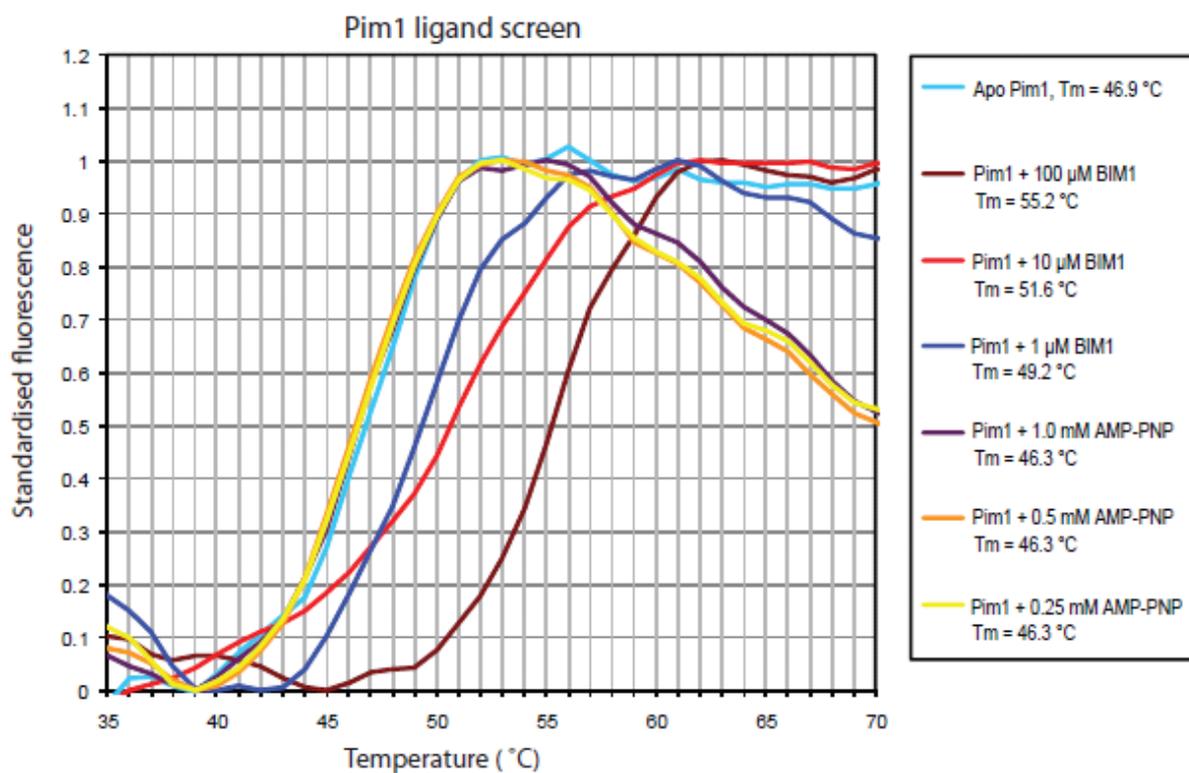
Figure 3.26. Pim1 ligand binding by thermofluor.

Binding of BIM-1 and AMPPNP to Pim1 were tested by thermofluor[TM]. An insignificant ΔTm is associated with Pim1 in the presence of 0.25, 0.5, and 1.0 mM AMPPNP indicative of a weak interaction. However in the presence of 1.0, 10.0, and 100.0 μM BIM-1 there is a significant increase in the Tm of Pim1 by 2.3, 4.7 and 8.3 °C respectively suggesting that a high affinity thermally stable complex is formed with BIM-1.

Pim1-BIM-1 inhibitor complex did not improve the ^{1}H, ^{15}N HSQC spectrum

Pim1 binding to the BIM-1 inhibitor was previously validated by thermofluorTM and suggested to bind tightly. For NMR analysis 1 mM BIM-1 was added to Pim1 to ensure complete saturation and this inhibition of Pim1. Figure 3.27 illustrates the ^{1}H, ^{15}N HSQC of the Pim1-BIM-1 complex and reveals that the strategy of forming a single stable complex did not improve the spectrum. Some amide resonances experience CSPs thus indicating an interaction between Pim1 and BIM-1. However no significant increase in the amide peak count and intensity was observed. This implied that the stable Pim1-BIM-1 complex did not significantly perturb the dynamic behaviour for most of the Pim1 backbone amides and did not significantly reduce its suspected tendency to aggregate.

Another strategy that might improve Pim1 ^{1}H, ^{15}N HSQC would be to increase the temperature above 25 °C because temperature can also affect the dynamic properties of a protein. However, due to the low thermal stability of Pim1 even after screening other buffer solution conditions by thermofluor^{TM} this was not possible. Despite showing functional activity Pim1 was thus abandoned as a kinase target by NMR due to the intractability of assignment.
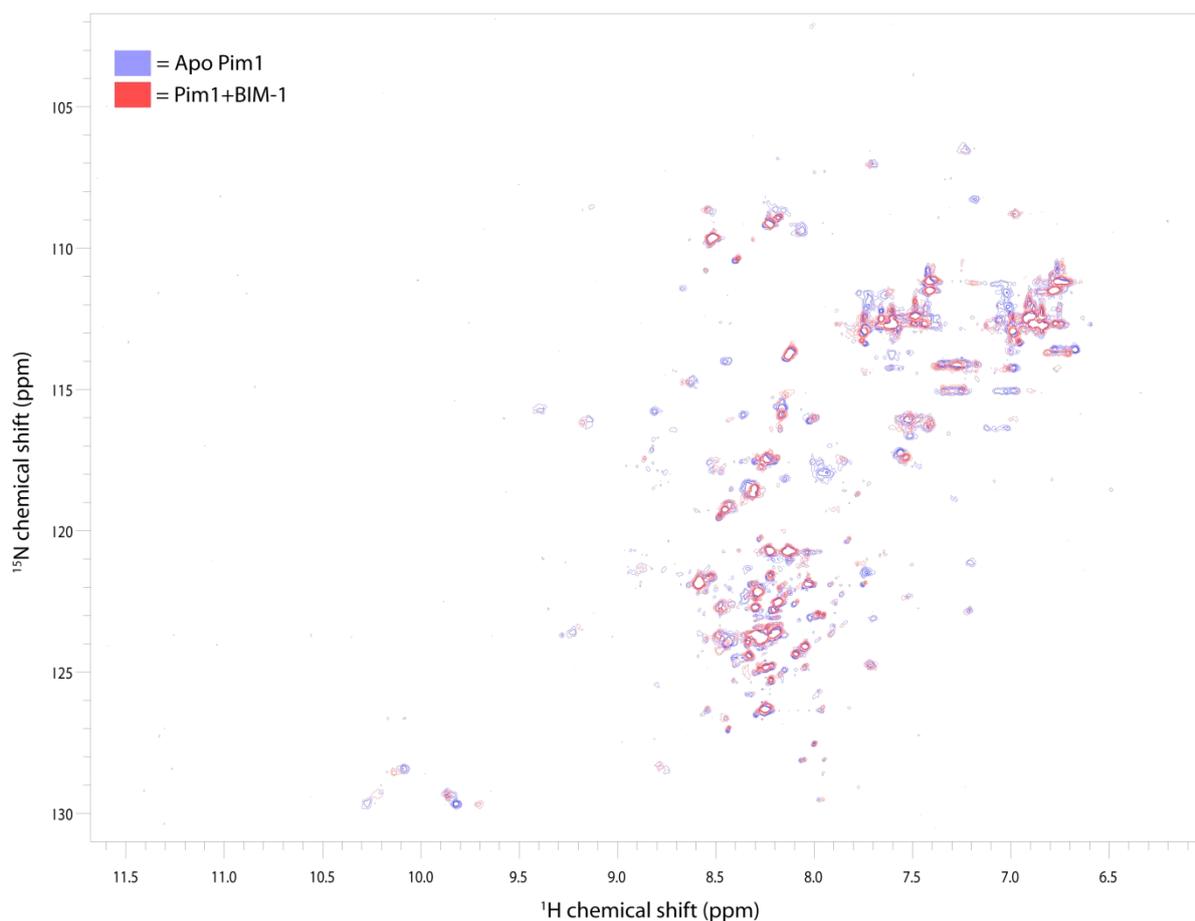
Figure 3.27. apo Pim1 and Pim1-BIM-1 complex HSQC.

Superimposition of the apo Pim1 ^{1}H, ^{15}N HSQC (blue) with the Pim1-BIM-1 complex ^{1}H, ^{15}N HSQC (red) reveals CSPs induced by BIM-1 interacting with Pim1. slow exchange indicative of tight binding after the addition of 1 mM BIM-1. The interaction of BIM-1 with Pim1 however does not increase the peak count. The spectra were acquired with 0.2 mM ^{1}H, ^{15}N Pim1 at 25°C in 50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na azide, 10 % ^{2}H$_2$O on a Varian 800 MHZ. The BIM-1 compound was solubilised in the same buffer conditions.
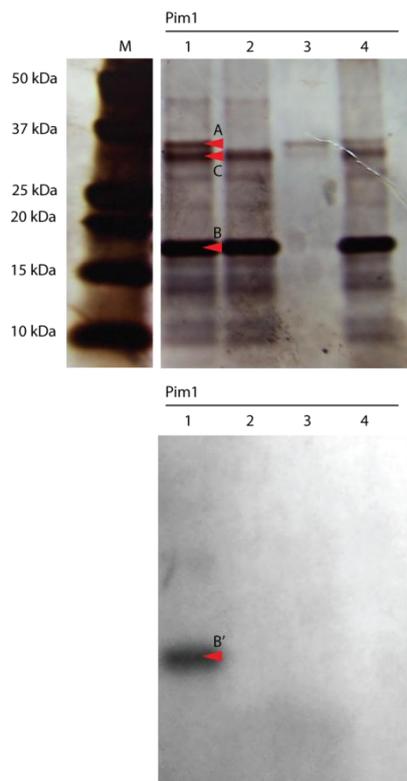
Figure 3.28. Kinase radioactivity assay presented as a silverstain and autoradiograph.

In both the silver stain (upper figure) and autoradiograph (lower figure) lanes 1, 5, 9, 13 correspond to Kinase incubated with substrate Histone H3 in the presence of γ^{32}P-ATP and reaction buffer. Lanes 2, 6, 10, 14 correspond to substrate Histone H3 in the presence of γ^{32}P-ATP and reaction buffer but no kinase. Lanes 3, 7, 11, 15 correspond to kinase incubated with γ^{32}P-ATP and reaction buffer but no substrate Histone H3. Lanes 4, 8, 12, 16 correspond to kinase incubated with substrate Histone H3 and reaction buffer but no γ^{32}P-ATP. Box A highlights the detection of TOPK-Invitrogen autophosphorylation with γ^{32}P-ATP. Box B and D highlights the detection of Histone H3 phosphorylated with γ^{32}P-ATP, therefore confirming active TOPK-Invitrogen and Pim1 respectively. Box C highlights autophosphorylation of TOPK-Invitrogen but in the absence of Histone H3. Our *E.coli* expressed TOPK constructs did not exhibit any activity. In summary TOPK-Invitrogen and Pim1 are active, TOPK and TOPK10-322 are inactive. Red and yellow arrows denote monomeric and dimeric forms of Histone H3 respectively.

TOPK and Pim1 are unsuitable targets for NMR

Despite optimisation of buffer conditions, formation of a stable kinase-inhibitor complex and construct truncation, we evaluate that TOPK and Pim1 are not amenable targets for NMR because both kinases have a large number of backbone amide resonances that broaden out into intermediate exchange. We believe this phenomenon is due to their inherent dynamics which cannot be altered based on our current optimisation strategies. A secondary issue with Pim1 is its poor thermal stability and tendency to precipitat which were emphasised during thermofluor[TM] experiments. Although Pim1 was constitutively active, our TOPK *E.coli* expressed constructs were inactive and was also found to lack binding activity to any ligands despite a high throughput ligand screen conducted by the SGC. This would suggest that the TOPK nucleotide binding site was inaccessible in bacterially expressed form.

Analysis of CaMK1D, CaMK1G, CSNK1G3 and PAK5 as potential NMR targets

Due to the poor ^{1}H, ^{15}N HSQC spectra produced by TOPK and Pim1, backbone assignments could not be accomplished and subsequent solution structure characterisation could not progress further. We therefore continued our analysis of kinases by NMR using different kinases. Several cDNA kinase targets were provided by the SGC cloned into the pNIC28a-Bsa expression vector that included CaMK1D, CaMK1G, CSNK1G3, PAK5, and VRK2. PAK4 was cloned into a pGEX vector. All of these kinase targets were transformed into their default expression host cell lines and expressed and purified as described in chapter II. The kinase targets which yielded soluble amounts of ^{1}H, ^{15}N labelled protein sufficient for a ^{1}H, ^{15}N HSQC were CaMK1D, CaMK1G, CSNK1G3 and PAK5, whereas PAK4 and VRK2 failed to express and were not pursued further. All the subsequent preliminary kinase ^{1}H, ^{15}N HSQC spectra were acquired in 50 mM Na Phosphate pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02% Na Azide as the default NMR buffer and their results are described in the following sections below. These protein kinases were also screened in different buffer conditions using the thermofluor[TM] assay in order to assess their thermal stability in different buffer types and pH for potential buffer solution optimisation.

Evaluation of PAK5 for NMR studies

PAK5 purified to 95% purity and is monomeric by size exclusion folded

Full length PAK5 is an 80 kDa protein encompassing an N-terminal p21 binding domain (PBD), AID and C-terminal kinase catalytic domain. However the PAK5 construct provided by the SGC only encodes the catalytic domain and we refer to this as PAK5 (Figure 3.29, 3.30). PAK5 was expressed and purified to 95% purity based on SDS PAGE (Figure 3.31-3.33). Size exclusion purification of PAK5 (Figure 3.31) suggests that it is monomeric because it elutes as a single peak in a position that correlates to a 34 kDa species based on comparison with the elution of protein standards.

PAK5 is folded and with an estimated 80 % of the total expected backbone amide resonances observable in its ^{1}H, ^{15}N HSQC

A ^{1}H, ^{15}N HSQC spectrum of PAK5 shown in figure 3.34 illustrates that PAK5 is folded due to evidence of amide resonance dispersion. Significantly there are also a large number of backbone amide resonances that are visible. Most of these resonances show strong to moderate signal intensity but a large subset of amide resonances also display very weak intensities thus indicating that the protein conformation is dynamic in solution. In addition to this the weak intensities may also be due to fast T2 relaxation, a phenomenon associated with large proteins. Initial peak count estimates that there are 225 amide resonances with the total number of expected backbone amides equal to 280. Therefore the spectrum illustrates 80 % of the total expected backbone amide peaks. In the bottom left corner of the spectrum where Trp side chain amide Hε resonances reside, only one Trp side chain is readily observable whereas the other two exhibit very weak signal intensity and have folded over into the top left corner of the spectrum (11.5-10.5 ppm ^{1}H dimension, 105-104 ppm ^{15}N dimension). They occur in this position of the spectrum because they were not acquired within the spectral width of the HSQC in the ^{15}N dimension.
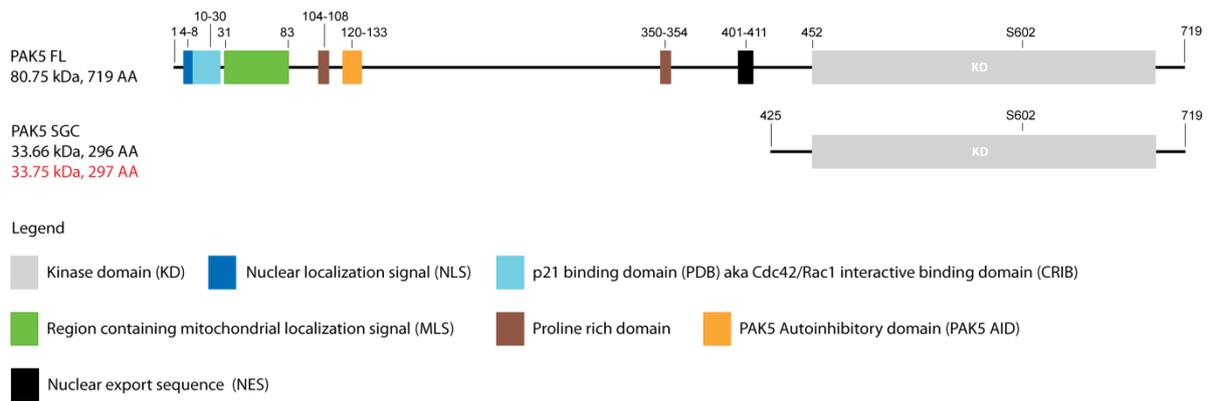
Figure 3.29. PAK5 full length and PAK5 KD construct.

Our PAK5 construct only encodes the kinase catalytic domain from residue 425-719. In PAK5 full length the nuclear localisation signal and proline rich region and C-terminal potential integrin binding domain are shown as red, yellow, and black vertical lines respectively. The green box represents the N-terminal p21-binding domain (PDB) also known as the Cdc42/Rac which contains the CRIB motif. The gray box represents the autoinhibitory domain (AID).

```
            430          440          450          460          470
     SRVSHE QFRAALQLVV SPGDPREYLA NFIKIGEGST GIVCIATEKH
            480          490          500          510          520
TGKQVAVKKM DLRKQQRREL LFNEVVIMRD YHHDNVVDMY SSYLVGDELW
            530          540          550          560          570
VVMEFLEGGA LTDIVTHTRM NEEQIATVCL SVLRALSYLH NQGVIHRDIK
            580          590          600          610          620
SDSILLTSDG RIKLSDFGFC AQVSKEVPKR KSLVGTPYWM APEVISRLPY
            630          640          650          660          670
GTEVDIWSLG IMVIEMIDGE PPYFNEPPLQ AMRRIRDSLP PRVKDLHKVS
            680          690          700          710
SVLRGFLDLM LVREPSQRAT AQELLGHPFL KLAGPPSCIV PLMRQYRHH
```

Predicted pI = 7.7

Figure 3.30. PAK5 primary sequence.

The primary sequence encoded for our PAK5 construct and its predicted pI are shown.
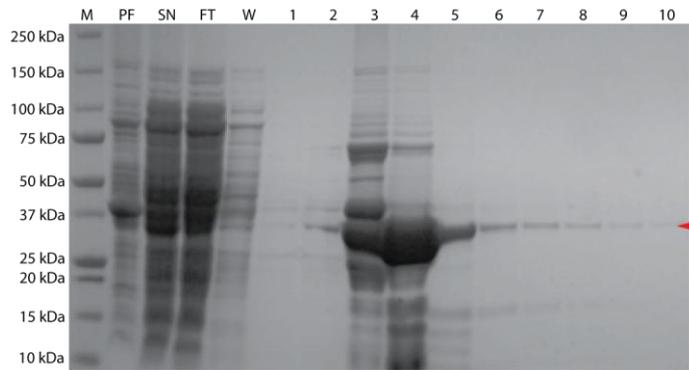
Figure 3.31. SDS PAGE of His$_6$-PAK5 Ni^{2+} NTA purification.

Elution fractions were pooled from fraction number 2-9.
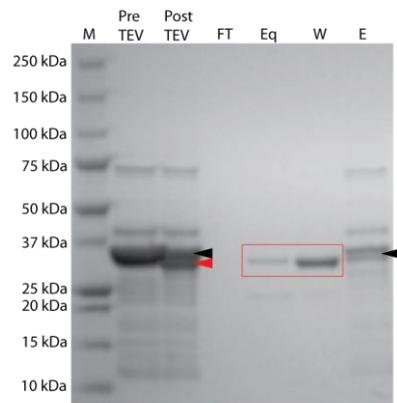
Figure 3.32. SDS PAGE of PAK5 post His$_6$-tag cleavage.

PAK5 in the W and Eq fractions were pooled after overnight incubation with TEV followed by Ni^{2+} NTA purification to purify from incomplete His$_6$-tag cleavage (black arrow), free His$_6$-tag, and other protein contaminants.
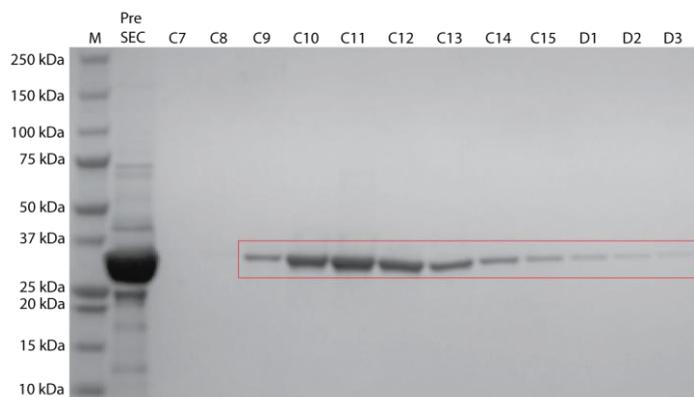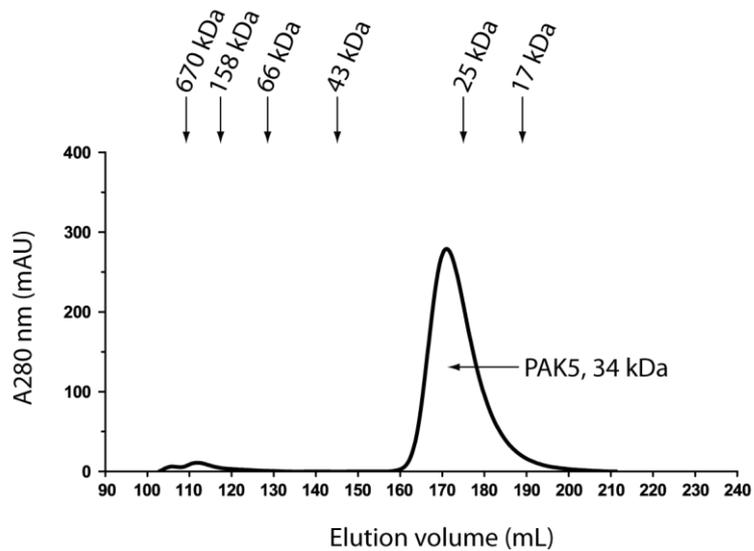
Figure 3.33. PAK5 size exclusion elution profile.

PAK5 is 34 kDa and appears monomeric by size exclusion purification (S75 column) because the position it elutes at (165-185 ml) corresponds to a 34 kDa species. The correlation between molecular weight and position of elution is based on comparison with the elution of protein standards. PAK5 was eluted in 50 mM Na phosphate pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide. Fractions highlighted by the red box were pooled together.
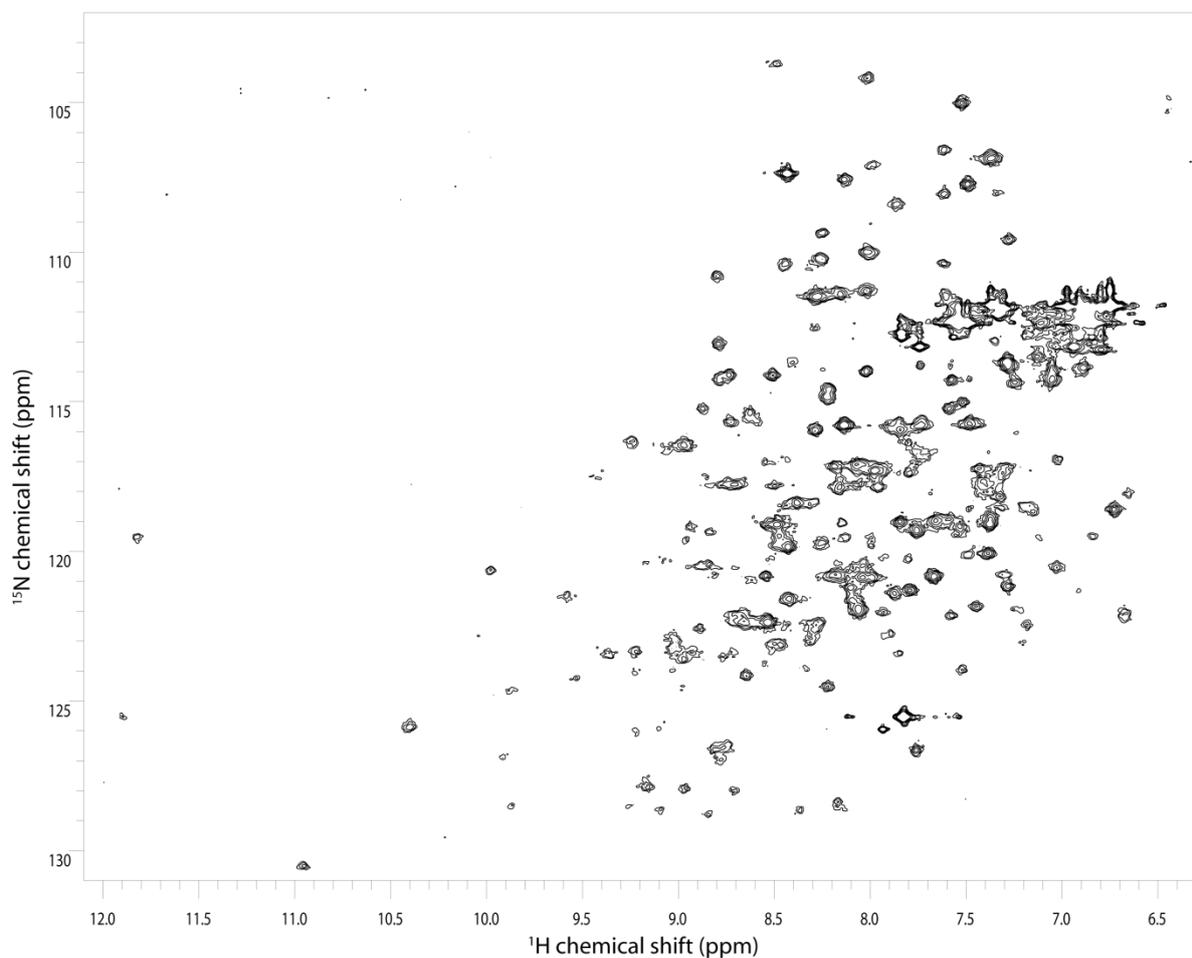
Figure 3.34. PAK5 ^{1}H, ^{15}N HSQC.

Amide resonance dispersion is evident in ^{1}H, ^{15}N HSQC of PAK5 indicating that it is folded. A few of the amide resonances display weak intensity possibly due to dynamics approaching an intermediate exchange timescale or potentially because of fast T2 relaxation due to the large molecular weight of the protein. The spectrum was acquired with 0.2 mM ^{1}H, ^{15}N PAK5 at 25°C in 50 mM Na Phosphate pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na azide, 10 % ^{2}H$_2$O on a Varian 800 MHZ.

PAK5 dynamics suggest a stable conformation

An analysis of PAK5 B-factors was conducted to probe any correlation between the PAK5 B-factors and its ^{1}H, ^{15}N HSQC. Figure 3.35 illustrates the B-factor representation of PAK5 revealing that the core of the C-terminal lobe has predominantly low B-factors indicative of minimal flexibility. Likewise the N-terminal lobe and hinge region show low B-factors. Only a couple of loops and β1 strand display intermediate B-factors (20 Å^2) in the N-terminal lobe. Not surprisingly the most mobile regions of PAK5 are the N-termini and C-termini with both demonstrating high B-factors. Globally PAK5 appears relatively rigid suggesting limited dynamics. This correlates reasonably well with the PAK5 ^{1}H, ^{15}N HSQC (Figure 49) because a large number of backbone amide resonances are observable with mostly strong to moderate intensity suggesting limited dynamics. Therefore it is highly likely these resonances correlate to the rigid portions of PAK5 that show low B-factors. A subset of backbone amide resonances broaden out beyond detection presumably because they exhibit dynamics on an intermediate exchange time scale. We postulate that these resonances are likely to correspond to the mobile loop regions described by intermediate and high B-factors.

Thermofluor^{TM} suggests that PAK5 exhibits moderate-high thermal stability in buffers that range from pH 6-8

An evaluation of PAK5 thermal stability in different buffer conditions was necessary as a strategy for improving the PAK ^{1}H, ^{15}N HSQC amide peak count. Data from thermofluor^{TM} experiments are presented in figure 3.36. It is evident that Tm of PAK5 did not differ significantly in the buffers tested. The range of Tm for PAK5 falls mainly between 55.4-56.2 °C suggesting that it is equally stable in all the tested buffer types exhibiting moderate-high stability and that there is no obvious correlation between Tm and pH. Only in 50 mM Bis-Tris pH 6.0 was the Tm outside of this narrow range with a Tm of 53.4 °C. Thermal melting curves with a Tm that shift by only 1-2 °C are not considered significant. Due to the lack of correlation between the buffer conditions and pH, optimisation of the PAK5 NMR buffer was difficult to predict.
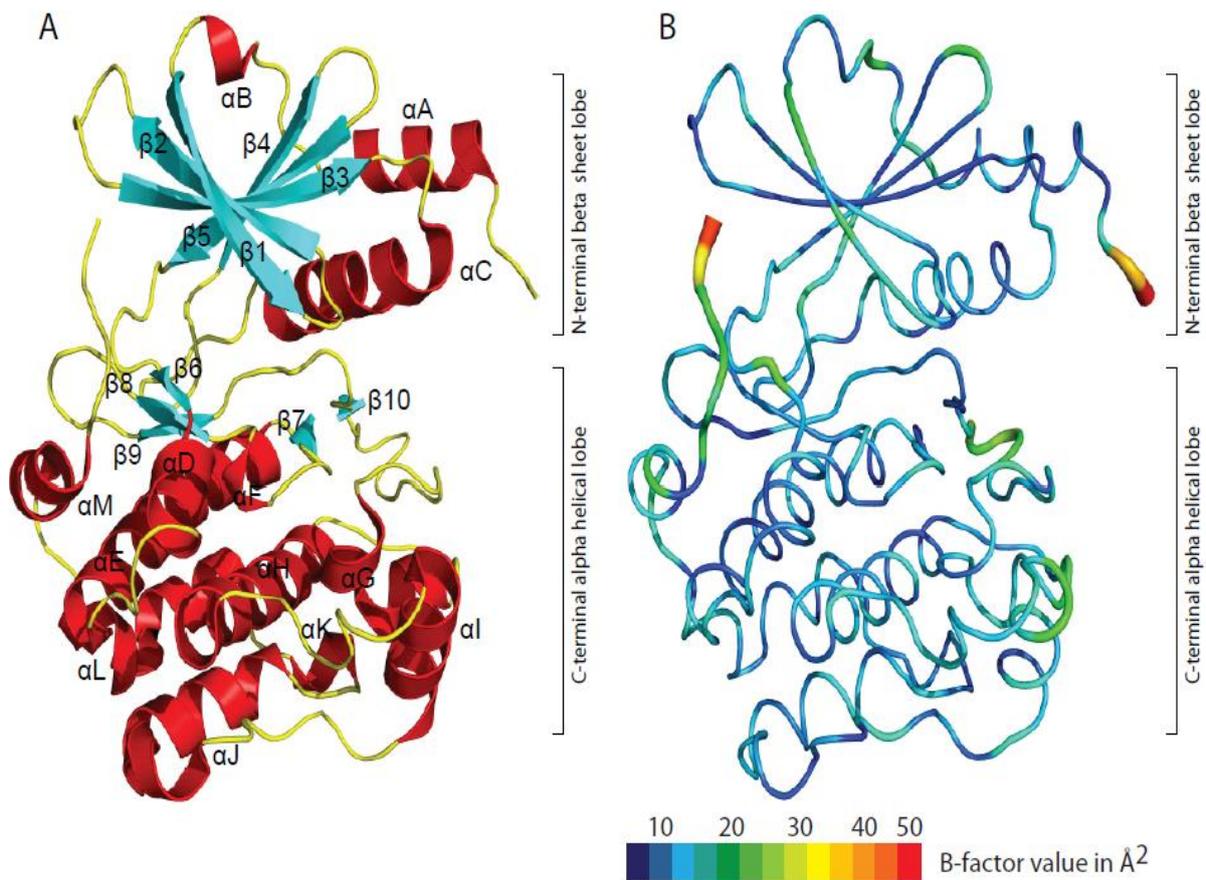
Figure 3.35 apo PAK5 crystal structure and B-factors.

The apo PAK5 crystal structure illustrates that the catalytic domain adopts the typical kinase fold comprised of an N-terminal ATP binding lobe and C-terminal substrate binding lobe. The B-factor representation of PAK5 illustrates the dynamic and rigid regions of the protein kinase. The N-terminal and C-terminal regions display the highest B-factor values thus corresponding to the most mobile parts of PAK5. The core regions of both lobes appear rigid as they are described by low B-factors.
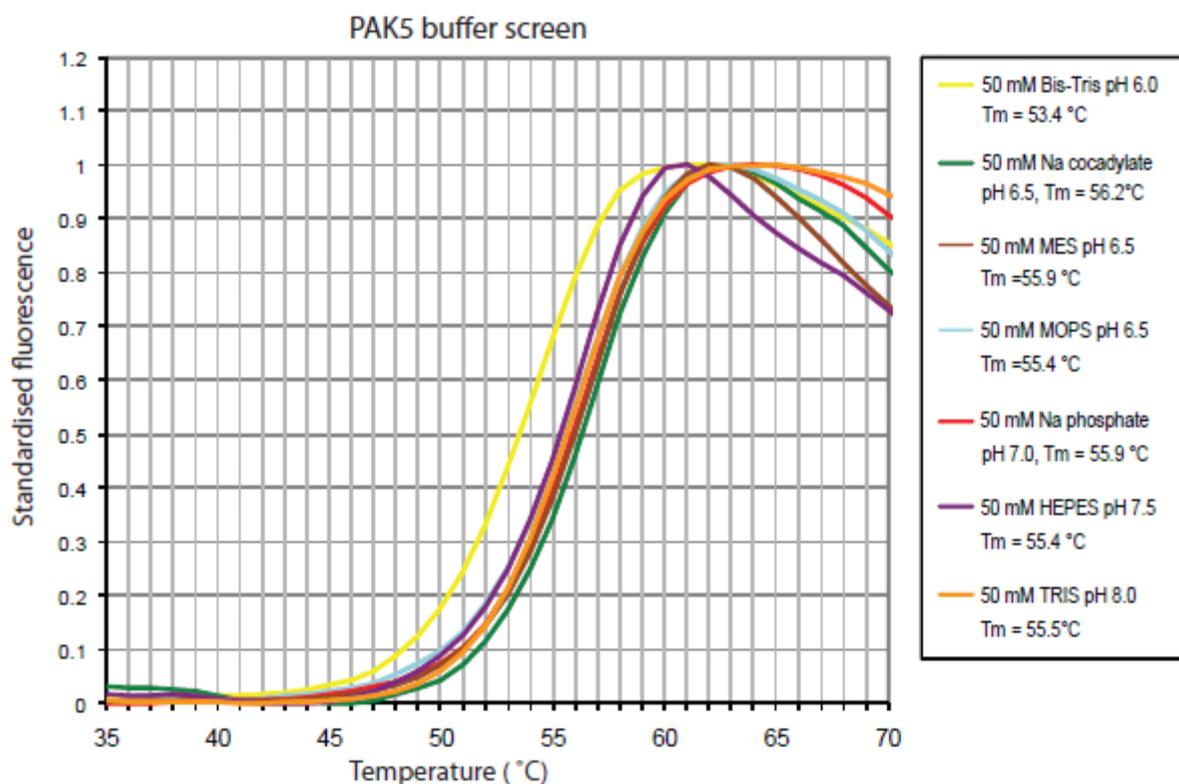
Figure 3.36. PAK 5 thermofluor pH screen.

Evaluation of PAK5 thermal stability in different buffer types and pH was conducted by thermofluor[TM]. With the exception of 50 mM Bis-Tris pH 6.0 which had a Tm of 53.4 °C, the Tm of PAK5 is consistently similar in the all buffer types and pH showing a narrow Tm range of 55.4-56.2 °C. PAK5 has a predicted pI of 7.7 but according to thermofluor[TM], the pI must be higher than pH 8.0 or lower than pH 6.0 because a Tm around 55 °C suggests a stable protein far from its pI assuming the results are not false positives.

PAK5 ^{1}H, ^{15}N HSQC amide peak count did not increase at pH 7.0

Although the PAK5 ^{1}H, ^{15}N HSQC showed 80 % of the total expected backbone amide peaks further optimisation was necessary to increase this in order to achieve a more complete backbone assignment. Furthermore a large subset of peaks displayed weak and moderate intensity which is not ideal for backbone assignment. Optimisation was attempted by lowering the pH from 7.5 to 7.0 assuming no compromise in stability based on the thermofluor[TM] results. The PAK5 ^{1}H, ^{15}N HSQC spectrum at pH 7.0 (Figure 3.37) reveals that the signal intensity of some amide resonances does increase but was not significant enough for additional amide resonances to be observed.

Unfortunately at pH 7.0 there was an increase in PAK5 precipitation compared to pH 7.5 during experimental acquisition of the ^{1}H, ^{15}N HSQC at 25 °C. This implies that the thermal stability of PAK5 is lower than suggested by thermofluor[TM]. A ^{1}H, ^{15}N TROSY-HSQC of PAK5 (Figure 3.38) illustrates that there is a slight improvement over the ^{1}H, ^{15}N HSQC with better line shapes and resolution due to the TROSY effect. Further optimisation would require deuteration which might address the lower than expected total peak count if amide resonance broadening was due to fast T2 relaxation associated with its large molecular weight. Resonance intensities should also improve with deuteration however further optimisation was not attempted because the ^{1}H, ^{15}N HSQC demonstrated by CaMK1D was superior and required less optimisation.
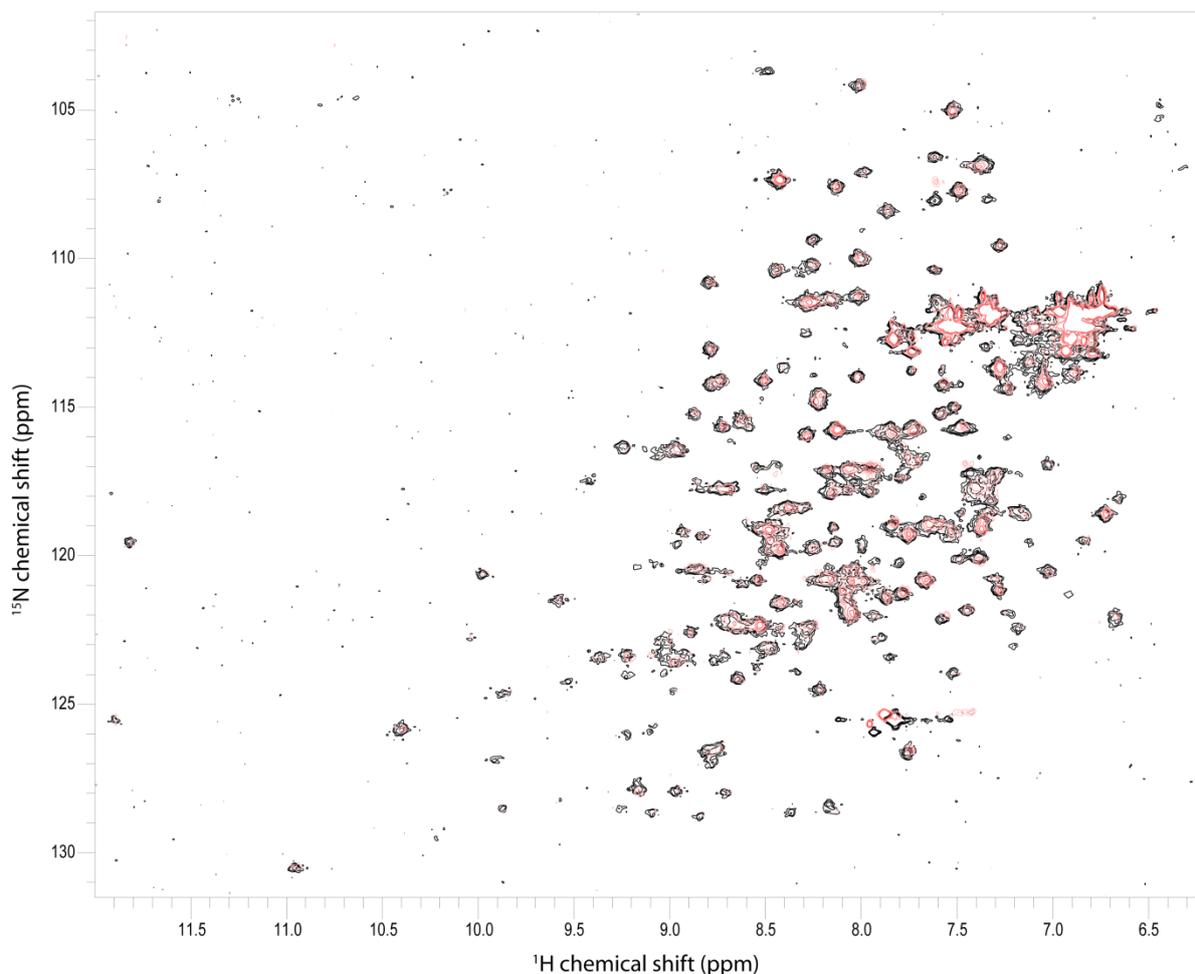
Figure 3.37. PAK5 ^{1}H, ^{15}N HSQC at pH 7.0 and pH 7.5.

An overlay of the PAK5 ^{1}H, ^{15}N, HSQC at pH 7.5 (black) and at 7.0 (red) illustrates that the spectra are nearly identical. No significant improvements in terms of additional peaks were observed at pH 7.0 compared to pH 7.5 but an increase in precipitation resulted at pH 7.0. Buffer solution conditions were the same as in previous PAK5 spectra.
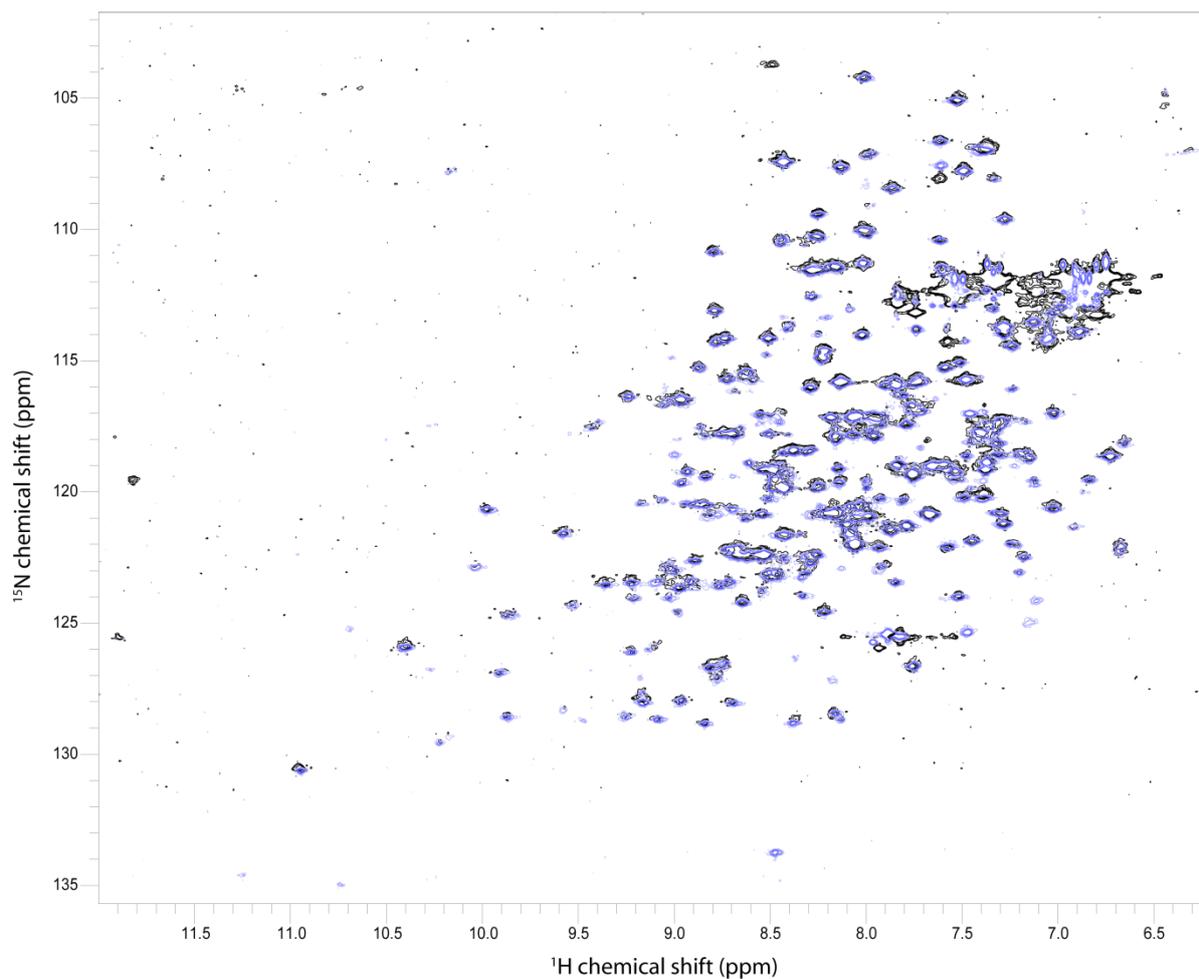
Figure 3.38. PAK5 ^{1}H, ^{15}N HSQC and TROSY HSQC at pH 7.0.

Superimposition of the PAK5 ^{1}H, ^{15}N HSQC (black) in its default buffer at pH 7.0 with the PAK5 ^{1}H, ^{15}N TROSY-HSQC (blue) in the same buffer at pH 7.0 illustrates an improvement in line shape due to the TROSY effect. Notably the signal intensity for a few peaks in the TROSY spectrum is reduced and is also due to the TROSY effect.

Evaluation CaMK1D for NMR studies

CaMK1D purified to 95 % purity and is monomeric by size exclusion

Initial expression and purification of ^{1}H, ^{15}N CaMK1D revealed yields of about 10 mg L^{-1} of culture (M9 minimal media) could be obtained and purified to 95 % purity based on SDS PAGE whilst size exclusion purifcation suggests that it purifies as a monomeric 37.8 kDa species (Figure 3.41-3.43). Furthermore CaMK1D demonstrated activity in our NADH coupled assay when in the presence of Ca^{2+}/CaM indicating that CaMK1D was folded and functional.

CaMK1D is folded with an estimated 85 % of the total expected backbone amide resonances observable in its ^{1}H, ^{15}N HSQC

Autoinhibited CaMK1D is comprised of 334 residues (Figure 3.39, 3.40) and the total number of expected amide resonances is 322. Our preliminary ^{1}H, ^{15}N HSQC of CaMK1D (Figure 3.44) illustrates a spectrum with a large number of observable amide resonances. An estimate of the peak count suggests 275 backbone amide peaks were visible correlating to an 85 % of the total expected backbone amide peaks. Apart from the centre of the spectrum where there is considerable resonance overlap the rest of the resonances are well dispersed indicating that the construct is folded. Strong signal intensity is associated with the majority of the amide resonances implying a stable conformation but some amide resonances with intermediate and weak intensities are also apparent suggesting that CaMK1D contains dynamic elements. CaMK1D contains five Trp residues and four Trp side chain amide Hε resonances are initially observable in the expected bottom left area of the spectrum. The remaining unobserved Trp side chain amide Hε resonance was only visible at low contour level because it displayed very weak signal intensity. It is possible that this particular Trp side chain amide Hε resonance was experiencing dynamic exchange that was approaching an intermediate exchange timescale.
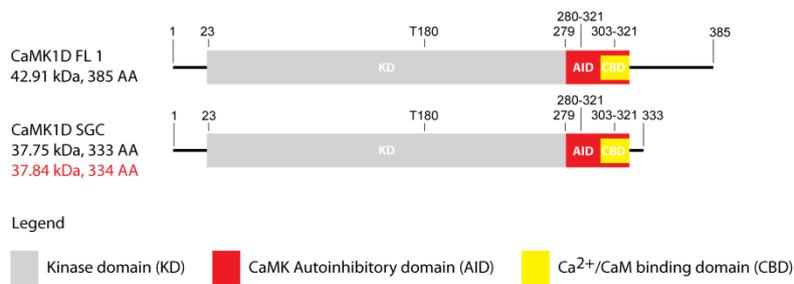
Figure 3.39. CaMK1D full length and CaMK1D SGC construct.

The SGC CaMK1D constructs encode the kinase catalytic domain and the regulatory AID including the CBD. The C-terminal stretch of residues of unknown function is truncated and lacks the remaining 52 residues.

```
1          10         20         30         40         50

MARENGESSS SWKKQAEDIK KIFEFKETLG TGAFSEVVLA EEKATGKLFA

          60         70         80         90         100

VKCIPKKALK GKESSIENEI AVLRKIKHEN IVALEDIYES PNHLYLVMQL

         110        120        130        140        150

VSGGELFDRI VEKGFYTEKD ASTLIRQVLD AVYYLHRMGI VHRDLKPENL

         160        170        180        190        200

LYYSQDEESK IMISDFGLSK MEGKGDVMST ACGTPGYVAP EVLAQKPYSK

         210        220        230        240        250

AVDCWSIGVI AYILLCGYPP FYDENDSKLF EQILKAEYEF DSPYWDDISD

         260        270        280        290        300

SAKDFIRNLM EKDPNKRYTC EQAARHPWIA GDTALNKNIH ESVSAQIRKN

         310        320        330

FAKSKWRQAF NATAVVRHMR KLHLGSSLDS SNA


Predicted pI = 6.1
```

Figure 3.40 CaMK1D primary sequence.

The primary sequence encoded for our CaMK1D construct and its predicted pI are shown. Thr 198 (highlighted in red) is phosphorylated by CaMKK for full activation.
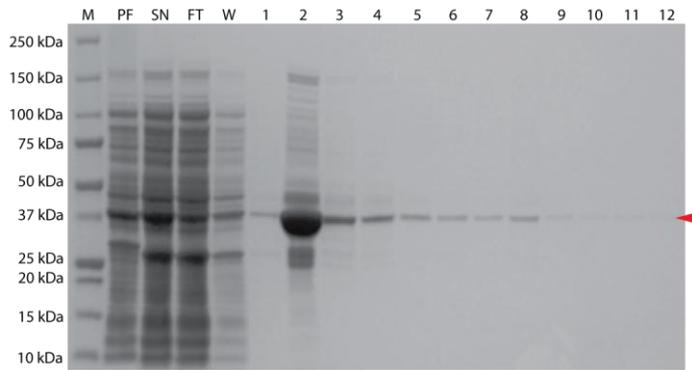
Figure 3.41. SDS PAGE of His$_6$-CaMK1D Ni^{2+} NTA purification.

Elution fractions were pooled from fraction number 1-12.
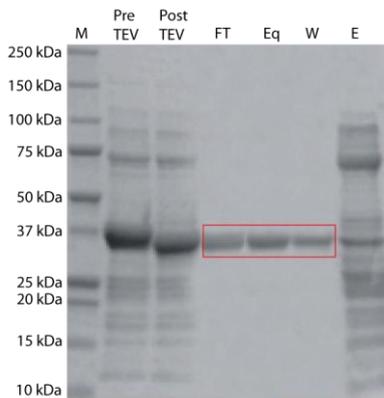
Figure 3.42. SDS PAGE of CaMK1D post His$_6$-tag cleavage.

CaMK1D in the FT, Eq, and W fractions were pooled after overnight incubation with TEV followed by Ni^{2+} NTA purification to purify from incomplete His$_6$-tag cleavage, free His$_6$-tag, and other protein contaminants.
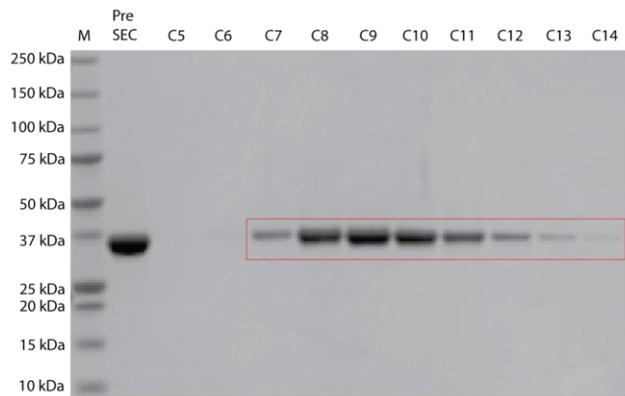
Figure 3.43. SDS PAGE of CaMK1D from size exclusion purification.

CaMK1D is 37.8 kDa protein that elutes from about 150–180 ml. SDS PAGE of these fractions indicates 95 % purity. A higher molecular weight contaminant elutes from 130-140ml. CaMK1D was eluted in 50 mM Na phosphate pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide.
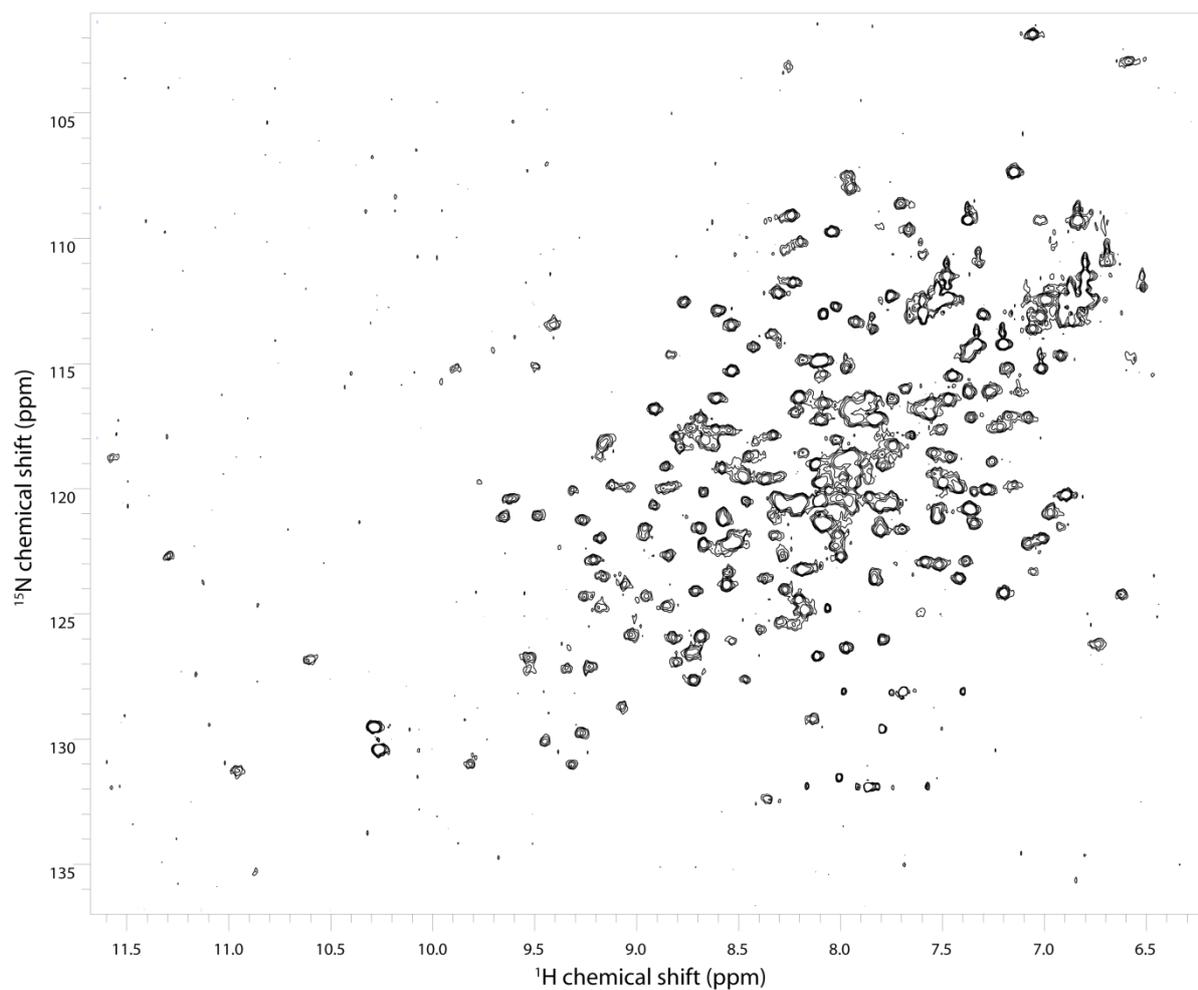
Figure 3.44. CaMK1D ^{1}H, ^{15}N HSQC.

The ^{1}H, ^{15}N HSQC of CaMK1D demonstrates amide resonance dispersion indicative of a folded protein with a large majority of the resonances displaying strong intensity. The spectrum was acquired with 0.2 mM ^{1}H, ^{15}N CaMK1D at 25°C in 50mM Na Phosphate pH 7.5, 150mM NaCl, 0.5mM TCEP, 0.02% Na Azide on a Varian 800 MHZ.

CaMK1D dynamics suggests a stable conformation

We previously inferred from the ^{1}H, ^{15}N HSQC of CaMK1D that the protein was in a relatively stable conformation and attempted to examine the CaMK1D B-factors to elucidate any potential correlation with the NMR spectrum. Figure 3.45 illustrates the B-factor representation of CaMK1D and highlights the dynamic and flexible regions of the protein. It is evident that both the N-terminal and C-terminal lobes exhibit high and intermediate levels of flexibility. One of the regions showing the highest B-factors corresponds with the αC helix. This is not surprising because the αC helix is a mobile regulatory element in protein kinases and usually adopts a disordered or inactive conformation in kinases that are in an inactive state (Huse and Kuriyan 2002). Another region of high flexibility appears near the base of the C-terminal lobe and corresponds to helix αH and the adjacent loop N-terminal to it.

Compared to PAK5, the global structure of CaMK1D appears to be more dynamic and flexible in terms of its B-factors even in the presence of the GSK-3 XIII inhibitor. This however does not correlate with a poor ^{1}H, ^{15}N HSQC like Pim1, conversely the ^{1}H, ^{15}N HSQC of apo CaMK1D is superior to PAK5 interms of a higher peak count and consistency of peak intensities. This suggests that in solution the CaMK1D backbone amide resonance intensities do not correlate with its B-factors. However this could suggest that the dynamics of the amide groups is manifested on a slow or fast exchange time scale as opposed to intermediate exchange broadening. A majority of the resonances are likely to be in a single stable conformation because most of the core regions of the N-terminal and C-terminal lobe are defined by low B-factors. We postulate that stabilising role of the AID has a major influence on the dynamics of the resonances and is likely to reduce the number of conformations that CaMK1D can sample due to its autoinhibitory role as a pseudosubstrate. Furthermore low B-factors describe the AID which reinforces the notion that it is a rigid element. It is important to note that the B-factors reflect CaMK1D in complex with GSK-3 XIII whereas our ^{1}H, ^{15}N HSQC of CaMK1D is for the apo form therefore an accurate correlation between the NMR spectrum and B-factors is not possible.
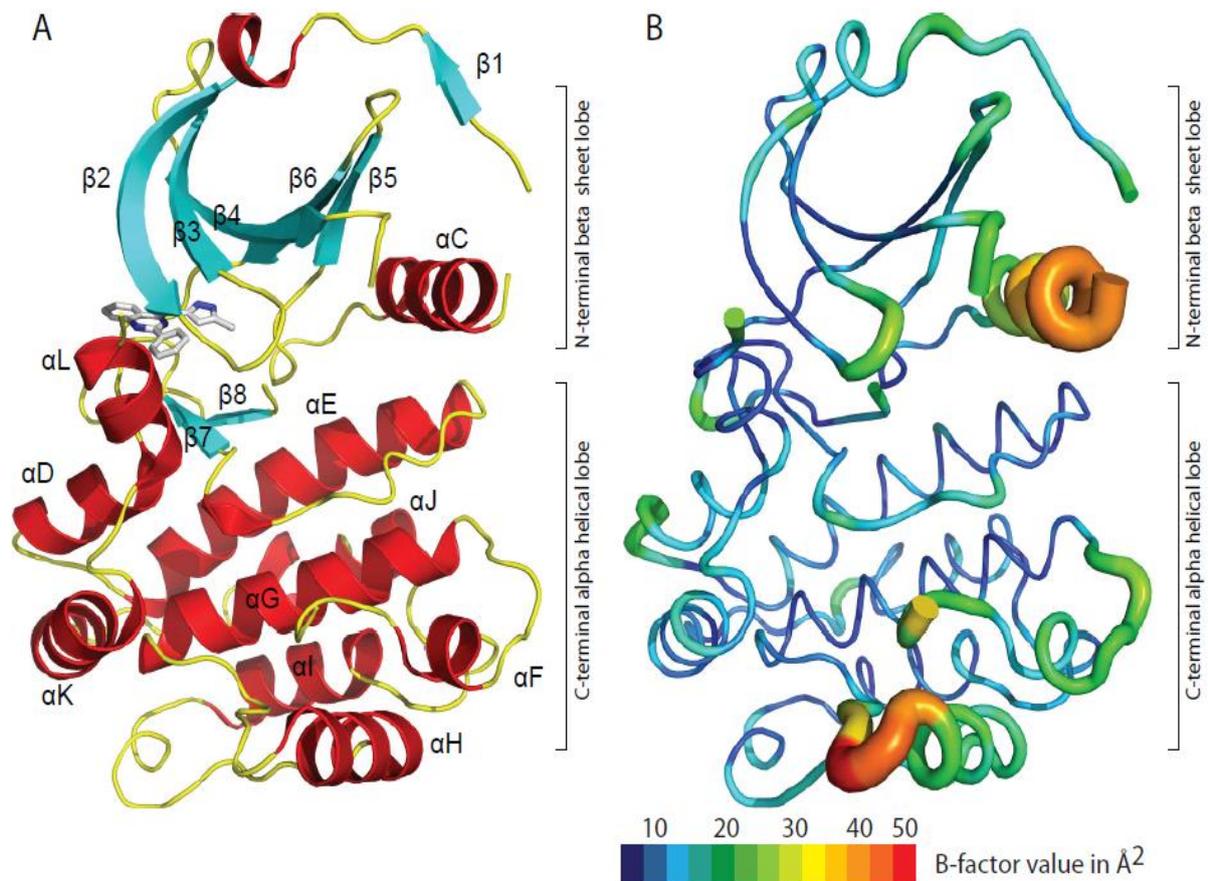
Figure 3.45 CaMK1D-GSK3 XIII inhibitor crystal structure and B-factors.

The crystal structure of CaMK1D in complex with GSK-3 XIII inhibitor (PDB:2JC6) exhibits a canonical kinase fold comprised of an N-terminal ATP binding lobe and C-terminal substrate binding lobe that is followed by a regulatory autoinhibitory domain. The activation loop, flexible N-termini and C-termini are not observable in this crystal structure due to limited electron density. The B-factor representation of CaMK1D illustrates the dynamic and rigid regions of the protein kinase. Notably the majority of the AID is very rigid and indicative of a stable pseudosubstrate interaction.

Thermofluor^{TM} suggests that CaMK1D exhibits moderate-high thermal stability in buffers that range from pH 6-8 and correlates positively with increasing pH

The thermal stability of CaMK1D was assessed in different buffer conditions using thermofluor^{TM} to screen for conditions that might optimise its NMR spectra. Figure 3.46 shows that the Tm of CaMK1D generally increases with high pH. For example the Tm of CaMK1D increases from 48.1-57.1 °C as the pH is shifted from pH 6.0 to pH 8.0. Presumably the rational for CaMK1D stability correlating with increasing pH is because at high pH we are further away from the predicted CaMK1D pI of 6.1 i.e. there is a greater overall surface net charge on CaMK1D thus increasing its solubility and stability. Moreover the range of Tm values suggests that CaMK1D has moderate-high stability. Notably the Tm of CaMK1D in 50 mM Na phosphate pH 7.0 and 50 mM HEPES pH 7.5 was 55 and 56 °C respectively suggesting that these conditions can potentially produce moderate-high stability at pH that is appropriate for NMR. Therefore Na phosphate pH 7.0 and HEPES pH 7.5 were used to generate our default CaMK1D NMR buffer which contained 50 mM Na phosphate pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02% Na Azide.
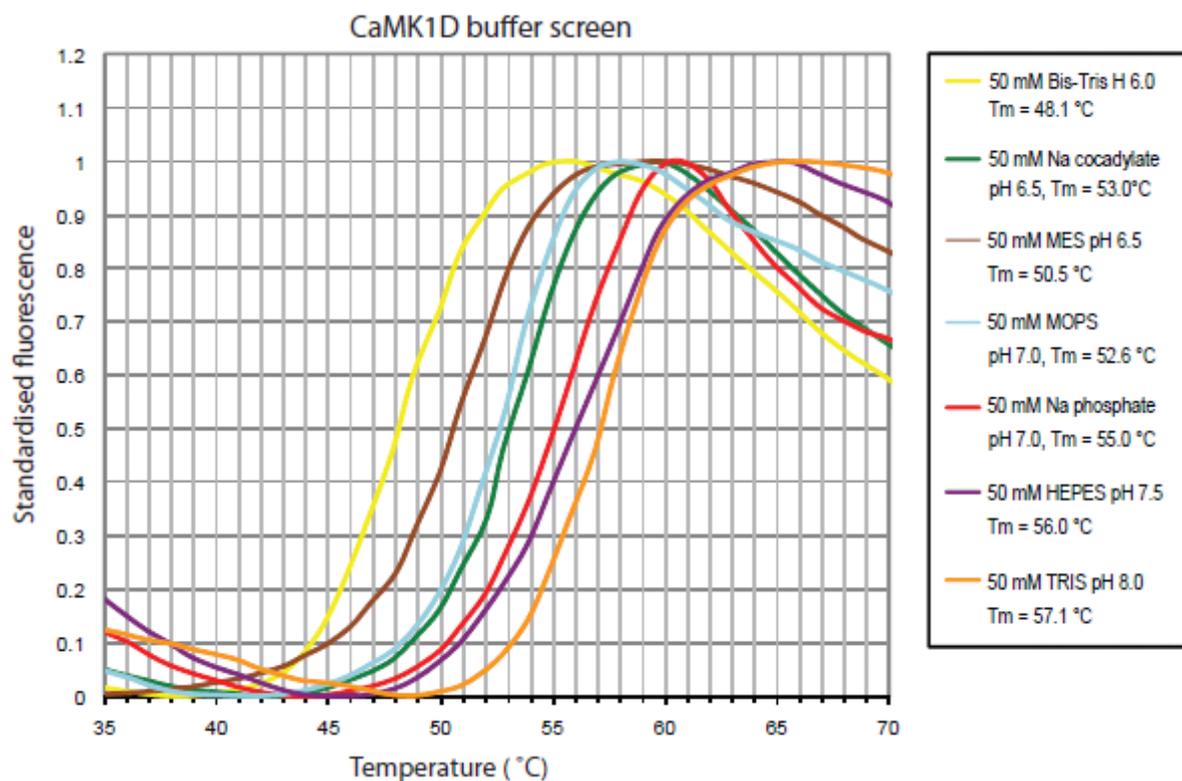
Figure 3.46. CaMK1D thermofluor pH screen.

Using thermofluor[TM] the average CaMK1D Tm was determined in different buffer solutions. CaMK1D stability appears to correlate positively with increasing pH from 6.0-80. Consequently CaMK1D appears to be most stable in 50 mM TRIS pH 8.0 with a Tm of 57.1 °C.

CaMK1D ^{1}H, ^{15}N HSQC amide peak count did not increase at pH 7.0 but did increase the amide peak intensities

Our initial ^{1}H, ^{15}N HSQC spectrum of CaMK1D was very promising and was therefore selected as the main kinase target for study by NMR. Additional optimisation was made by lowering the pH from 7.5 to 7.0 to increase the signal intensity of amide resonances and validate if additional resonances could be observable. In addition the NaCl concentration was also lowered to 75 mM to optimise the sensitivity of the cryoprobe. These buffer changes were made on the basis of thermofluor^{TM} data which suggested that CaMK1D had an equivalent Tm in pH 7.0 and pH 7.5 and similarly for NaCl at 75 mM and 150 mM.

Lowering the pH to 7.0 resulted in a ^{1}H, ^{15}N HSQC (Figure 3.47) that demonstrated resonances with greater signal intensity but additional resonances were not apparent. A few resonances exhibit pH dependent chemical shift perturbations. Notably a resonance in the top right corner of the spectrum (6.6 ppm ^{1}H dimension, 103 ppm ^{15}N dimension). Subsequently a ^{1}H, ^{15}N TROSY-HSQC at pH 7.0 was acquired for further optimisation and reveals improved line shape and resolution. An overlay of this spectrum with the ^{1}H, ^{15}N HSQC spectrum at pH 7.0 shown in figure 3.48 highlights these differences
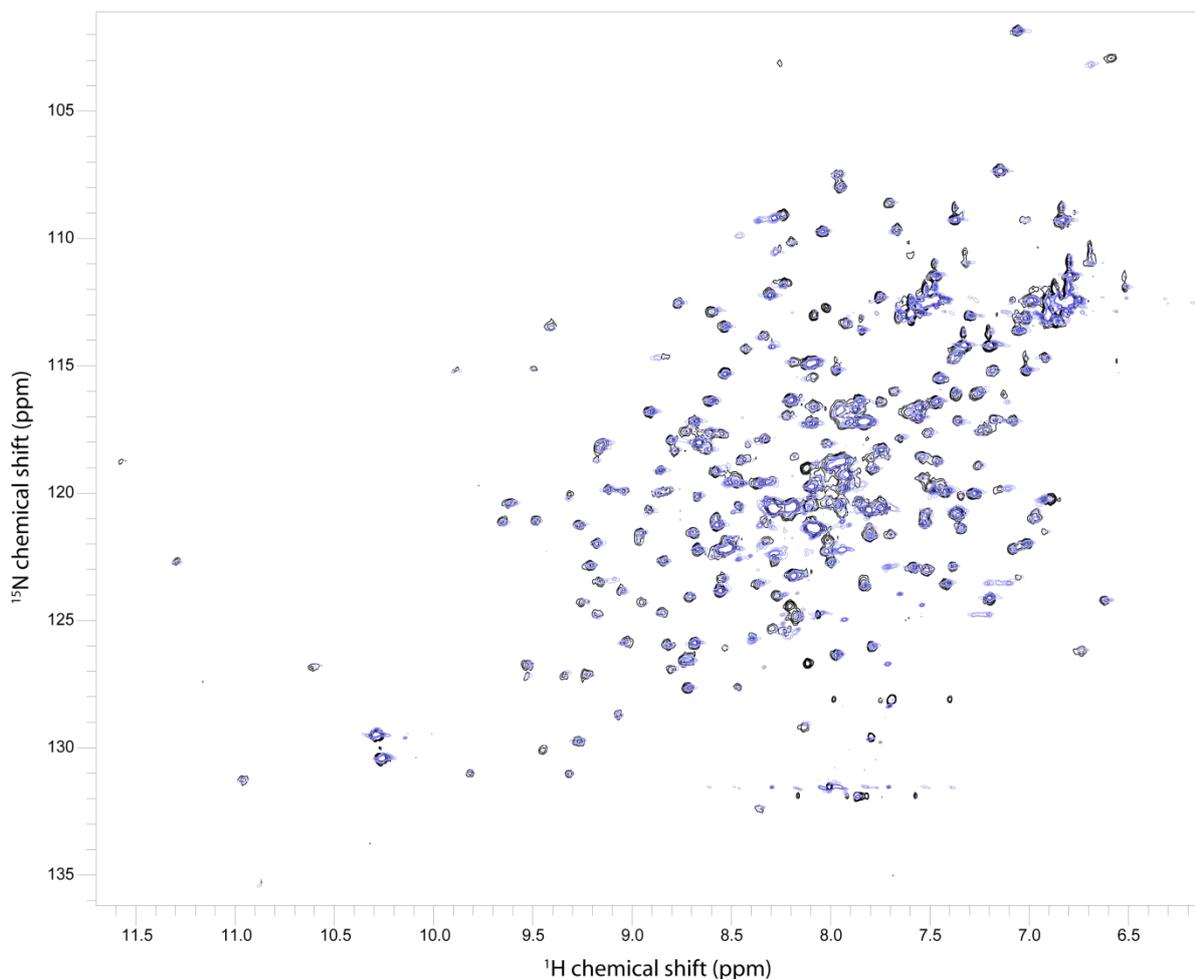
Figure 3.47. CaMK1D ^{1}H ^{15}N HSQC, pH 7.0, 75 mM NaCl.

In order to increase the peak count and optimise the sensitivity of the cryoprobe the default CaMK1D NMR buffer solution pH and NaCl was lowered from pH 7.5, 150 mM NaCl (black) to pH 7.0, 75 mM NaCl (blue). The superimposed spectra indicates that at pH 7.0 and 75 mM NaCl the resonance intensities are greater due to increased protonation of the exchangeable amide groups and sensitivity enhancement of the cryoprobe however additional peaks were not evident. Some peaks at pH 7.0, 75 mM NaCl (blue) also did not superimpose with the default conditions (black) suggesting that these amide resonances where exhibiting pH dependent CSPs.
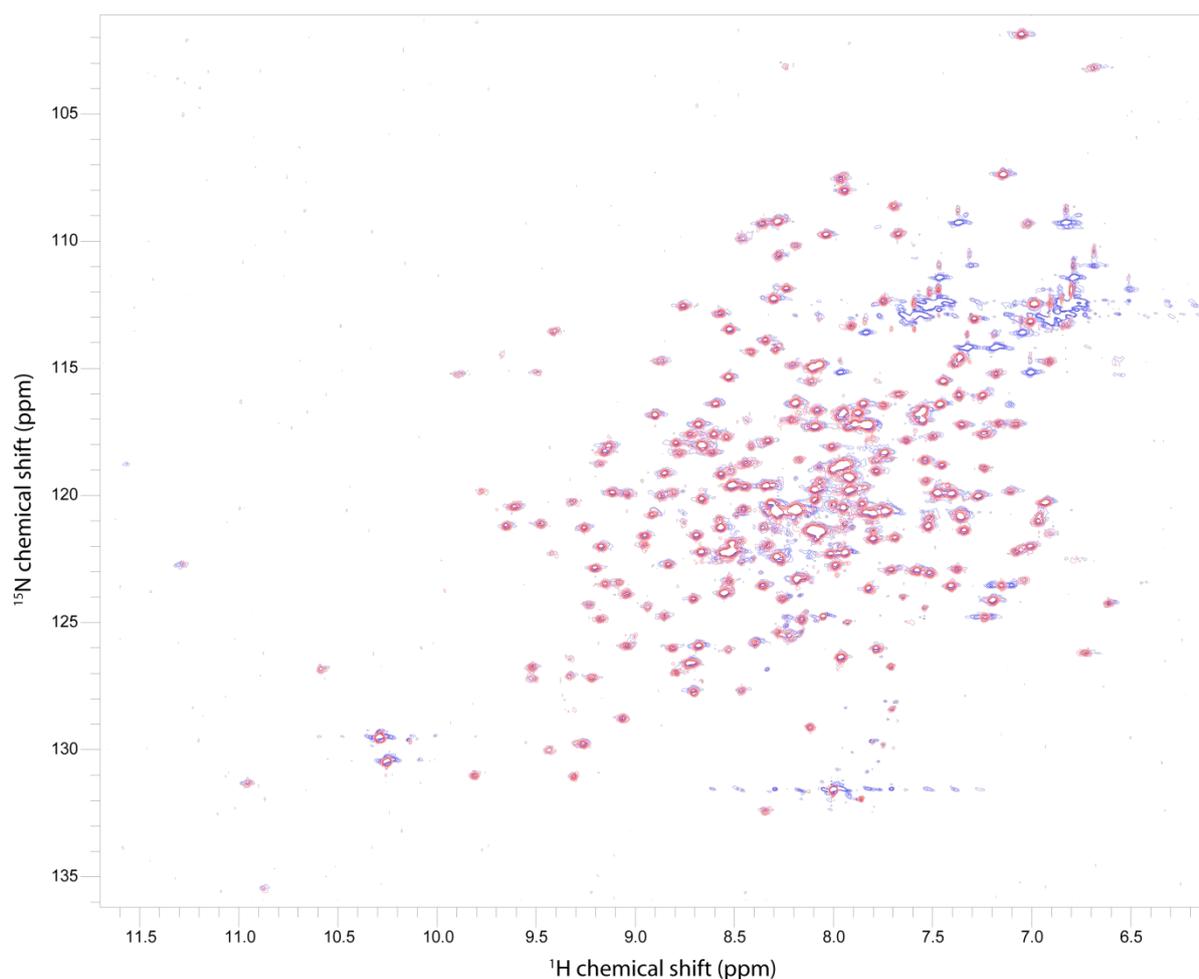
Figure 3.48. CaMK1D ^{1}H ^{15}N HSQC and TROSY HSQC, pH 7.0, 75 mM NaCl.

Under the optimised buffer solution conditions using pH 7.0 and 75 mM NaCl, the ^{1}H, ^{15}N TROSY-HSQC of CaMK1D (red) was acquired. This spectrum superimposed with the ^{1}H ^{15}N HSQC of CaMK1D (blue) acquired in the same optimised buffer solution demonstrates improved line shape and resolution of backbone amide resonances whilst resonance intensities of the Asn and Gln amides broaden out.

Evaluation of CaMK1G for NMR studies

CaMK1G is folded but the majority of expected backbone amide resonances are unobservable in its ^{1}H, ^{15}N HSQC

CaMK1G (Figure 3.49, 3.50) was purified to 90 % purity according to SDS PAGE and appeared to be monomeric by size exclusion (Figure 3.51-3.53). In contrast to CaMK1D, the ^{1}H, ^{15}N HSQC of CaMK1G (Figure 3.54) is inferior in all aspects. Although CaMK1G the spectrum suggests that CaMK1G is folded due to the presence of amide resonance dispersion and two dispersed Trp side chain amide Hε resonances there are a large number of expected backbone amide resonances missing. Theses observations are similar to Pim1 suggesting intermediate exchange dynamics and aggregation. The intensity of the few observable resonances ranges from moderate to very weak implying dynamic behaviour approaching intermediate exchange. Precipitation of the CaMK1G sample was evident after acquiring the ^{1}H, ^{15}N HSQC spectrum at 25 °C which was likely to have reduced the intensity of the few observable resonances. It is somewhat unexpected that the CaMK1G spectrum is significantly inferior to CaMK1D considering that they high sequence identity. This suggests that the subtle differences in sequence between CaMK1D and CaMK1G are enough to significantly alter their intrinsic dynamic behaviour and aggregation properties.

Interestingly we note that the CaMK1G construct is slightly truncated at the C-terminal end of the CBD but hypothetically this should not affect the stability of the protein because the majority of the pseudosubstrate interactions are mediated by the AID which is N-terminal to the CBD but the ^{1}H, ^{15}N HSQC and low thermal stability of CaMK1G (Figure) suggest this is not the case. Instead it would appear that slight truncation of the CBD has destabilised the protein significantly which would imply that the CBD is implicated in key pseudosubstrate interactions. If this hypothesis is correct then the next likely strategy would be construct optimisation to improve the stability of CaMK1G and potentially it's ^{1}H, ^{15}N HSQC spectrum either by truncating the regulatory domain further to remove the destabilizing effect or by extending it beyond the CBD to increase stabilizing pseudosubstrate interactions. Moreover

adjacent to the CBD is the C-terminal extension of CaMK1G which is functionally important in the targeting of CaMK1G to membranes. We postulate that some of these residues near the CBD may also form further stabilizing interactions across the N-terminal lobe but residues further upstream from this point are likely to be disordered, dynamic and flexible in the absence of a membrane surface.
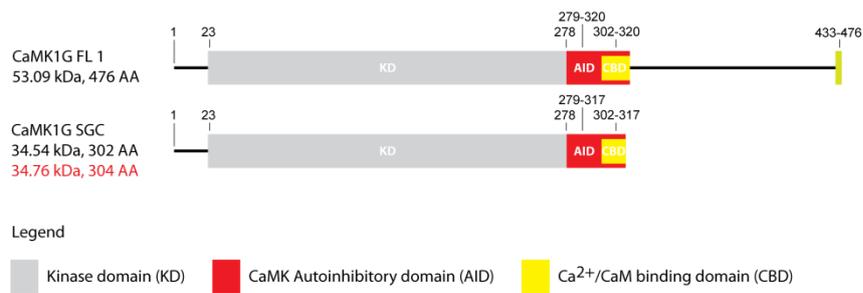
Figure 3.49. CaMK1G full length and CaMK1G SGC construct.

Our CaMK1G construct encodes the kinase catalytic domain and the regulatory AID which includes the overlapping CBD. The extended C-terminal stretch of residues and a small region of the CBD are both truncated.

```
          20          30          40          50          60

        QTTNIR KTFIFMEVLG SGAFSEVFLV KQRLTGKLFA LKCIKKSPAF

          70          80          90         100         110

RDSSLENEIA VLKKIKHENI VTLEDIYEST THYYLVMQLV SGGELFDRIL

         110         120         130         140         150

ERGVYTEKDA SLVIQQVLSA VKYLHENGIV HRDLKPENLL YLTPEENSKI

         160         170         180         190         200

MITDFGLSKM EQNGIMSTAC GTPGYVAPEV LAQKPYSKAV DCWSIGVITY

         210         220         230         240         250

ILLCGYPPFY EETESKLFEK IKEGYYEFES PFWDDISESA KDFICHLLEK

         260         270         280         290         300

DPNERYTCEK ALSHPWIDGN TALHRDIYPS VSLQIQKNFA KSKWRQAFNA


AAVVHH


Predicted pI = 5.9
```

Figure 3.50. CaMK1G primary sequence.

The primary sequence encoded for our CaMK1G construct and its predicted pI are shown.
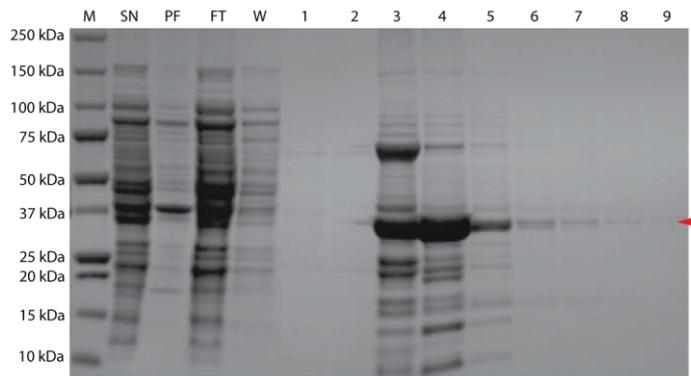
Figure 3.51. SDS PAGE of His$_6$-CaMK1G Ni^{2+} NTA purification.

Elution fractions were pooled from fraction number 3-9.
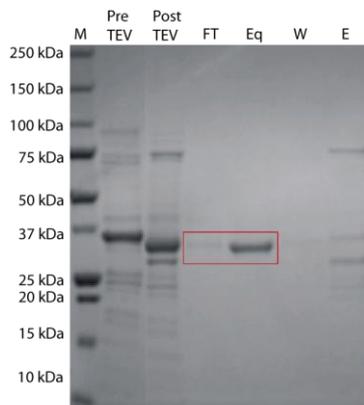
Figure 3.52. SDS PAGE of CaMK1G post His$_6$-tag cleavage.

CaMK1G in the FT and Eq fractions were pooled after overnight incubation with TEV followed by Ni^{2+} NTA purification to purify from incomplete His$_6$-tag cleavage, free His$_6$-tag, and other protein contaminants.
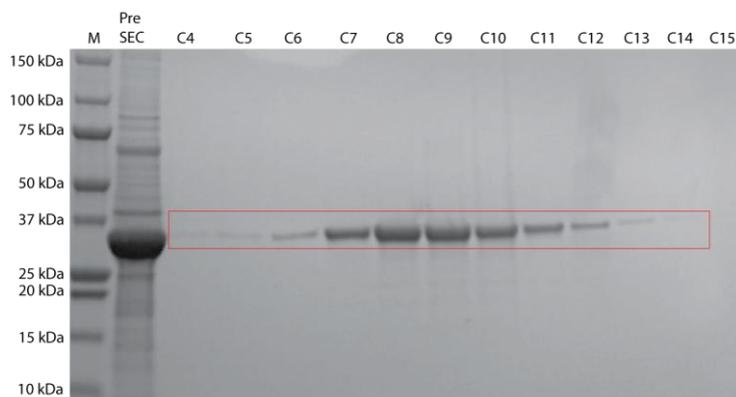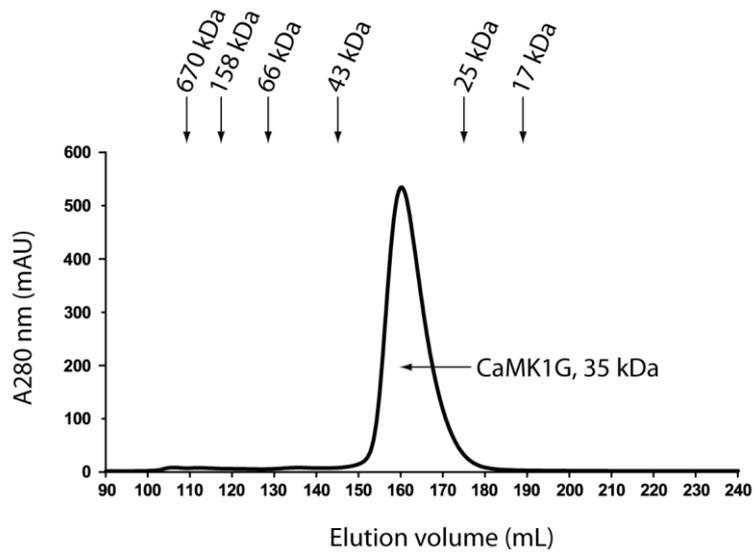
Figure 3.53. CaMK1G size exclusion elution profile and SDS PAGE.

The 35 kDa CaMK1G elutes from about 155–180 ml and produces a monomeric peak that corresponds to 90% pure protein on SDS PAGE. Fractions pooled are indicated by the red box.
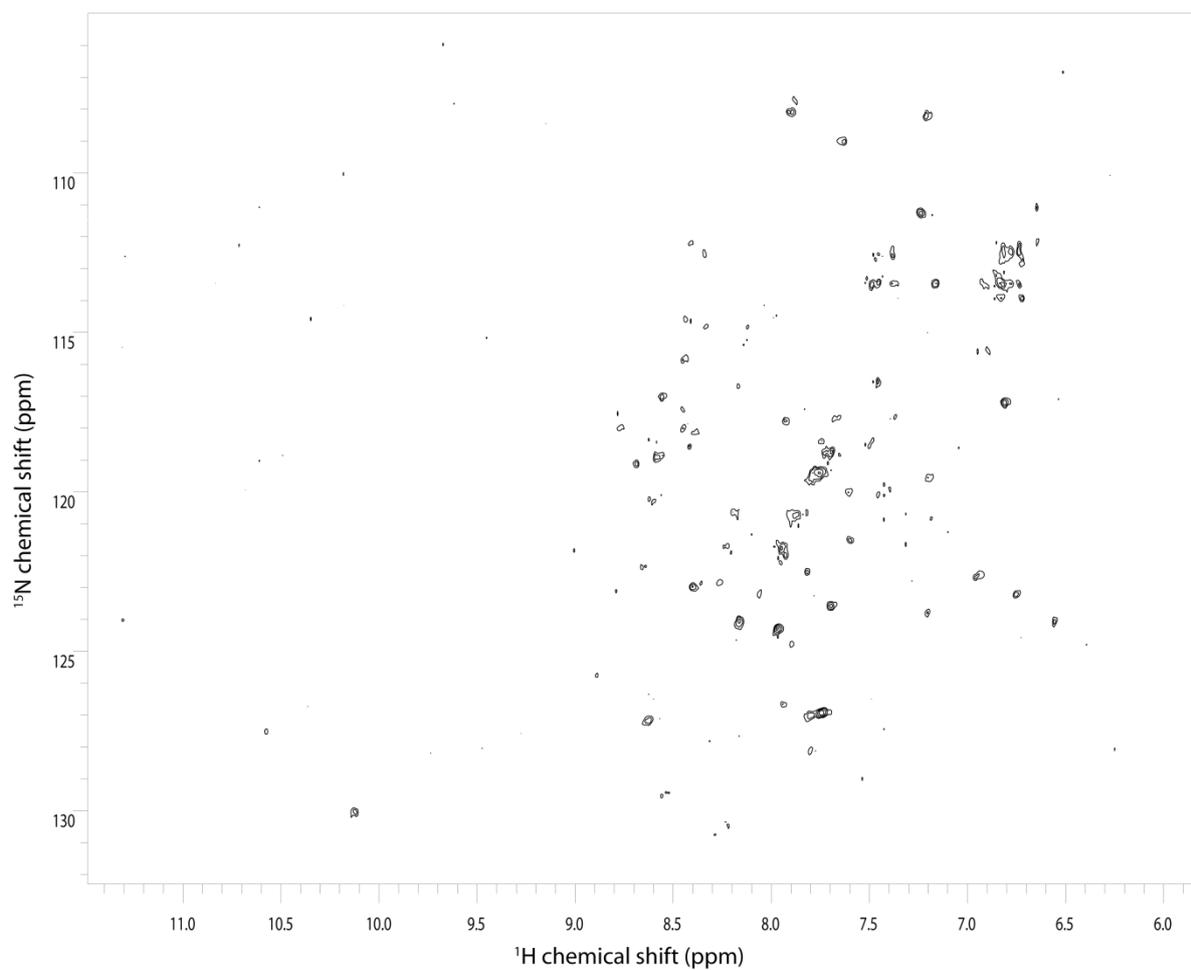
Figure 3.54. CaMK1G ^{1}H, ^{15}N HSQC.

The ^{1}H, ^{15}N HSQC for CaMK1G displays a very low number backbone amide resonances suggesting that the majority of the resonances have broadened out due to intermediate exchange. The spectrum was acquired in 50 mM Na Phosphate pH 7.5, 150mM NaCl, 0.5mM TCEP, 0.02% Na Azide at 25°C on a Varian 800 MHZ.

CaMK1G dynamics suggests a relatively mobile conformation

The B-factor values for CaMK1G (Figure 3.55) indicates that the whole protein is relatively mobile and dynamic with many loops and secondary structure elements across both lobes displaying intermediate B-factor values. In particular high B-factor values are also evident in the N-terminus, C-terminus, the regulatory αC helix, αI helix, and a loop between β1 and β2. In contrast to the crystal structure of CaMK1D, the activation loop of CaMK1G is well defined with the first few residues of the activation loop starting from DFG exhibiting low B-factor values. A comparison between CaMK1D and CaMK1G of the core regions of the N-terminal and C-terminal lobes implies that CaMK1G is more dynamic because its shows more intermediate B-factors in these areas.

The region between the lobes comprising the active site cleft also appears more dynamic than in CaMK1D suggesting that there is more interlobe conformational plasticity associated with CaMK1G in its autoinhibited state than CaMK1D. The plasticity of these residues is presumably manifested in the ^{1}H, ^{15}N HSQC with many resonances broadening out into intermediate exchange. Furthermore the AID of CaMK1G in comparison to CaMK1D appears to be less rigid as it is highlighted by intermediate and high B-factors towards the C-terminal end of the first helix of the AID. The AID of CaMKIs is typically defined by a helix-loop-helix structure but it is evident in the crystal structure of CaMK1G that the latter part of the AID comprising the second helix was not determined. This missing region comprises of 14 residues corresponding to the sequence KWRQAFNAAAVVHH and we assume that they were not defined structurally due to poor electron density indicative of extreme flexibility and dynamics. The stability of CaMK1G is therefore expected to be lower than CaMK1D due to lower stability and increased mobility of the CaMK1G AID.

We hypothesise that due to the higher intrinsic dynamic behaviour of CaMK1G and its less rigid AID, it is possible that CaMK1G is able to sample towards a more catalytically competent conformations than CaMK1D. In addition to this the activation loop of CaMK1G is well defined in its crystal structure, an observation that is usually

associated with an active kinase primed for cataylsis because in active kinases the activation loop is often stabilized to enagage in phosphate transfer and substrate interactions. In contrast the activation loop of CaMK1D is not well defined suggesting that it is adopting a more catalytically incompetent conformation. Although CaMKIs require activation loop phosphorylation for maximal activity against most substrates the observation of the stabilized CaMK1G activation loop in its autoinhibited state suggests that this conformation is primed for activity on a basal level.

It would thus be interesting to compare the activation and catalytic loop of CaMK1G with those of constitutively active proteins like Pim1 to examine if CaMK1G is indeed sampling a catalytically competent state reflected by its activation loop. However the possibility of CaMK1G basal activity independent of Ca^{2+}/CaM is unlikely because the AID is still engaged to the substrate binding. This is assuming that the AID although partially flexible is still occluding the substrate binding site via the pseudosubstrate mechanism.

Thermofluor^{TM} suggests that CaMK1G exhibits low thermal stability in buffers that range from pH 6-8

An evaluation of CaMK1G thermal stability in different buffer conditions was conducted using thermofluor^{TM} to explore solution conditions that could improve its low amide peak count and intensities. Our results in figure 3.56 show that CaMK1G does not exhibit much variation in Tm with buffer type and pH whilst showing no obvious correlation between Tm and pH. The differences in Tm between the diffeent buffers was only 0.5 °C which is considered insignifcant and indcates a narrow range. For example, in 50 mM Na phosphate pH 7.0 and 50 mM HEPES pH 7.5, the Tm of CaMK1G was 40.5 and 40.2 respectively °C. Furthermore these Tm values are within the low-moderate Tm stability range. This suggests that CaMK1G is significantly less stable than CaMK1D in the equivalent buffers and is consistent with our hypothesis that the increased flexibility of the CaMK1G and its AID makes CaMK1G less stable than CaMK1D. Due its intrinsic intermediate exchange dynamics and low stability, CaMK1G was subsequently dropped as an NMR target.
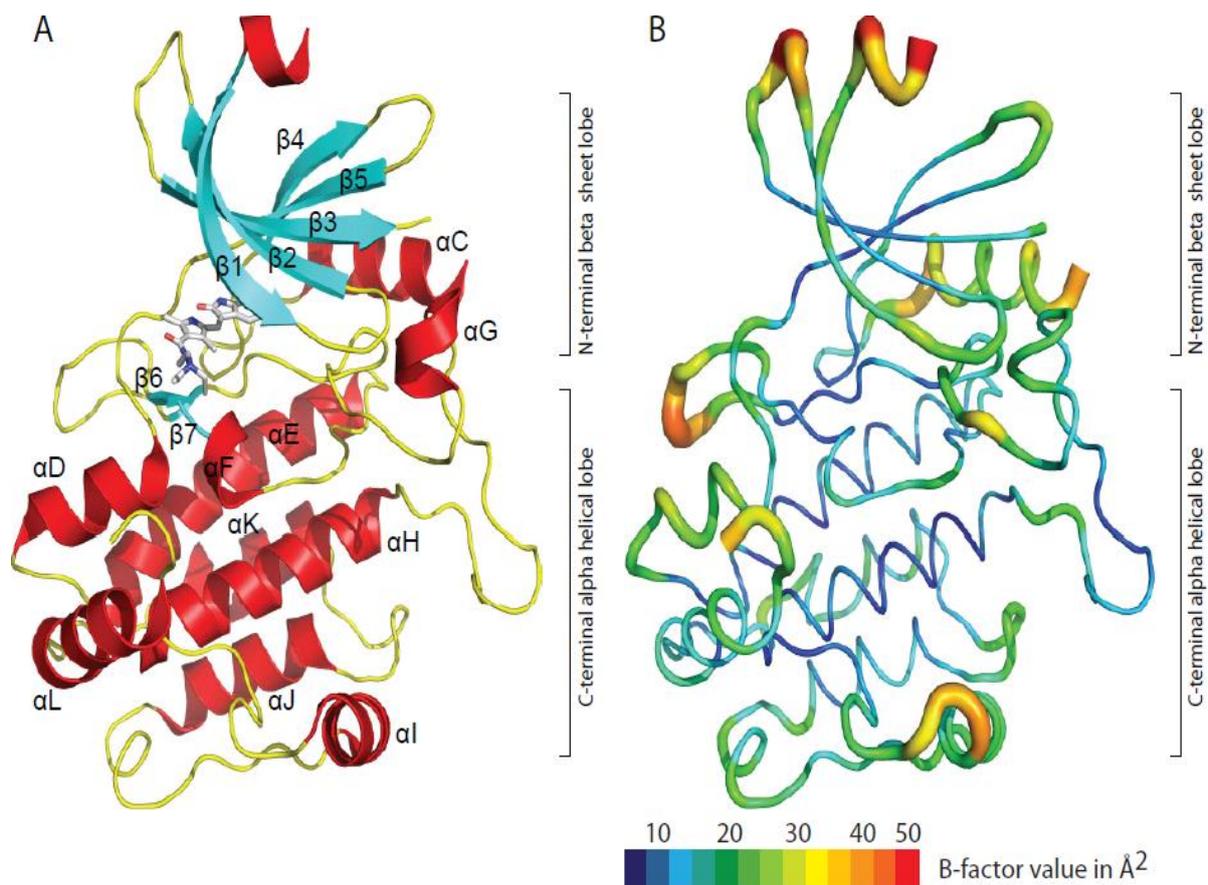
Figure 3.55 CaMK1G-inhibitor crystal structure and B-factors.

The CaMK1G crystal structure in complex with (PDB:2JAM) exhibits a canonical kinase fold comprised of an N-terminal ATP binding lobe and C-terminal substrate binding lobe that is followed by a regulatory AID. The latter part of the AID was not determined. However the activation loop is observable in this crystal structure. The CaMK1G B-factors suggest that globally CaMK1G is a relatively dynamic protein. The latter part of the AID is described by intermediate and high B-factors indicative of increased flexibility.
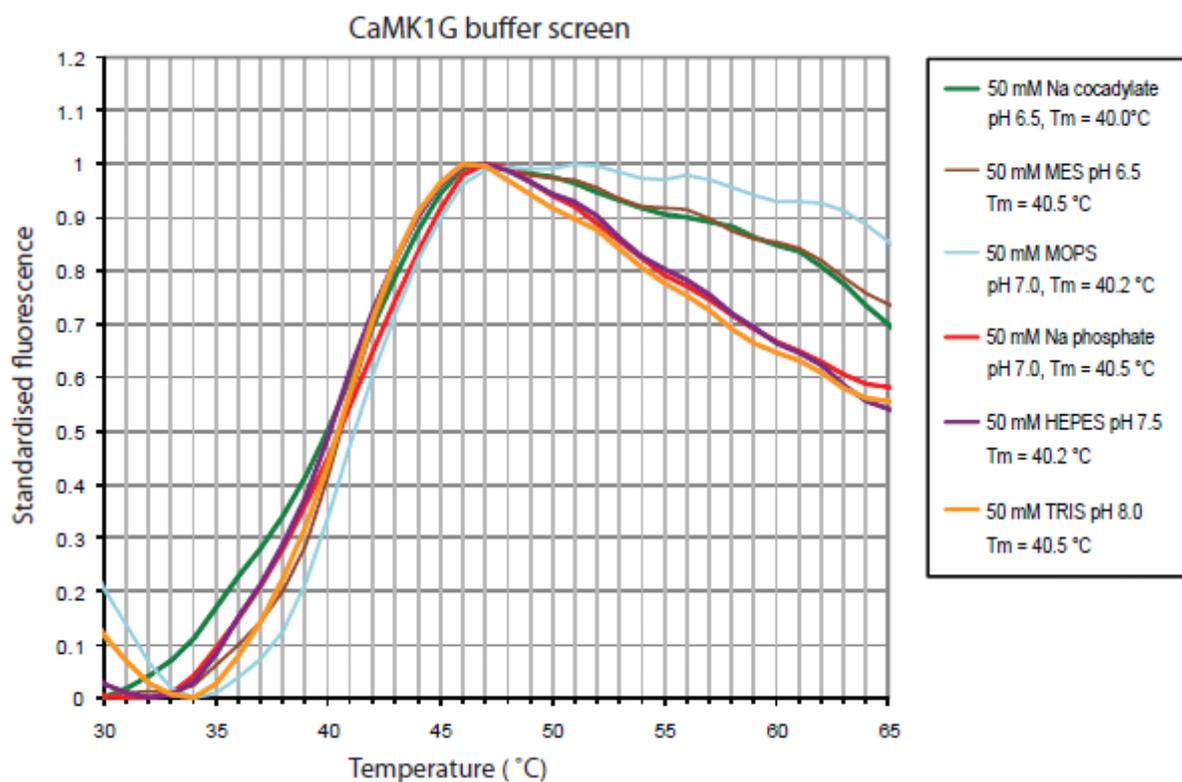
Figure 3.56. CaMK1G thermofluor pH screen.

The thermal stability of CaMK1G was assessed in different buffers types and pH by thermofluorTM. Our results show that the Tm of CaMK1G does not correlate with pH. In addition the Tm of CaMK1G is similar in all buffer conditions with a range of Tm from 40.0-40.5. °C which suggests low-moderate stability.

Evalutation of CSNK1G3 for NMR studies

CSNK1G3 is folded but the majority of amide resonances are not observed in the ^{1}H, ^{15}N HSQC spectrum

CSNK1G3 (Figure 3.57, 3.58) was purified to 95 % purity based on SDS PAGE (Figure 3.59-3.61). Purification by size exclusion chromatography suggested that CSNK1G3 was a monomer in solution. Analysis of the ^{1}H, ^{15}N HSQC for CSNK1G (Figure 3.62) reveals amide resonance dispersion thus confirming that the protein is folded. However most of the backbone amide resonances have broadened beyond practical detection presumably due to intermediate exchange dynamics and aggregation. The few resonances that are observable demonstrate strong, moderate and weak intensities suggesting that CSNK1G3 exhibits dynamic behaviour. Our CSNK1G3 construct only contains one Trp residue and a single Trp side chain amide H$^\varepsilon$ resonance is seen the bottom left corner of the spectrum. Sample precipitation was evident after acquiring the ^{1}H, ^{15}N HSQC at 25 °C indicating low stability.

Figure 3.57. CSNK1G3 full length and CSNK1G3 SGC constructs.

Our CSNK1G3 construct encodes the kinase catalytic domain and a small portion of the C-terminal domain.

```
       40          50          60          70          80

    GVLMVG  PNFRVGKKIG  CGNFGELRLG  KNLYTNEYVA  IKLEPMKSRA

       90         100         110         120         130

PQLHLEYRFY  KQLGSGDGIP  QVYYFGPCGK  YNAMVLELLG  PSLEDLFDLC

      140         150         160         170         180

DRTFSLKTVL  MIAIQLISRM  EYVHSKNLIY  RDVKPENFLI  GRPGNKTQQV

      190         200         210         220         230

IHIIDFGLAK  EYIDPETKKH  IPYREHKSLT  GTARYMSINT  HLGKEQSRRD

      240         250         260         270         280

DLEALGHMFM  YFLRGSLPWQ  GLKADTLKER  YQKIGDTKRA  TPIEVLCENF

      290         300         310         320         330

PEMATYLRYV  RRLDFFEKPD  YDYLRKLFTD  LFDRKGYMFD  YEYDWIGKQL

      340         350         360

PTPVGAVQQD  PALSSNREAH  QHRDKMQQSK  NQ


Predicted pI = 9.0
```

Figure 3.58. CSNK1G3 primary sequence.

The primary sequence encoded for our CSNK1G3 construct and its predicted pI are shown.
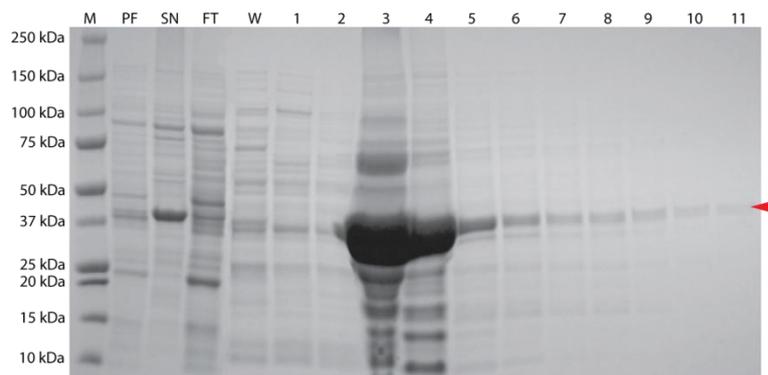
Figure 3.59. SDS PAGE of His$_6$-CSNK1G3 Ni^{2+} NTA purification.

Elution fractions were pooled from fraction number 3-11.
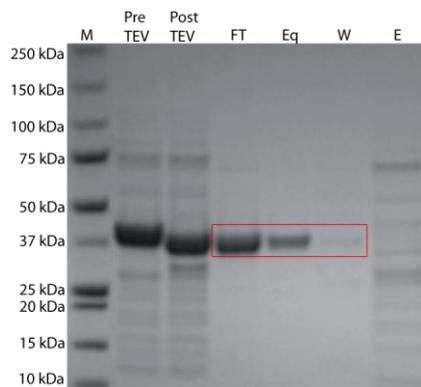
Figure 3.60. SDS PAGE of CSNK1G3 post His$_6$-tag cleavage.

CSNK1G3 in the FT, Eq, and W fractions were pooled after overnight incubation with TEV followed by Ni^{2+} NTA purification to purify from incomplete His$_6$-tag cleavage, free His$_6$-tag, and other protein contaminants.
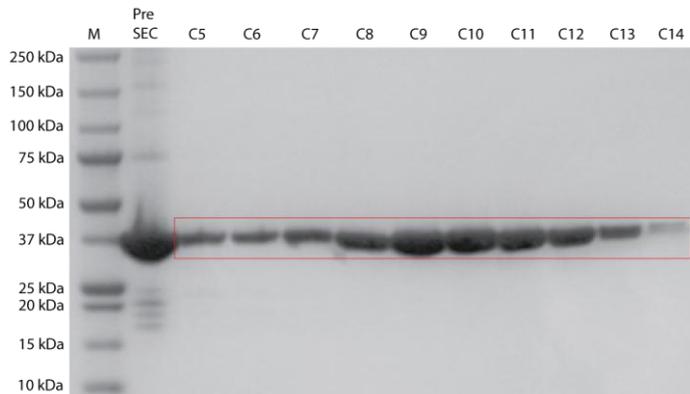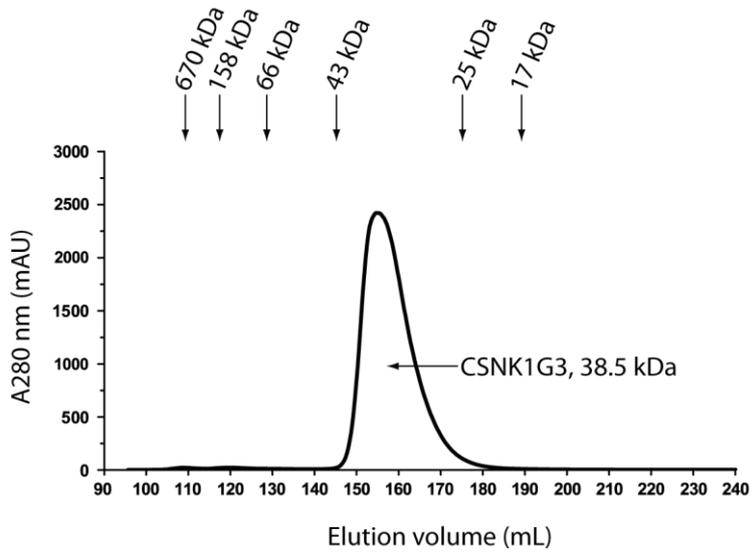
Figure 3.61. CSNK1G3 size exclusion elution profile and SDS PAGE.

The 38.5 kDa CSNK1G3 elutes from 145–180 ml and produces a monomeric peak based on comparison with the elution of protein standards by size exclusion. ractions that were pooled are indicated by the red box.
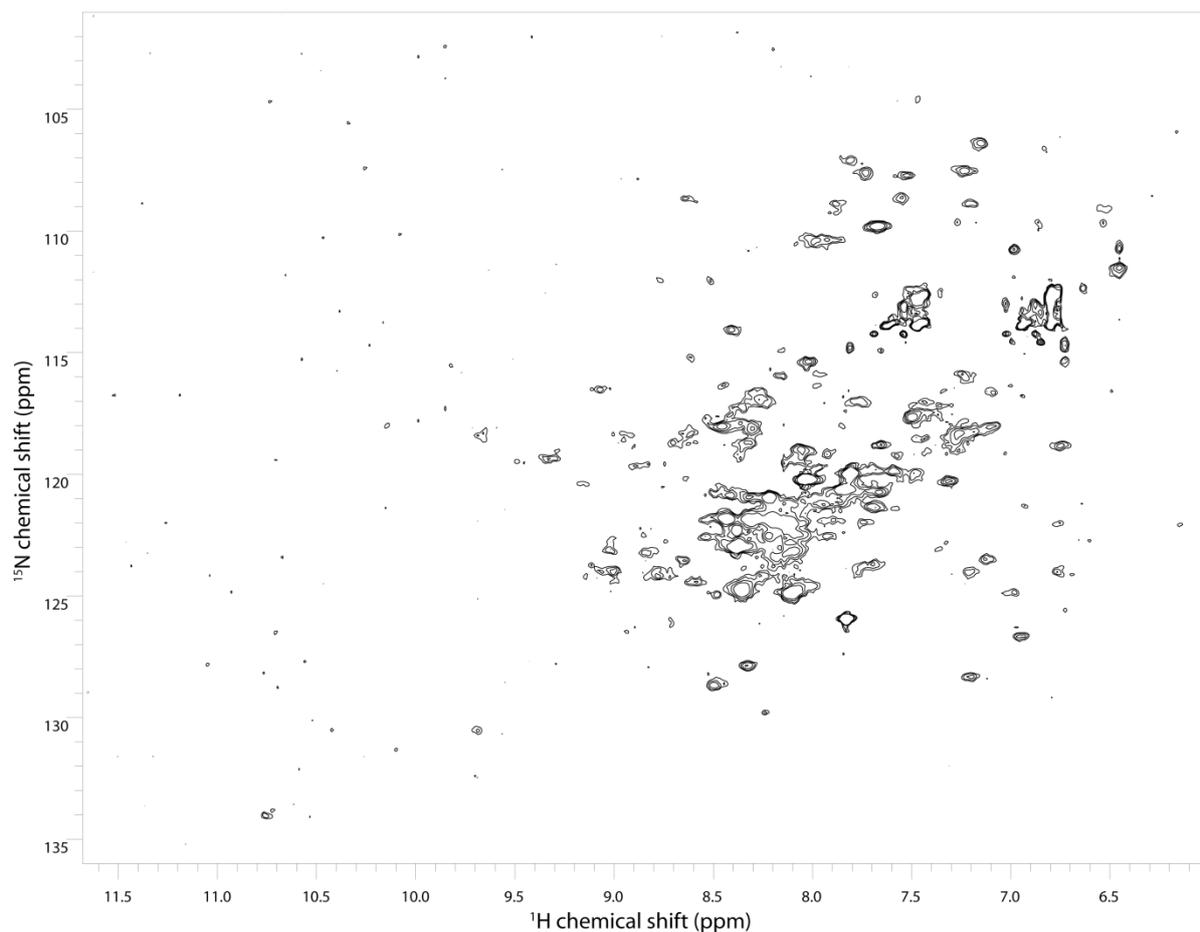
Figure 362. CSNK1G3 ^{1}H, ^{15}N HSQC.

The spectrum illustrates that CSNK1G3 has a folded protein backbone but a large number of resonances have broadened out into intermediate exchange. The spectrum was acquired in 50 mM Na Phosphate pH 7.5, 150mM NaCl, 0.5mM TCEP, 0.02% Na Azide at 25°C on a Varian 800 MHZ.

CSNK1G3 dynamics suggests a relatively mobile conformation

The high B-factors of CSNK1G3 (Figure 3.63) indicate that the CSNK1G3 backbone exhibits dynamic behaviour throughout both lobes. This conforms to the dynamic behaviour inferred from the CSNK1G3 ^{1}H, ^{15}N HSQC spectrum. CSNK1G3 is reported to be constitutively active although it can undergo a psuedosubstrate mode of inhibition mediated by autophosphorylation of its regulatory C-terminal domain. Our CSNK1G construct only encodes the catalytic domain therefore we should anticipate a spectrum comparable to Pim1 assuming that CSNK1G3 is constitutively active like Pim1. If CSNK1G3 is indeed constitutively active then it is likely to dynamically sample multiple catalytically competent conformations. Heterogeneity in the phosphorylation state of CSNK1G3 may also have resulted in the poor spectrum presumably through modulating the dynamics of CSNK1G3 but the issue of sample heterogeneity was not confirmed by mass spec.

Thermofluor^{TM} suggests that CSNK1G3 exhibits low-moderate thermal stability in buffers that range from pH 6-8

In order to improve the amide peak count of CSNK1G3 we decided to optimise its buffer solution by analysing its thermal stability in different buffer types using thermofluor^{TM} (Figure 3.64). Melting curves produced for CSNK1G3 in the different buffer types and pH indicated that the Tm ranged from 40.0-52.0 °C. However there is no correlation between Tm and pH. Considering that CaMK1G has a predicted pI of 9.0 we might have expected an increasing in Tm with decreasing pH but this is not evident. The highest Tm from the buffers tested was produced by 50 mM Na phosphate pH 7.0. This is comparable to the Tm for CaMK1D in the same buffer and was used as the default buffer for CSNK1G3 including 150 mM NaCl as a default stabilizer.

However after acquiring a ^{1}H, ^{15}N HSQC at 25 °C samples recipitation became evident suggesting that the Tm of CSNK1G3 is lower than predicted by thermofluor^{TM}. Furthermore the CSNK1G3 predicted pI of 9.0 would suggest that at lower pH the stability of the protein should increase but this was observed in the

thermofluor^{TM} data. We postulate that due to the constitutively active nature of CSNK1G3 and the absence of the inhibitory domain the construct is intrinsically dynamic and unstable compared to its autoinhibited form. Consequently optimisation of the ^{1}H, ^{15}N HSQC for CSNK1G3 and its thermal stability were not progressed further and CSNK1G3 was dropped as an NMR target.
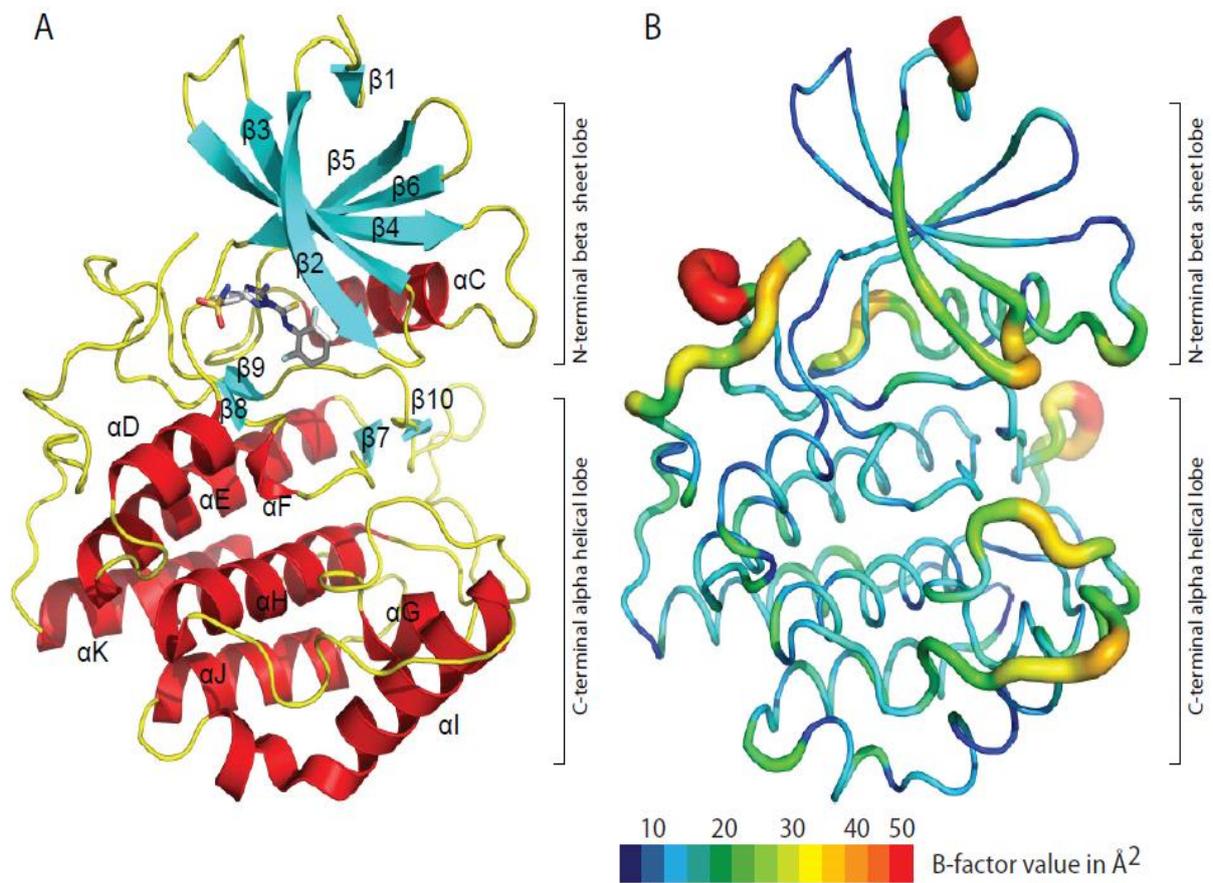
Figure 3.63 CSNK1G3-inhibitor crystal structure and B-factors.

The CSNK1G3 crystal structure in complex with (PDB: 2CHL) showing the catalytic domain that is comprised of an N-terminal ATP binding lobe and C-terminal substrate binding lobe. The activation loop is observable in this crystal structure. The CSNK1G3 B-factors illustrate that CSNK1G3 is a relatively dynamic protein, particularly at the termini and loops.
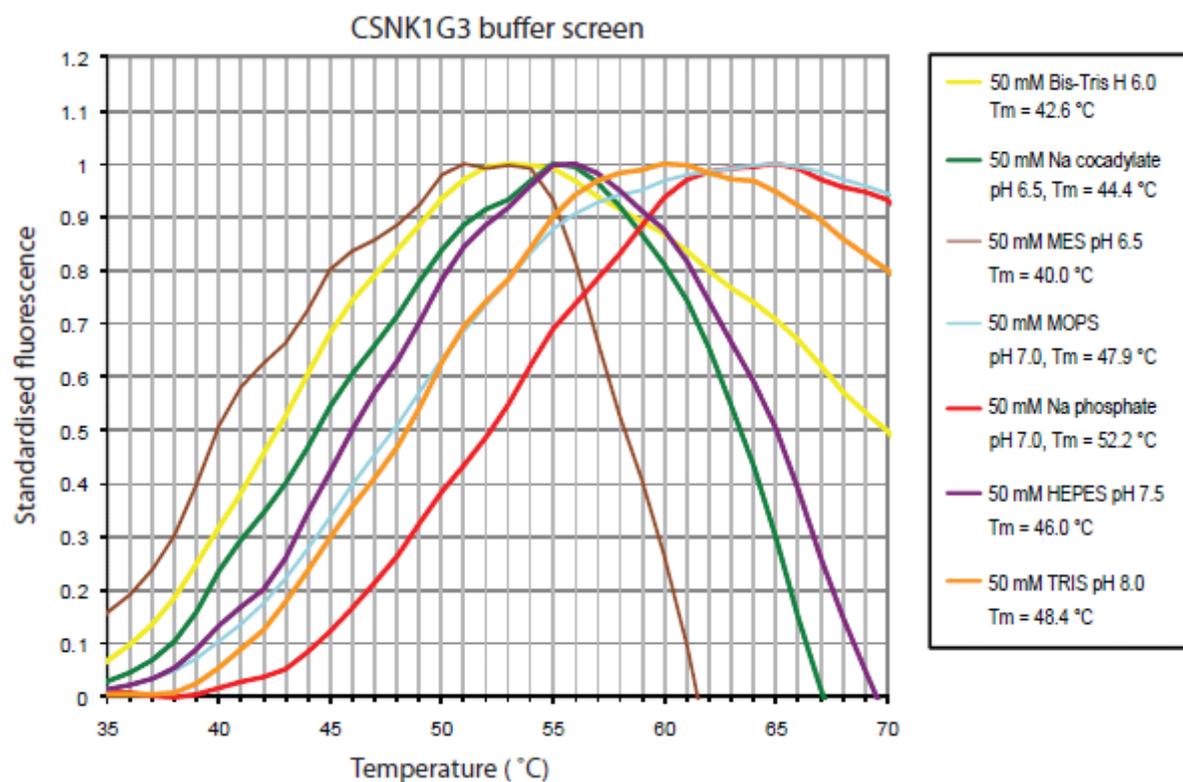
Figure 3.64. CSNK1G3 thermofluor pH screen.

The stability of CSNK1G3 was characterised in different buffer conditions and found to produce a range of Tm but without any correlation between stability and pH. Notably theTm of CSNK1G3 in 50mM Na Phosphate pH 7.0 was deduced to be 52.2 °C. However, under actual NMR experiment conditions the protein precipitated at 25 °C despite using this buffer type and pH in the default NMR. This suggests that the predicted Tm by thermofluor is an overestimate of the actual Tm of CSNK1G3.

Expression of CaMK1D and mutants

CaMK1D – an NMR amenable kinase target

Following our evaluation of six kinases (CaMK1D, CaMK1G, CSNK1G3, PAK5, Pim1 and TOPK) in terms of their expression, stability, solubility and spectral quality it was decided that CaMK1D was the best candidate for study by NMR. The next aim was to optimise the expression levels of milligram quantities of 95 % pure and isotopically labeled soluble CaMK1D and mutants in order to initiate the backbone assignment of CaMK1D and subsequent structural, functional and ligand interaction studies.

Additional CaMK1D mutant constructs (Figure 4.0) were designed by using site directed mutagenesis (Stratagene). An activation loop mutant termed CaMK1D_T180E was designed by mutating Thr 180 to Glu in order to mimic the phosphorylation of Thr 180 in maximally active CaMK1D. CDB mutants CaMK1D_W306S and CaMK1D_M319A were also produced to investigate the functional importance of these residues in mediating interactions between Ca^{2+}/CaM and the CBD. Expression and purification results for CaMK1D and mutants are discussed in the following sections.
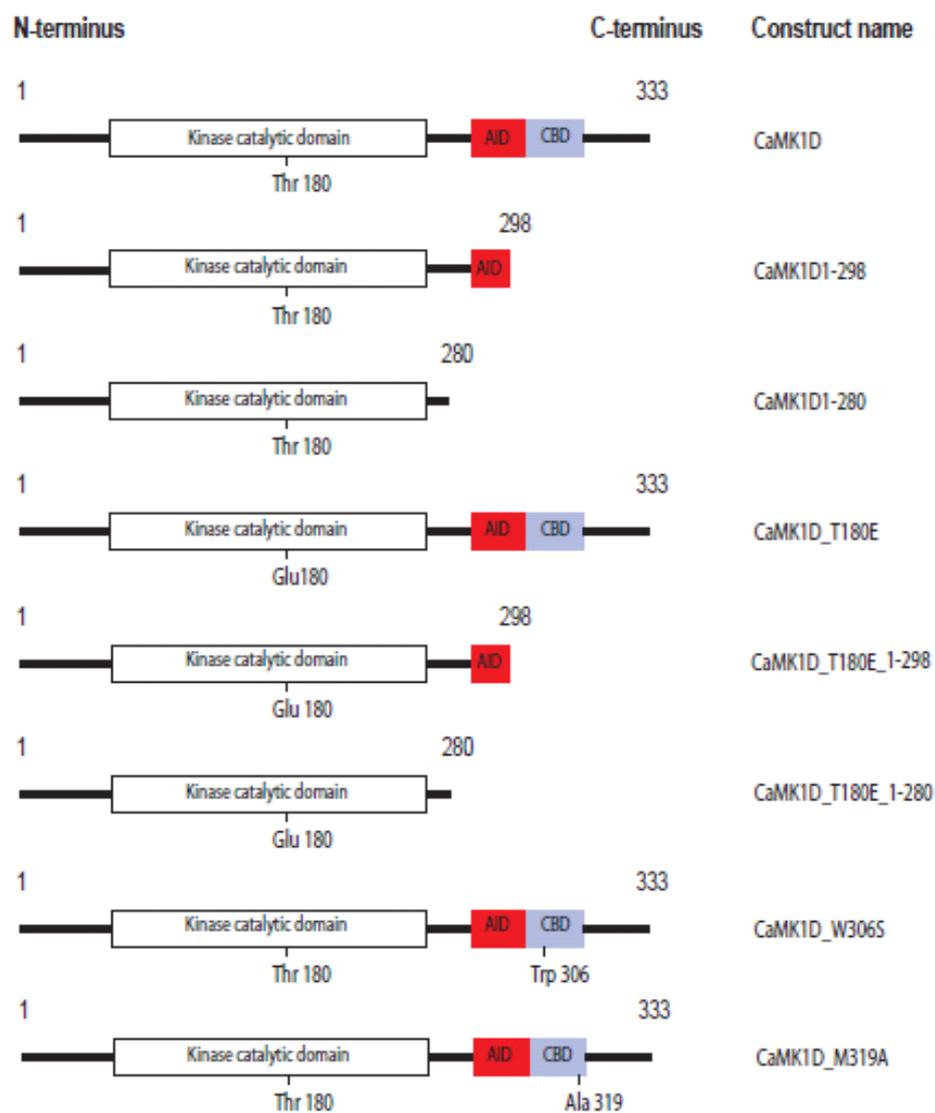
Figure 4.0. CaMK1D contructs for NMR.

A schematic summarising all of the CaMK1D constructs that were expressed and purified.

Expression tests of CaMK1D and mutants in LB suggests an optimal post induction time of 16 hours

In order to maximise the yield of soluble protein, the post induction time for CaMK1D and the activation loop mutant CaMK1D_T180E were determined in Luria Broth (LB) by comparing 8 hours against 16 hours post induction time (Figure 4.1). Analysis of the post induction times by SDS PAGE indicates that after 16 hours at 18°C, whole cell expression levels of CaMK1D are greater compared to 8 hours post induction. The same outcome was demonstrated for CaMK1D_T180E indicating that the mutation did not decrease expression levels. Expression tests of the CBD mutants CaMK1D_W306S and CaMK1D_M319A in LB confirmed that these constructs could also be over expressed after 16 hours induction at 18°C and that the mutations did not decrease expression levels. Thus the optimal post induction time for over expressing CaMK1D, CaMK1D_T180E and the CBD mutants was judged to be 16 hours. The default cell line Rosetta (DE3) was used for the expression test of CaMK1D and CaMK1D_T180E in LB because the cell line provides an abundance of rare codons for *E. coli* expression. Expression tests of CaMK1D_W306S and CaMK1D_M319A in LB were conducted using BL21 (DE3) RIPL codon plus, because the expression test for these constructs were conducted after the original Rosetta (DE3) cell line had been optimised as described in the following section.
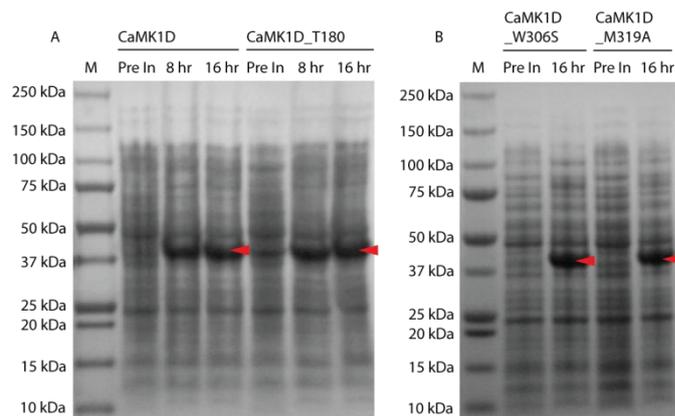
Figure 4.1 A. SDS PAGE of CaMK1D and CaMK1D_T180E whole cell expression.

CaMK1D was over expressed following 8 and 16 hours post induction in LB. The position of the His$_6$-tagged CaMK1D band (40.2 kDa) is denoted by the red arrow. **B. SDS PAGE showing whole cell over expression of CaMK1D_W306S, and CaMK1D_M319A.** CaMK1D CBD mutants CaMK1D_W306S and CaMK1D_M319A were over expressed for 16 hours in LB. The position of the His$_6$-tagged CaMK1D_W306S and CaMK1D_M319A bands (both 40.2 kDa) is denoted by the red arrow.

SDS PAGE abbreviations: M = marker, Pre = preinduction, 8hr = 8 hour post induction, 16hr = 16 hour post induction

CaMK1D requires rare codon supplmentation provided by Rosetta (DE3) for optimal expression

The default cell line for CaMK1D expression was Rosetta (DE3). This was based on the protocol established by the SGC. Rosetta (DE3) is an *E. coli* strain designed to enhance the expression of eukaryotic proteins which are not optimised for *E.coli* codon usage by encoding rare tRNAs that recognise rare codons sometimes found in eukaryotic genes. Analysis of codon usage frequency of CaMK1D using the graphical codon usage analyser software (GCUA) revealed for example that codons encoding several Leu residues have a low frequency usage in *E.coli* (Figure 4.2) which could potentially decrease the translation efficiency and yield of soluble protein. The CaMK1D construct was created by the SGC but was not codon optimised for *E. coli* BL21 (DE3) expression therefore CaMK1D expression in Rosetta (DE3) was preferable for the production of unlabelled CaMK1D in LB and ^{1}H, ^{15}N-labelled CaMK1D in M9 minimal media. Alternatively a synthetic CaMK1D codon optimised sequence for *E. coli* BL21 (DE3) expression could also be generated but this was not pursued. M9 minimal media was used for the production of isotopically labelled proteins because it was cost effective compared to other commercially available media such as bioexpress (Isotec).

Deuteration of CaMK1D is necessary

Our CaMK1D construct is 37.8 kDa which makes structural analysis by NMR very difficult because of increased resonance broadening and overlap associated with large proteins (Foster, McElroy et al. 2007). In order to proceed with sequential backbone assignment of CaMK1D by NMR it was necessary to incorporate deuteration and implement TROSY experiments in order to address these potential large molecular weight issues.

Rosetta (DE3) exhibit a prolonged lag phase in media containing ^{2}H$_2$O

Initial ^{2}H, ^{15}N and ^{2}H, ^{15}N, ^{13}C labelled CaMK1D samples were produced using Rosetta (DE3) cells but the transformed cell line was found to exhibit a prolonged lag phase of growth particularly in the presence of ^{2}H$_2$O (deuterium oxide). *E.coli* cells

grown in ^2H_2O containing media experience stress that can result in an increased lag phase, reduced biomass and protein synthesis mainly due to the inhibitory effects of ^2H_2O on the TCA cycle (Hochuli, Szyperski et al. 2000). In our experience Rosetta (DE3) generally exhibits a longer lag phase compared to BL21 (DE3) and in the presence of ^2H_2O the Rosetta (DE3) lag phase is increased further. Therefore cell line optimisation was attempted for more efficient production of ^2H ^{15}N and ^2H, ^{15}N, ^{13}C labelled CaMK1D samples.

BL21 (DE3) codon plus identified as an optimal cell line for growth and expression in deuterated media

The default SGC CaMK1D expression protocol (SGC) was used as a starting point and translated to a small scale expression test that compared CaMK1D expression levels in Rosetta (DE3), BL21 (DE3), and BL21 (DE3) RIPL codon plus in M9 minimal media containing 70 % and 99 % ^2H_2O content (Figure 77). The BL21 (DE3) and BL21 (DE3) RIPL codon plus cultures were tested for expression due to their shorter lag phases relative to Rosetta (DE3). Like Rosetta (DE3), BL21 (DE3) RIPL codon plus also encodes rare tRNAs for the recognition of rare eukaryotic codons allowing for efficient translation of eukaryotic proteins. In 70 % and 99 % ^2H_2O, Rosetta (DE3) exhibited the longest lag phase in comparison to BL21 (DE3) and BL21 (DE3) RIPL codon plus, typically taking an additional 2–4 hours respectively to reach approximately an O.D$_{600nM}$ of 0.4. Thus BL21 (DE3) and BL21 (DE3) RIPL were better adapted to the ^2H_2O environment with shorter lag phases of growth.

Whole cell expression levels of CaMK1D in 70 % ^2H_2O compared to 99 % ^2H_2O did not show a significant difference between BL21 (DE3) and BL21 (DE3) RIPL codon plus cells. This suggested that the cells became adapted to the high ^2H_2O content and maintained their growth and expression. Rosetta (DE3) had greater expression level in 70 % ^2H_2O than 99% ^2H_2O. A low level of CaMK1D expression level in Rosetta (DE3) in 99 % ^2H_2O was most likely due to poor induction because the level of whole cell protein in the pre and post induction samples appeared the same.

A comparison of CaMK1D whole cell expression between the cell lines shows that BL21 (DE3) RIPL codon plus (in 70 % and 99 % ^2H_2O) and Rosetta (DE3) (in 70 % ^2H_2O only) exhibit similar levels of over expression whereas over expression in BL21 (DE3), appears lower than both of these cell lines. This is perhaps expected because BL21 (DE3) does not encode rare tRNAs that recognise rare eukaryotic codons thus translation is not efficient and can stall if the pool of rare tRNAs is depleted. To summarise, the expression of CaMK1D in BL21 (DE3) RIPL codon plus and Rosetta (DE3) in 70% ^2H_2O appeared equivalent by SDS PAGE but Rosetta (DE3) has a longer lag phase thus making BL21 (DE3) RIPL codon plus the preferred cell line for growth and expression in deuterated M9 minimal media.
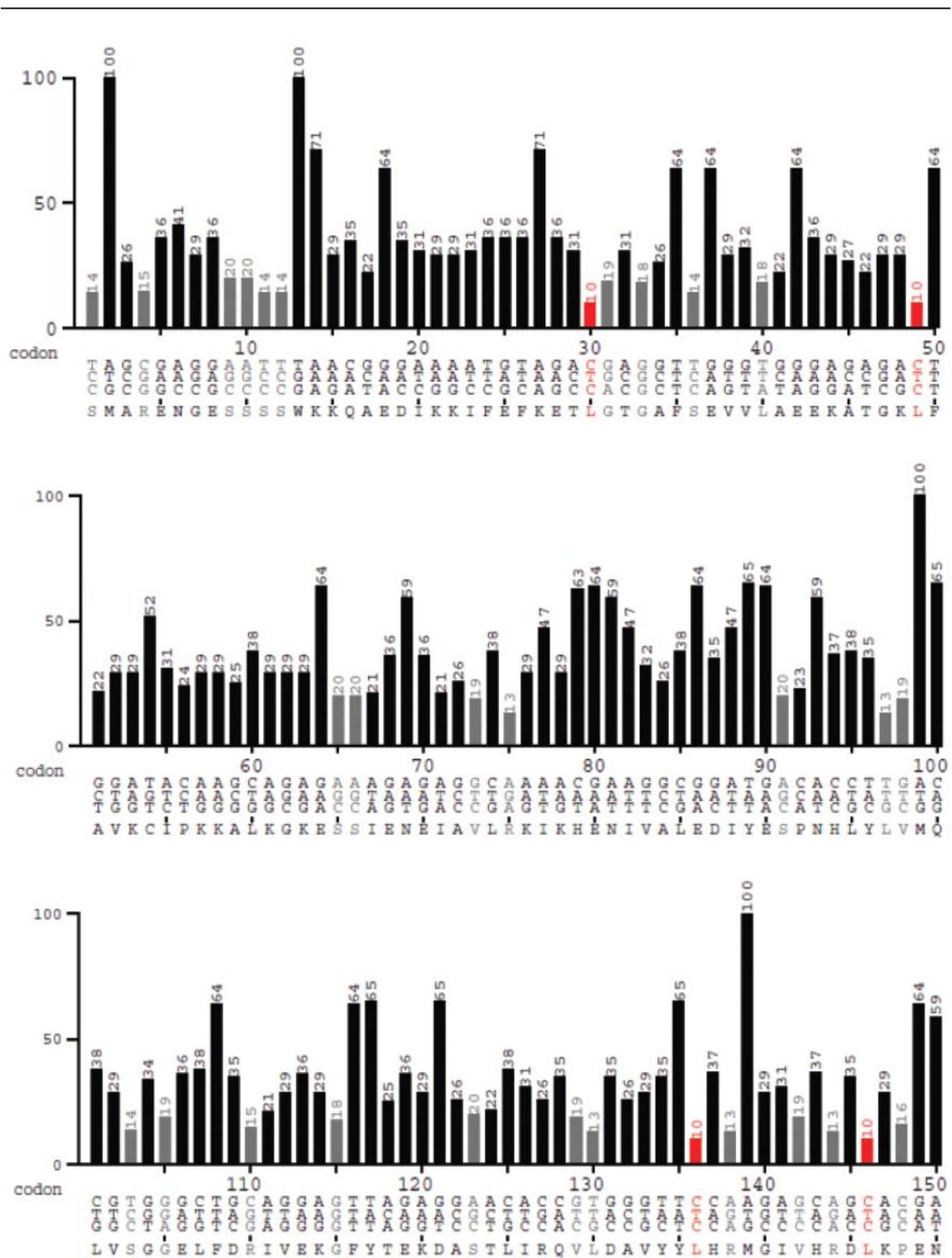
Figure 4.2. CaMK1D residue 1-150 codon frequency usage in *E.coli*.

The analysis is shown for residues 1-150 and is based on the human CaMK1D cDNA provided by the SGC. Rare codons are shown in red.
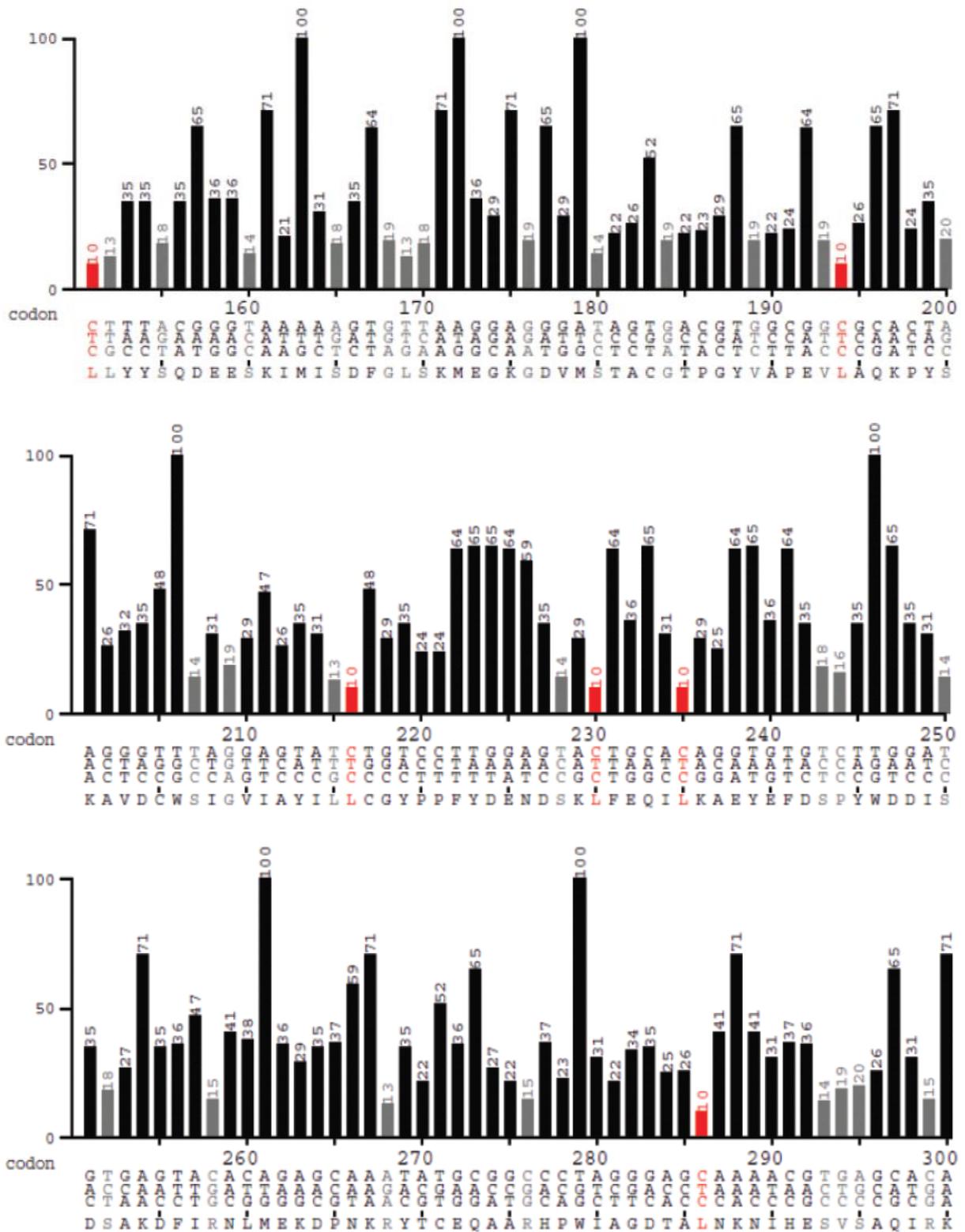
Figure 4.2.1. CaMK1D resiue151-300 codon frequency usage in *E.coli*.

The analysis is shown for residues 151-300 and is based on the human CaMK1D cDNA provided by the SGC. Rare codons are shown in red.
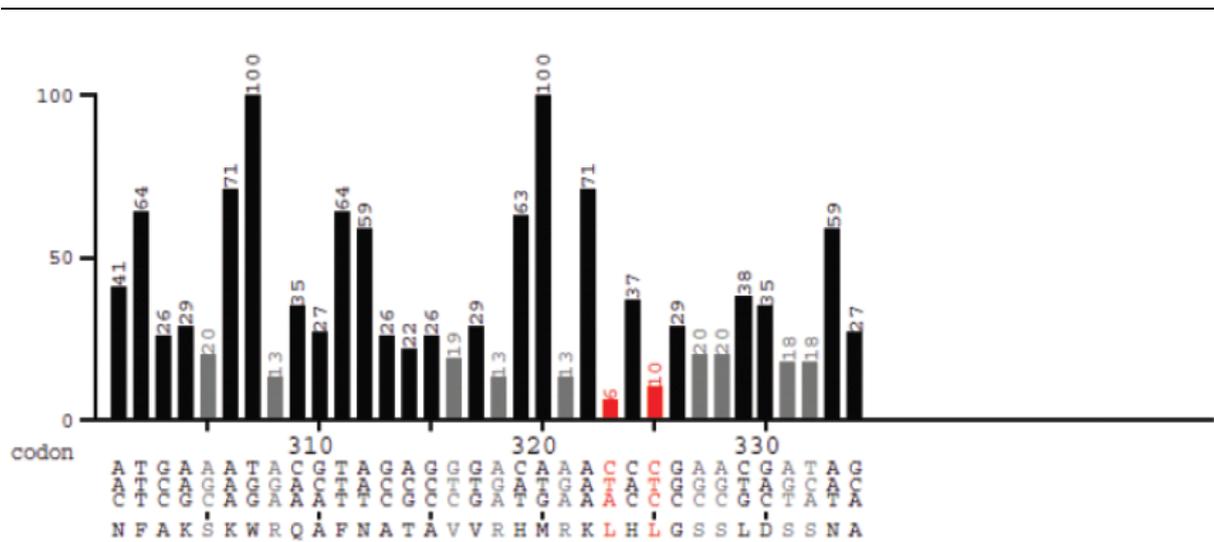
Figure 4.2.2. CaMK1D residue 301-334 codon frequency usage in *E.coli*.

The analysis is shown for residues 301-334 and is based on the human CaMK1D cDNA provided by the SGC. Rare codons are shown in red. Analysis was performed using the graphical codon usage analyser (GCUA). The y-axis shows the frequency of the codon in *E.coli*.
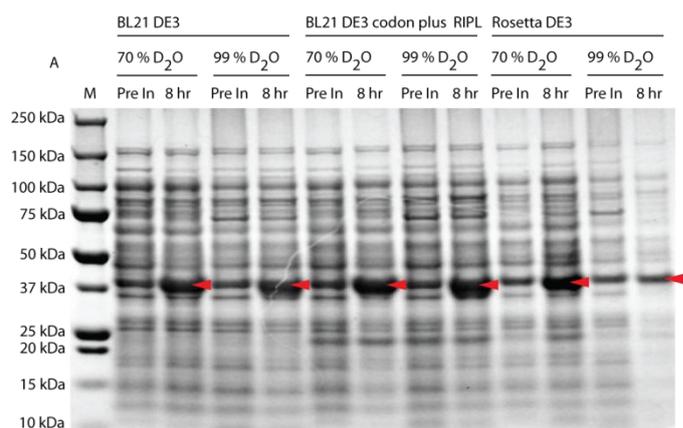
Figure 4.3. SDS PAGE of CaMK1D expression test in different cell lines and ^2H_2O content in M9 media.

Whole cell expression levels of CaMK1D (indicated by the red arrow) were compared in BL21 (DE3), BL21 (DE3) RIPL codon plus, and Rosetta (DE3) cell lines in M9 minimal media containing 70 % (v/v) and 99 % (v/v) ^2H_2O content. Consequently BL21 (DE3) RIPL codon plus was found to be the optimal cell line for expression of deuterated CaMK1D samples.
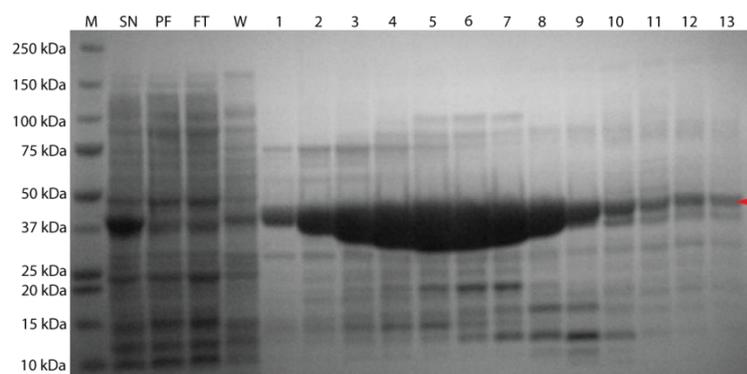
Figure 4.4. SDS PAGE of His$_6$-tagged CaMK1D purified by Ni^{2+} NTA affinity purification.

Elution fractions are shown containing His$_6$-tagged CaMK1D which appears to have a molecular weight of about 40 kDa (the migration His$_6$-tagged CaMK1D on SDS PAGE is indicated by the red arrow). His$_6$-tagged CaMK1D was eluted using a linear imidazole gradient. Contaminating *E.coli* proteins are also visible.
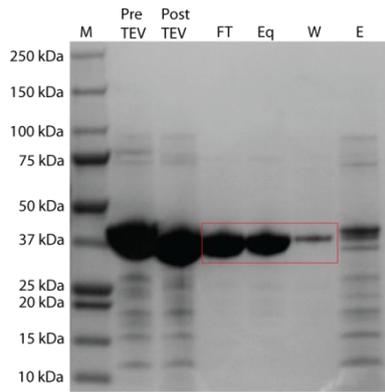
Figure 4.5. SDS PAGE of CaMK1D post His$_6$-tag cleavage.

Removal of contaminants that co-purify with CaMK1D in the initial Ni^{2+} NTA affinity purification was mediated by cleavage of the His$_6$-tag and dialysis of imidazole followed by a second Ni^{2+} NTA affinity purification step. The red box highlights the fractions containing CaMK1D that were purified from the contaminants.

Size exclusion chromatography purification of CaMK1D and mutants

CaMK1D purifies as a monomeric protein by size exclusion chromatography

After the second Ni^{2+} NTA-affinity purification step, CaMK1D was purified by SEC on a S75 26/60 column. The preliminary running buffer used was 50 mM Na phosphate pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide, but for future purifications in which samples were used for ligand interaction studies, the buffer was changed to 50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP 0.02 % Na Azide. This was necessary in order to compare ligand interaction studies of autoinhibited inactive CaMK1D with active CaMK1D•Ca^{2+}/CaM, because Ca^{2+} is precipated by the phosphate component from Na phosphate buffer. The CaMK1D SEC profile (Figure 56) displayed two resolved peaks with an elution volume from 150-180 ml that corresponded to 95% pure CaMK1D as determined by SDS PAGE (Figure 80). Based on comparison with the size exclusion elution profile of protein standards the position and shape of the CaMK1D peak suggests that CaMK1D is a monomer with an apparent molecular weight of 37.8 kDa. A high molecular weight *E.coli* protein contaminant is shown eluted from 120-140 ml (figure) and was well resolved from the CaMK1D peak (150-180ml). Analysis of the pre size exclusion CaMK1D sample by SDS PAGE reveals that relative to CaMK1D minor amounts of *E.coli* contaminating proteins of high and low molecular weight are still evident but were removed from CaMK1D after size exclusion because the CaMK1D size exclusion elution fractions appear $\geq$ 95 % pure. A final determination of the protein estimates approximately 25 mg of protein L^{-1} of culture for unlabelled protein and 20 mg of deuterated and isoptocially labelled protein L^{-1} of culture.

CaMK1D mutants exhibit similar purification results to CaMK1D by size exclusion

CaMK1D_T180E, CaMK1D_W306S and CaMK1D_M319A were found to exhibit similar size exclusion elution profiles to CaMK1D. The similarity therefore suggested that the mutations did not alter the oligomeric state of the protein or global shape. The purity and yields of the mutant constructs was also comparable to that of CaMK1D (SDS PAGE data not shown).
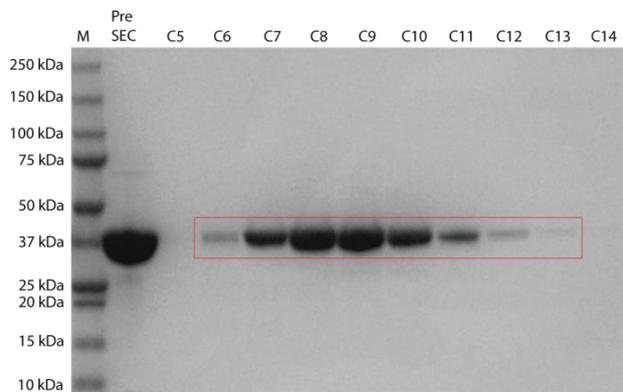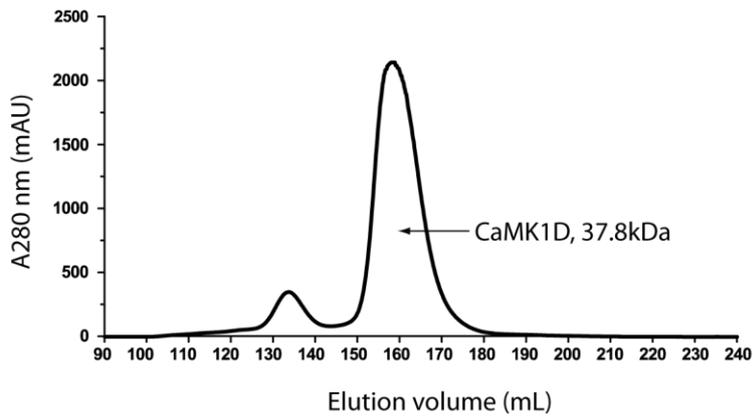
Figure 4.6. CaMK1D Size exclusion and SDS PAGE

CaMK1D appears homogenous and 95 % pure after size exclusion purification. Elution fractions shown correspond to the fractions under the CaMK1D size exclusion peak from approximately 150–180 ml (Figure). PSE = pre size exclusion CaMK1D.

Expression of CaMK1D C-terminal truncations

CaMK1D1-298 exhibits higher amounts of soluble protein than CaMK1D1-280 and likewise for T180E C-terminal truncation variants

In order to characterise active forms of CaMK1D lacking the AID sequence, several C-terminal truncations of CaMK1D and CaMK1D_T180E were generated by mutagenesis. Based on previous cell line optimisation results with CaMK1D, these constructs were expressed in BL21 (DE3) RIPL codon plus. Expression tests were conducted to assess the levels of over expressed protein. The whole cell protein, soluble and insoluble protein levels were tested because these constructs were not previously expressed or purified. Whole cell expression tests revealed that CaMK1D1-280, CaMK1D1-298, CaMK1D_T180E_1-280 and CaMK1D_T180E_1-298 exhibit greater over expression after 16 hours induction at 18°C compared to 4 and 8 hours (Figure 81).

The levels of soluble protein for CaMK1D1-280 and CaMK1D_T180E_1-280 appeared to be 50 % soluble and 50 % insoluble (Figure 82). The soluble fractions of CaMK1D1-298 and CaMK1D_T180E_1-298 exhibited about 75 % soluble protein compared to their respective insoluble fractions. This suggests that at 18 °C, 16 hours post induction, CaMK1D1-298 and CaMK1D_T180E_1-298 and are more stable compared to CaMK1D1-280 and CaMK1D_T180E_1-280. The difference in yield of soluble protein is likely due to the stabilizing inhibitory role of the AID because the longer constructs CaMK1D1-298 and CaMK1D_T180E_1-298 still contain part of the C-terminal AID which can impart some stability on the kinase structure. CaMK1D1-280 and CaMK1D_T180E_1-280 completely lack the AID thus the interactions of the AID with the substrate binding site and N-terminal ATP binding lobe seen in the crystal structure are absent and these shorter constructs are less stable and exhibit more tendency to precipitate.
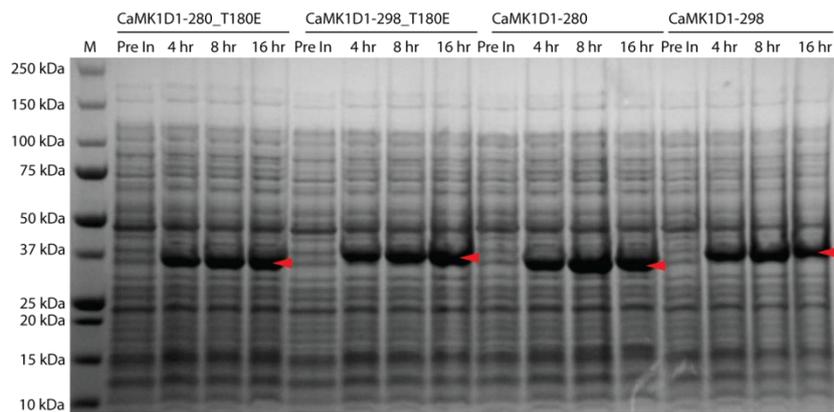
Figure 4.7. SDS PAGE of CaMK1D and CaMK1D_T180E C-terminal truncation mutants expression test.

Cultures were induced at 18°C for 4, 8, and 16 hours and post expression samples were taken for each time point respectively. Expression appears greater at 16 hours post induction (red arrow).
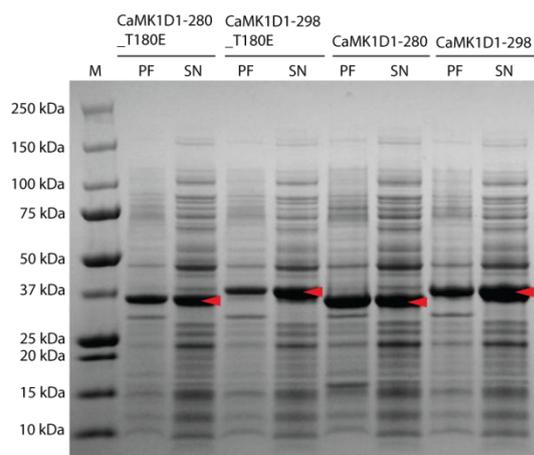
Figure 4.8. SDS PAGE of CaMK1D and T180E C-terminal truncation insoluble and soluble fractions.

Amounts of insoluble (PF = pellet fraction) and soluble protein (SN = supernatant) after 16 hours expression at 18 °C were analysed for CaMK1D1-280, 1-298, CaMK1D_T180E_1-280, and CaMK1D_T180E_1-298. The shorter constructs CaMK1D1-280 and CaMK1D_T180E_1-280 appear to have only 50% soluble protein whereas the longer constructs CaMK1D1-298 and CaMK1D_T180E_1-298 appear to have 75 % soluble protein.

Ni^{2+} NTA-affinity purification of CaMK1D C-terminal truncations

C-terminal truncated CaMK1D constructs co-elute with minor contaminants by Ni^{2+} NTA affinity

The C-terminal truncation constructs termed CaMK1D1-280, CaMK1D1-298 and their activation loop equivalents CaMK1D_T180E_1-280 and CaMK1D_T180E_1-298 were all purified via their N-terminal His$_6$-tag in the same manner described for CaMK1D using a Ni^{2+} NTA-affinity column. Only the results for CaMK1D1-280, and CaMK1D1-298 are described here because their respective activation loop mutants demonstrate similar purification results. Analysis by SDS PAGE of the elution fractions from Ni^{2+} NTA-affinity purification reaveals that both His$_6$-tagged CaMK1D1-280 and His$_6$-tagged CaMK1D1-298 co-eluted with *E.coli* contaminant proteins (Figure 83, 85). These contaminants were removed using the same method described for CaMK1D which involved a second Ni^{2+} NTA-affinity purification step after cleavage of the CaMK1D1-280 and CaMK1D1-280 His$_6$-tag with TEV and dialysis for the removal of imidazole.

Cleavage of the N-terminal His$_6$-tag and a second Ni^{2+} NTA-affinity purification improved the purification of C-terminal truncated CaMK1D constructs

After incubation with TEV protease we show by SDS PAGE that CaMK1D1-280 (31.8 kDa) migrates as a slightly lower molecular weight protein relative to its His$_6$-tagged form (34.2 kDa) thus illustrating successful cleavage of the His$_6$-tag. Likewise CaMK1D1-298 (33.8 kDa) migrates as a slightly lower molecular weight protein relative to its His$_6$-tagged form (36.2 kDa) after His$_6$-tag cleavage. SDS PAGE was also used to identify and confirm the fractions that contained CaMK1D1-280 and CaMK1D1-298 following the second Ni2+ NTA-affinity purification step and it was evident that CaMK1D1-280 and CaMK1D1-298 do not bind to the Ni^{2+} NTA column and therefore appear in the flow through (FT). The SDS PAGE (Figure 84, 86) also shows that CaMK1D1-280 and CaMK1D1-298 that was weakly bound to the column was eluted with equilibration (Eq) and wash (W) buffer whereas the majority of the contaminating *E.coli* proteins bound tightly and did not elute with CaMK1D1-280 and CaMK1D1-298 in the Eq and W fractions but eluted in the elution fraction which used elution buffer.
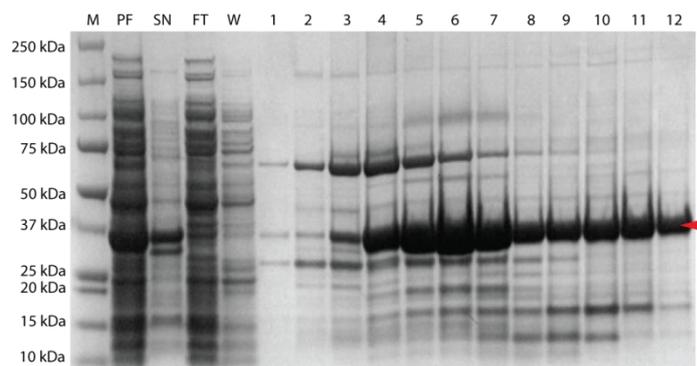
Figure 4.9. SDS PAGE of His$_6$-tagged CaMK1D1-280 from Ni^{2+} NTA affinity purification.

His$_6$-tagged CaMK1D1-280 was eluted with a linear gradient and migrates between the 37 and 25kDa molecular weight standards. Fraction number 3-12 were pooled.
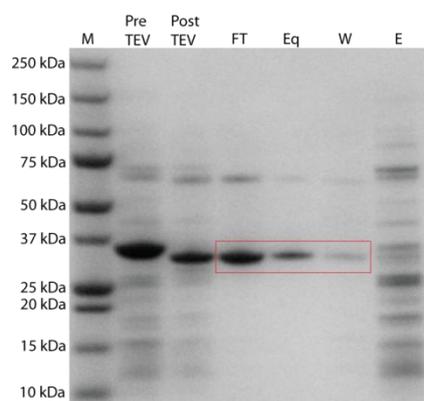
Figure 4.10. SDS PAGE of post His$_6$-tag cleavage.

CaMK1D1-280 was collected in the FT, Eq, and W fractions (highlighted by red box).
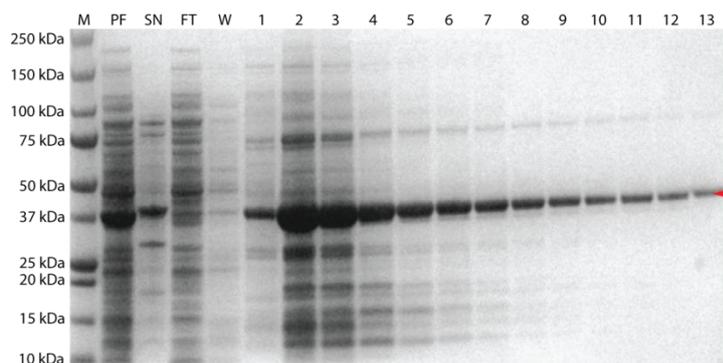
Figure 4.11. SDS PAGE of His$_6$-tagged CaMK1D1-298 from Ni^{2+} NTA affinity purification.

SDS PAGE showing the elution fractions of His$_6$-tagged CaMK1D$_1$-298 (indicated by red arrow) using 250 mM imidazole. Although not shown the results for CaMK1D_T180E_1-298 are similar to this.
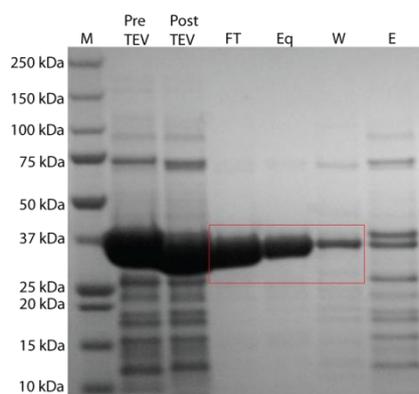
Figure 4.12. SDS PAGE of-CaMK1D1-298 post His$_6$-tag cleavage.

CaMK1D1-298 was collected in the FT, Eq, and W fractions.

Size exclusion purification of CaMK1D C-terminal truncations

C-terminal truncated CaMK1D constructs purifiy as monomeric proteins by size exclusion chromatography

Although the majority of the contaminant proteins were removed via the second Ni^{2+} NTA-affinity purification step, the purity of CaMK1D1-280 and CaMK1D1-298 seen by SDS PAGE was still < 95 % and aggregation in the samples was not eliminated by this technique. Therefore size exclusion purification chromatography was necessary in order to remove the remaining contaminants and aggregates that co-eluted with CaMK1D1-280 and CaMK1D1-298 from the second Ni2+ NTA-affinity purification step. The elution profiles of CaMK1D1-280 and CaMK1D1-298 both produce a single peak suggesting that these proteins behave as monomers in solution of 31.8 and 33.8 kDa when compared to the elution profile of protein standards (Figure 90, 92). The elution volume of the 31.8 kDa CaMK1D1-280 and the 33.8 kDa CaMK1D1-298 ranges from approximately 165–195 ml and 145-170 ml respectively. Higher molecular weight contaminants and aggregates elute off the size exclusion column before both C-terminal truncation mutants as exemplified in their respective size exclusion elution profiles. SDS PAGE of the CaMK1D1-280 and CaMK1D1-298 size exclusion fractions reveals 90 % and 95 % purity respectively (Figure 91, 93). We hypothesise that both constructs exhibit basal constitutive activity but evidence of His$_6$-tag phosphorylation similar to Pim1 was not observed.
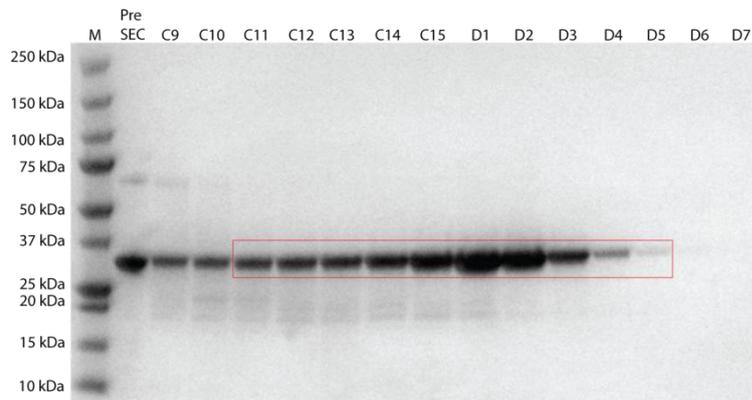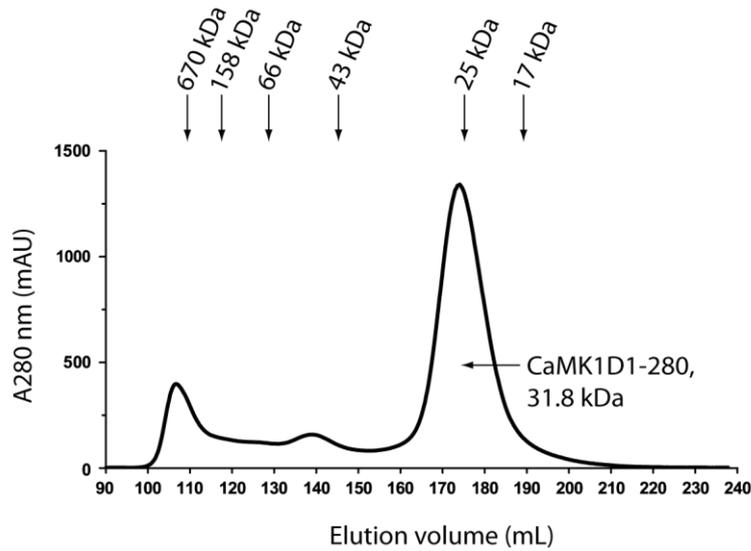
Figure 90. Size exclusion elution profile of CaMK1D1-280 and SDS PAGE. The 31.8 kDa CaMK1D1-280 elutes from approximately 165–190 ml. CaMK1D_T180E_1-280 also elutes from a similar position. CaMK1D1-280 appears 90 % pure after size exclusion purification. Elution fractions shown correspond to the fractions under the CaMK1D1-280 size exclusion peak. PSE = pre size exclusion CaMK1D1-280. The T180E variant of this truncation purifies in a similar manner.
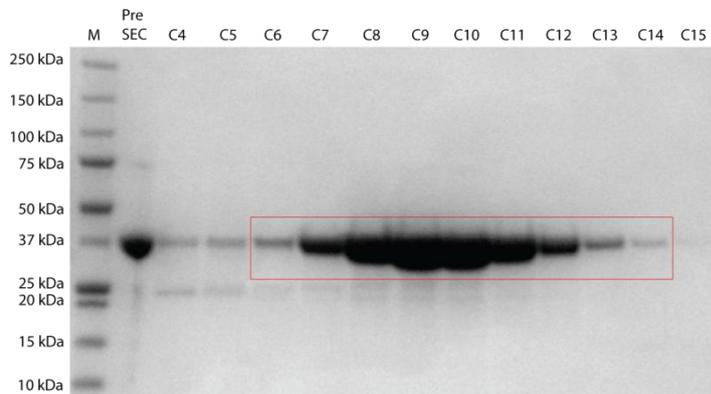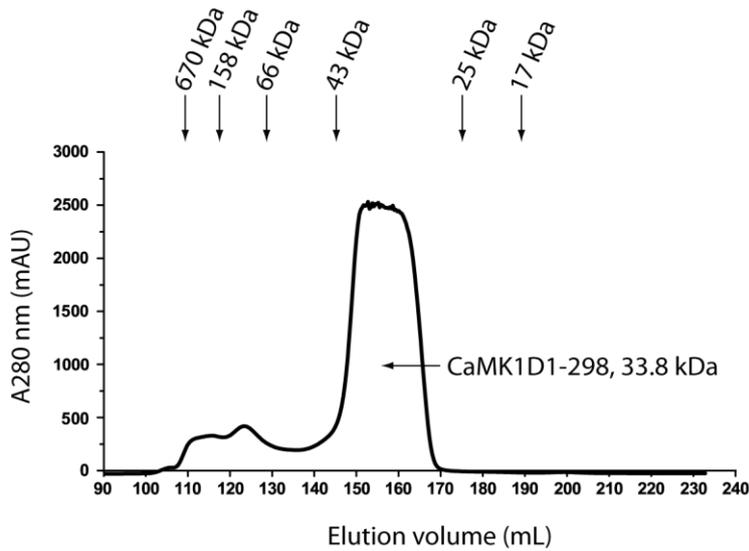
Figure 92. Size exclusion elution profile of CaMK1D1-298 and SDS PAGE. The 33.8 kDa CaMK1D1-298 elutes from approximately 145–170 ml. CaMK1D_T180E_1-298 also elutes from a similar position. CaMK1D1-298 appears 95 % pure after size exclusion purification. Elution fractions shown correspond to the fractions under the CaMK1D1-298 size exclusion peak. PSE = pre size exclusion CaMK1D1-298/CaMK1D_T180E_1-298.

CaM Expression

His$_6$-CaM is expressed as a stable and soluble protein at 37C

Expression and purification of His$_6$CaM was initiated in order to produce the active CaMK1D•Ca^{2+}/CaM complex for structural and functional characterisation. In order to optimise the yield of soluble stable protein, the expression of CaM afterr different post induction times was monitored by SDS PAGE. CaM is a thermal stable protein which allowed expression tests in LB to be conducted at a pre and post induction temperature of 37 °C. Analysis of the His$_6$-tagged CaM post induction times of 4, 8, and 16 hours by SDS PAGE (Figure 94) indicated that over expression of CaM increased with time and was greatest after 16 hours.
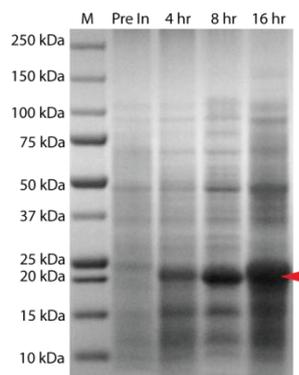
Figure 94. SDS PAGE showing His$_6$-tagged CaM expression test in LB growth media. Expression of CaM increases with time, appearing greatest after 16 hours post induction. His$_6$-tagged CaM appears to migrate with a molecular weight of around 20 kDa.

SDS PAGE abbreviations

M = marker SN = supernatant, PF = pellet fraction, FT = flow through, Eq = equilibration wash, W = wash, E = elution, PHF = pooled his trap fractions, ON = overnight, PSE = pre size exclusion, PAE = pre anion exchange.

Ni^{2+} NTA-affinity purification of CaM

CaM co-elutes with minor contaminants by Ni^{2+} NTA affinity

In order to initiate binding studies of CaMK1D-Ca^{2+}/CaM, pure CaM protein was needed. Therefore His$_6$-tagged CaM was purified in the same manner as described previously for all CaMK1D constructs using a Ni^{2+} NTA-affinity column. Analysis by SDS PAGE reveals that His$_6$-tagged CaM also co-elutes with some contaminants (figure) but these were removed by further purification steps described below. His$_6$-tagged CaM was also found in the wash (W) fraction due to saturation of the Ni^{2+} NTA-affinity column with His$_6$-tagged CaM. It is also evident from SDS PAGE (Figure 95) that the His$_6$-tagged CaM was completely soluble because His$_6$-tagged CaM was not present in the insoluble pellet fraction (PF).

Cleavage of the N-terminal His$_6$-tag and a second Ni^{2+} NTA-affinity purification improved the purification of CaM

His$_6$-tagged CaM was further purified from the contaminant proteins it co-eluted with by removal of the CaM His$_6$-tag and imidazole mediated by TEV protease and dialysis respectively. A second Ni^{2+} NTA-affinity followed these steps and analysis of the fractions on SDS PAGE (Figure 96) shows that after incubation with TEV protease CaM (16.8 kDa) migrates as a slightly lower molecular weight protein relative to its His$_6$-tagged form (19.4 kDa) thus indicating successful cleavage of the His$_6$-tag. SDS PAGE also indicates that CaM appears in the flow through (FT), equilibration (Eq), and wash (W) fractions suggesting that it binds weakly to the column whereas the contaminanting proteins bind with greater affinity hence their appearence in the elution fraction.
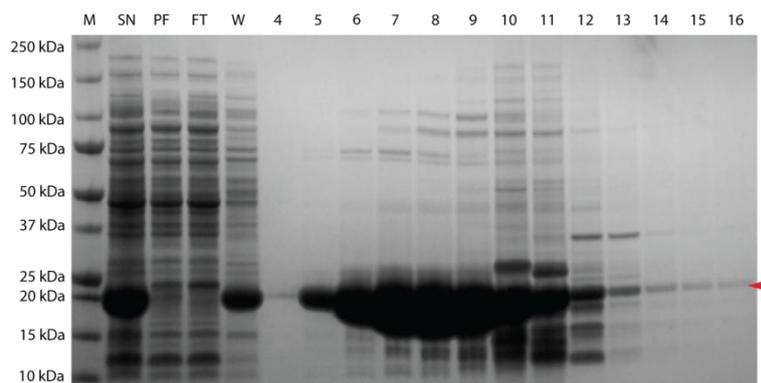
Figure 95. SDS PAGE of His$_6$-tagged CaM purified by Ni^{2+} NTA affinity purification. Elution fractions are shown containing His$_6$-tagged CaM which appears to have a molecular weight around 20 kDa. His$_6$-tagged CaM was eluted using a linear imidazole gradient.
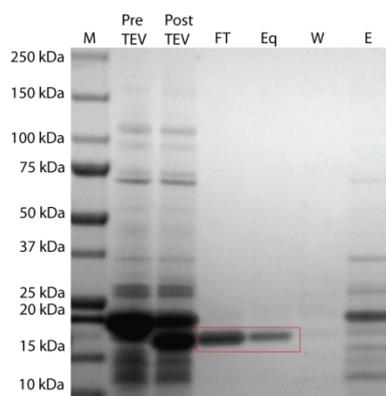
Figure 96. SDS PAGE showing TEV protease mediated removal of the His$_6$-tag from His$_6$-CaM. CaM was collected in the FT and Eq fractions.

Size exclusion purification of CaM

CaM Size exclusion chromatography purification suggests a monomeric protein

Minor contaminants and imidazole present in the pooled flow through, equilibration, and wash fractions of CaM from the second Ni^{2+} NTA-affinity purification step were removed by size exclusion chromatography. The elution profile of CaM produces a single peak suggesting that the protein behaves as a monomer in solution (Figure 97). The peak has a low absorbance at 280nM because CaM does not contain any Trp residues and the elution volume of the 16.8 kDa CaM appears to be from 170-200 ml. High molecular weight contaminants elute before CaM. Analysis of the CaM fractions post size exclusion chromatography by SDS PAGE reveals 90 % purity (Figure 98). Minor contaminants are evident in the first few fractions and were removed by a final purification step involving anion exchange.

Anion exchange purification of CaM

CaM purifies to 95 % purity after anion exchange chromatography

Analysis of the CaM fractions from anion exchange purification by SDS PAGE illustrates 95 % pure protein after the first 7 elution fractions (Figure 99). Activation of CaMK1D using Ca^{2+}/CaM in our activity assays confirmed that the purified CaM was folded and functional. The first 7 fractions were pooled and purified again by anion exchange to remove the minor contaminants.
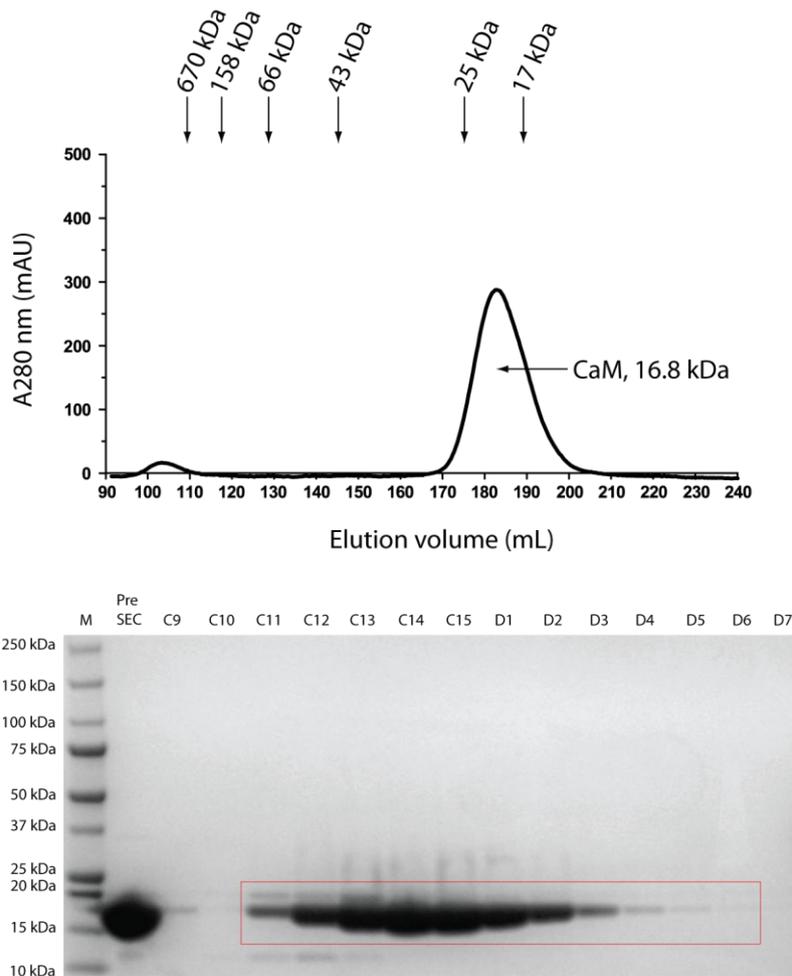
Figure 97. CaM size exclusion elution profile and SDS PAGE. The 16.8 kDa CaM purifies as a single peak by size exclusion purification (S75 column). Its elution from 170–200 ml corresponds to a 16.8 kDa species based on comparison with the elution of known protein standards. Size exclusion therefore suggests that CaM is monomeric. Minor contaminants are evident in the first few fractions shown by SDS PAGE. PSE = pre size exclusion CaM
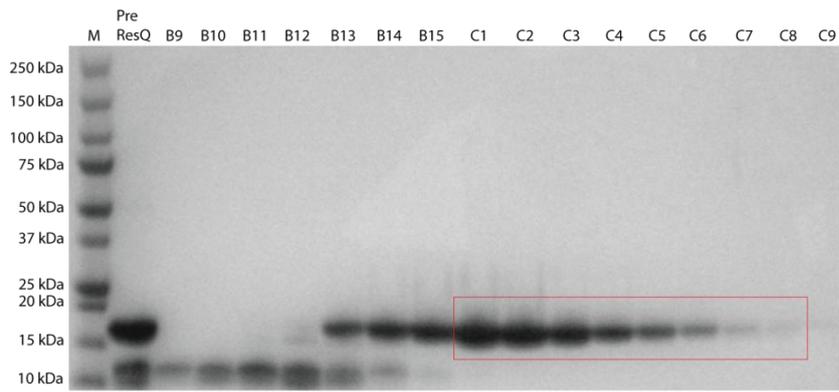
Figure 99. SDS PAGE of CaM anion exchange (resource Q) purification. CaM was purified to 95% purity by pooling the fractions indicated by the red box indicates pooled CaM fraction PAE = pre anion exchange CaM.

Deuterated CaMK1D and TROSY

^{2}H, ^{15}N CaMK1D is folded after amide ^{2}H exchange for ^{1}H and shows complete exchange

To begin characterisation of the CaMK1D solution structure and subsequent binding studies the next aim was to produce a ^{2}H, ^{15}N labelled CaMK1D sample from M9 minimal media containing 99 % ^{2}H$_2$O in order to acquire a higher resolution ^{1}H, ^{15}N TROSY-HSQC. A preliminary ^{1}H, ^{15}N TROSY-HSQC of ^{2}H, ^{15}N labelled CaMK1D post back exchange was acquired to evaluate the spectral improvements and feasibility of assigning the CaMK1D backbone using a deuterated sample. Acquiring the spectrum also helped to confirm the extent of back exchange because ^{2}H labelling introduces ^{2}H at exchangeable amide groups which need to be exchanged with ^{1}H in order to detect the ^{1}H amide resonances. Firstly the spectrum shows dispersion of backbone amide ^{1}H resonances thus confirming that the protein was folded as was implied by the 1D ^{1}H spectrum. Secondly assessment of the amide ^{1}H resonance line shape and signal to noise ratio reveals that there was an improvement when compared with the ^{1}H, ^{15}N HSQC of the non deuterated ^{15}N labelled CaMK1D sample. Thirdly the spectrum illustrates that the extent of amide ^{2}H exchange with ^{1}H for backbone and side chain amides was complete because an overlay of the ^{1}H, ^{15}N TROSY-HSQC spectra for ^{2}H, ^{15}N CaMK1D and ^{1}H, ^{15}N CaMK1D (Figure 101) indicates the same number of peaks (observable amide ^{1}H resonances). This suggests that back exchange was complete. Having the same number of peaks was important to maximise the number peak that can potentially be assigned. Furthermore all backbone triple resonance experiments are correlated to the amides seen in the ^{1}H, ^{15}N TROSY-HSQC spectrum therefore allowing the assignment of each amide resonance.

In some cases loss of peaks (amide ^{1}H resonance) in a deuterated sample can occur when exchange of amide ^{2}H for ^{1}H is not complete because amide ^{2}H resonances are not observable in ^{1}H, ^{15}N HSQC experiments but in our case solvent exchange process was effectively complete. The complete exchange of amide ^{2}H for ^{1}H usually requires unfolding and refolding the protein in order to facilitate the

exchange of deeply buried and highly stable amides. However, as already described this was avoided in our case because the alternative method of buffer exchanging into high pH buffer (or low pH for some proteins stable at acidic pH) was found to promote exchange effectively.
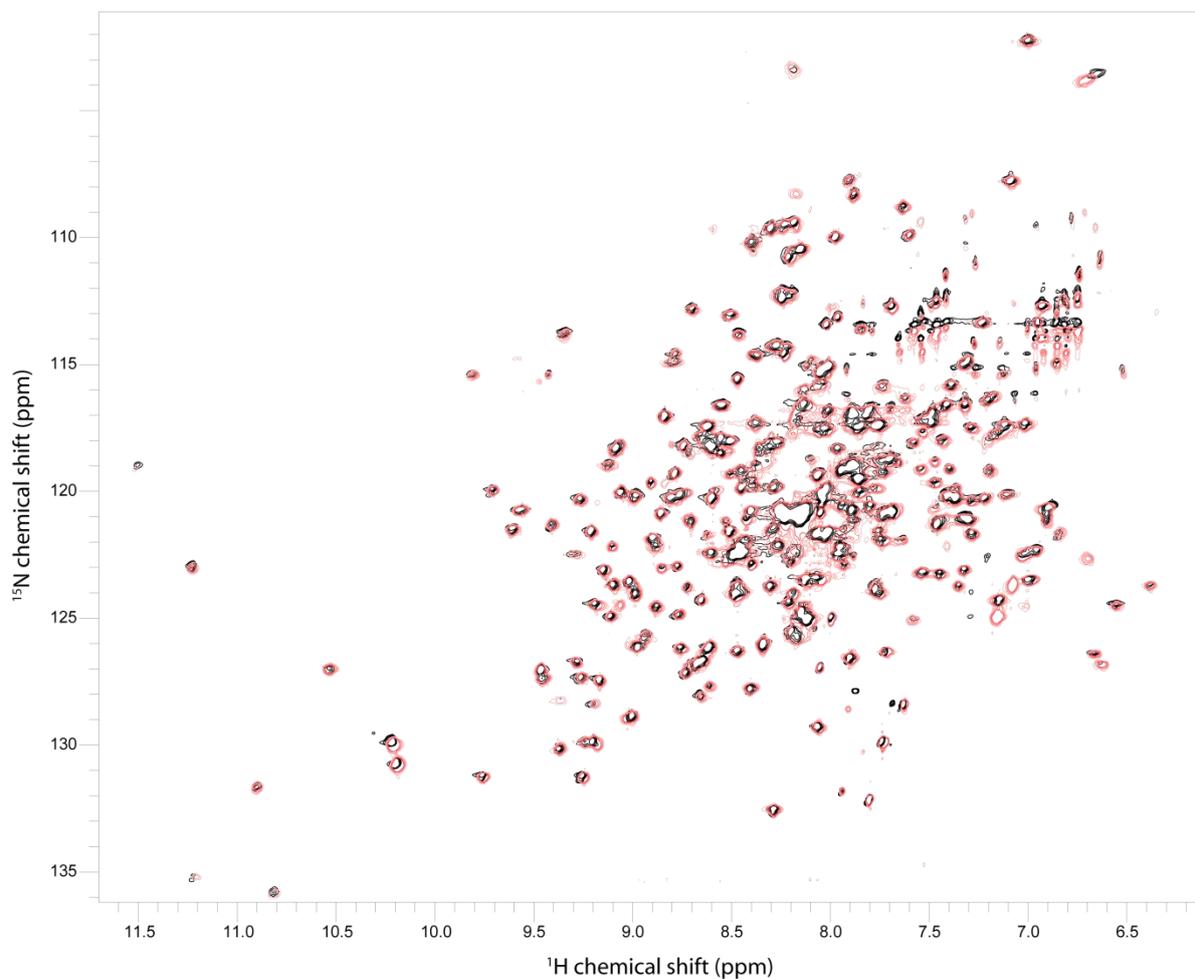
Figure 101. ^{1}H, ^{15}N TROSY-HSQC spectra of a ^{2}H, ^{15}N and ^{1}H, ^{15}N CaMK1D sample. Superimposition of the ^{1}H, ^{15}N TROSY-HSQC for ^{2}H, ^{15}N (red) and ^{1}H, ^{15}N CaMK1D (black) indicates that deuteration improves the resonance intensity of the ^{2}H, ^{15}N CaMK1D spectrum. Furthermore, the peak count of both spectra appears identical suggesting complete exchange of amide ^{2}H with ^{1}H in deuterated CaMK1D. Both spectra were acquired with 0.2 mM in 50 mM Na phosphate pH 7.0, 75 mM NaCl, 0.5 mM TCEP, Na Azide 0.02 %, 10 % ^{2}H$_2$O at 25 °C on a Varian 800 MHZ.

CaMK1D backbone assignment

Sequential assignment of CaMK1D backbone amide resonances using triple resonance experiments

To initiate the sequential backbone assignment of CaMK1D by NMR, a ^{2}H, ^{15}N, ^{13}C labelled CaMK1D sample was produced. TROSY versions of the triple resonance experiments HNCO, HN(CA)CO, HN(CO)CA, HNCA, HNCACB, and HN(COCA)CB were acquired on a Varian 800MHz equipped with a cryo-probe. Backbone amide directed sequential assignments were achieved by establishing the connectivities of carbonyl (CO), α (CA) and β (CB) carbon resonances for residues i and i-1 using the relevant triple resonance experiments. For a detailed explanation of the triple resonance experiments discussed here please refer to chapter II.

Firstly we used the CA resonance information provided by HNCA and HN(CO)CA experiments to build connectivities between residues i and i-1. In addition to this CB resonance information provided by the HNCACB and HN(COCA)CB was used to validate the connectivities between residue i and i-1. Importantly unique CA and CB resonances characteristic of a certain residue type were used to help identify residue types when possible. For example the CA resonances of Gly residues are unique and readily identifiable. Additional validation of connectivities between residue i and i-1 was achieved via CO resonance information provided by the HN(CA)CO and HNCO. By applying this strategy sequential unambiguous assignments were accomplished. For example figures 102 A, B, C, D, illustrates the sequential assignment of residues 21-26 corresponding to the sequence KIFEFK in CaMK1D based on CA, CB and CO resonance connectivities.
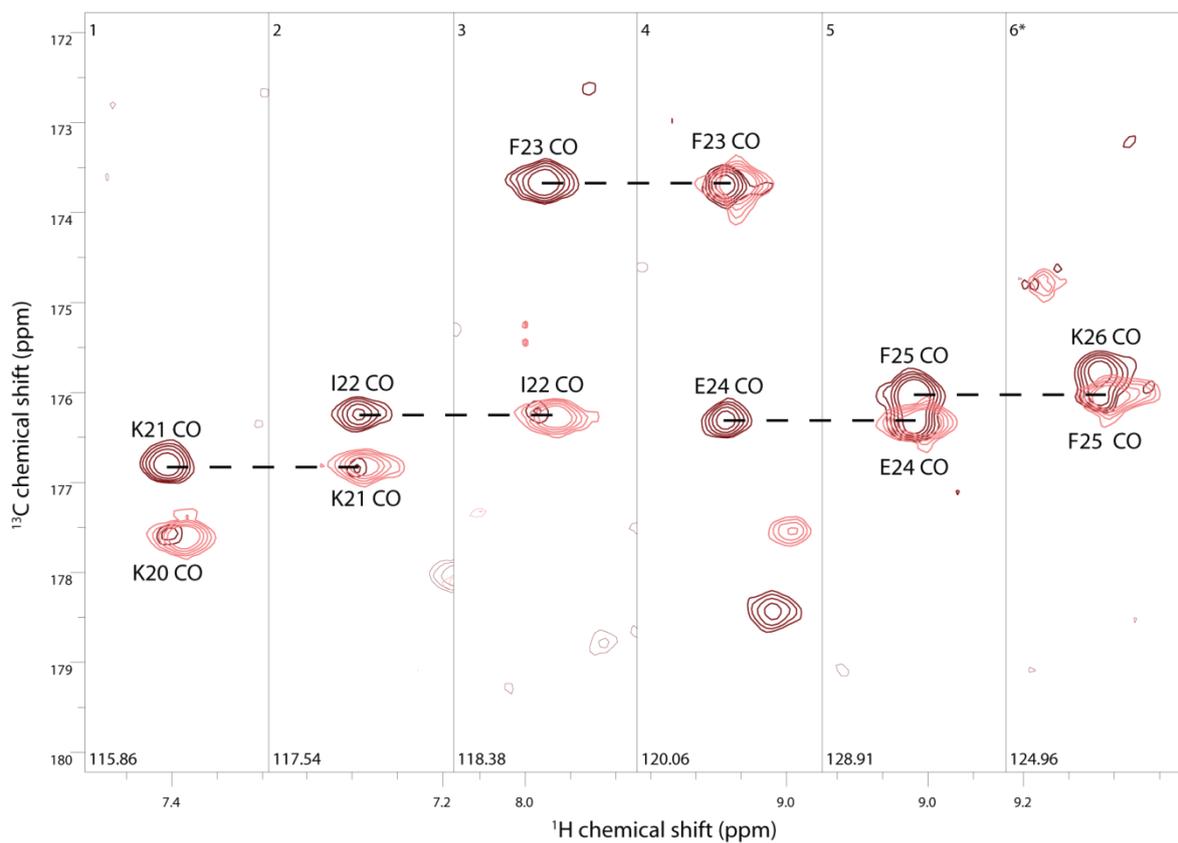
Figure 102 A. Sequential backbone assignment of CaMK1D using HNCO and HN(CA)CO spectra. Connectivity between residue i and i-1 carbonyl carbons were established were using HNCO (brown) and HNCACO spectra (pink). Sequential backbone assignments of carbonyl carbon resonances for CaMK1D residues KIFEFK (residues 21-26) are shown.
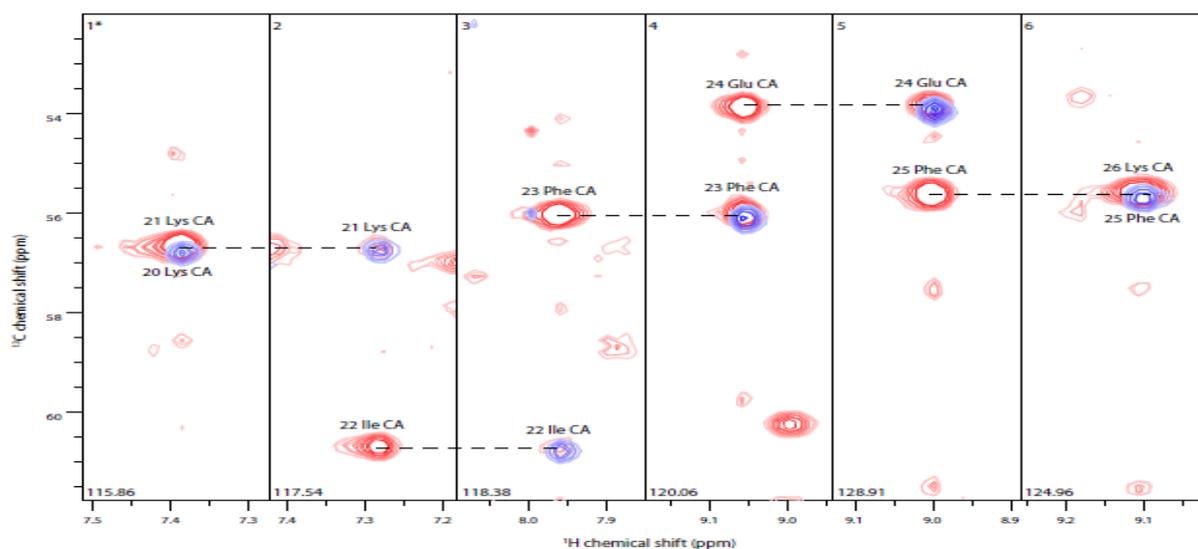
Figure 102 B. Sequential backbone assignment of CaMK1D using HN(CO)CA and HNCA spectra. Connectivity between residue i and i-1 α carbons were established were using HN(CO)CA (blue) and HNCA spectra (red). Sequential backbone assignments of α carbon resonances for CaMK1D residues KIFEFK (residues 21-26) are shown.
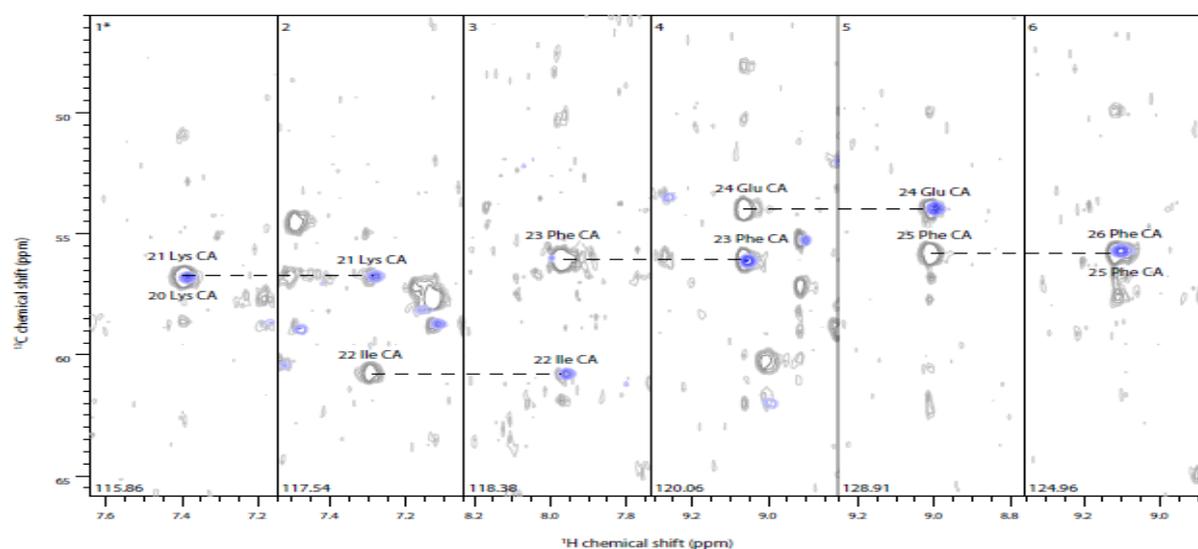
Figure 102 C. Sequential backbone assignment of CaMK1D using HN(CO)CA and HNCACB spectra. Connectivity between residue i and i-1 α carbons were established were using HN(CO)CA (blue) and HNCACB spectra (blue). Sequential backbone assignments of α carbon resonances for CaMK1D residues KIFEFK (residues 21-26) are shown.

Figure 102 D. Sequential backbone assignment of CaMK1D using HN(COCA)CB and HNCACB spectra. Connectivity between residue i and i-1 β carbons were established were using HN(COCA)CB (purple) and HNCACB spectra (orange). Sequential backbone assignments of β carbon resonances for CaMK1D residues KIFEFK (residues 21-26) are shown.

79% of the CaMK1D backbone was assigned

The ^1H, ^{15}N-TROSY HSQC of the 37.8 kDa deuterated CaMK1D protein in its inactive state was acquired as and assigned as shown in figure 103. The backbone amide resonances of 255 residues were assigned out of a possible 322, representing 79% coverage, including residues from every structural and functional element (Figure 104). Of those amino acids assigned, 98% include C', 100% include Cα, and 99% include Cβ chemical shifts. Of all 11 proline residues, 10 were assigned, with 80% including C', 60% including Cα and 40% including Cβ chemical shifts. The remaining Pro219 could not be assigned as it is N-terminal to another proline. Residues which could not be assigned predominantly due to intermediate exchange and resonance broadening were 1, 6, 9, 34-35, 56, 63-66, 68-70, 74-77, 82, 84-85, 99-102, 141, 143-144, 146, 164-169, 172, 174, 178-182, 186-187, 199-200, 212-213, 216, 251-254, 260, 303-305, 317-325 and 327. The unassigned residues are mostly found in regions that include for example the activation loop, catalytic loop, αC helix, hinge, CBD and C-terminus (Figure 104).

Incomplete assignments in the activation loop and C-terminus also coincided with the disorder and high B-factor values seen for these respective regions in the CaMK1D PDB: 2JC6 (Figure 58). Resonance assignments in these typically exposed and mobile elements were presumably compromised by resonance broadening caused by intermediate or slow time scale dynamics. Assignments were also complicated by resonance overlap near the centre of the ^1H, ^{15}N TROSY-HSQC spectrum (Figure 103). Several residues exhibited a second set of weaker or equivalent resonances, including for example Gly103, Gly217 and Asp223 suggesting additional conformers (Figure, green boxes). Isolated residue assignments were made for Glu67, Ala83, His142, Leu145 and Tyr 198 based on their unique chemical shifts, intra- and inter-residue correlations in the triple resonance spectra, and the CaMK1D structure.
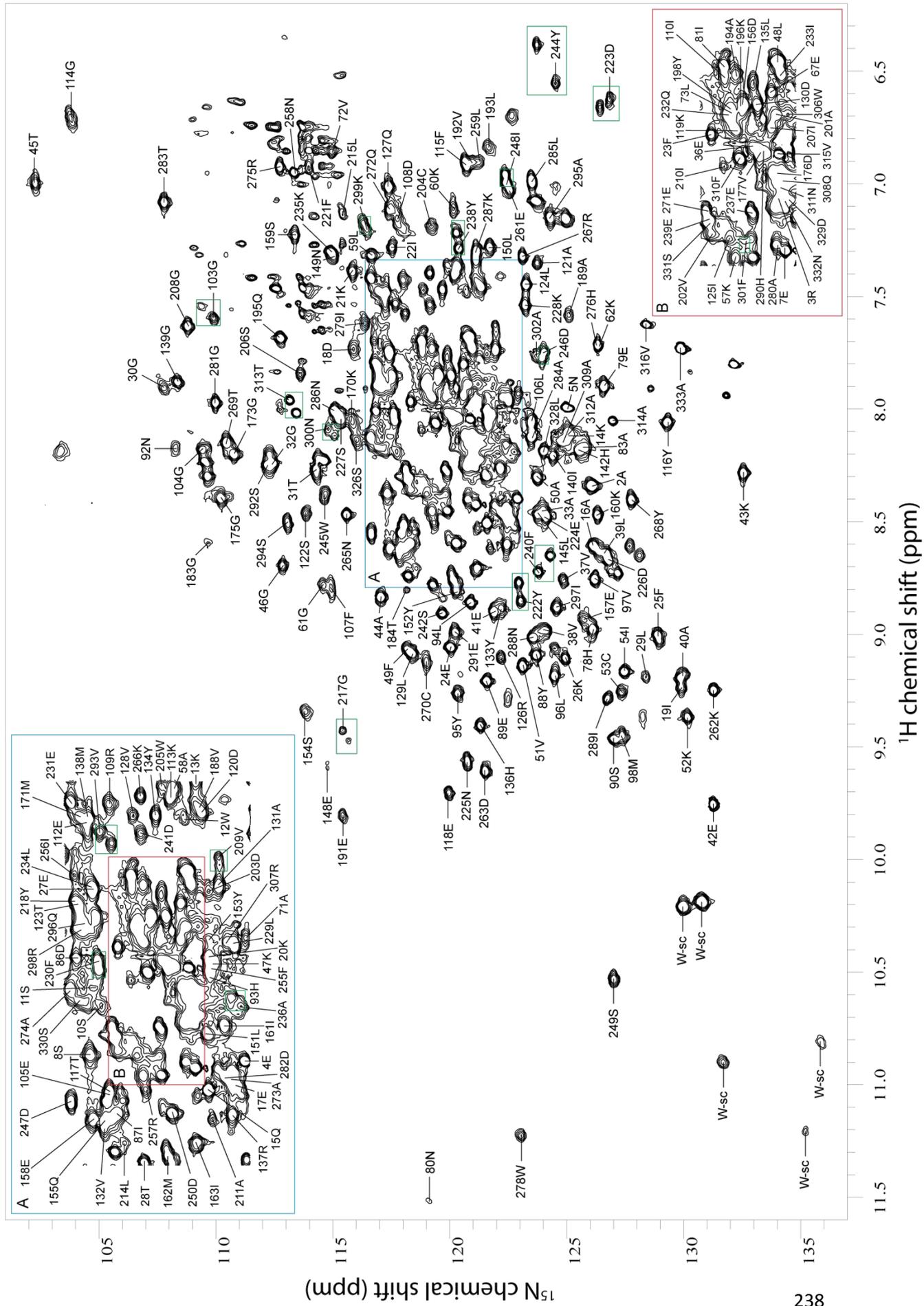
Figure 103. ^{1}H, ^{15}N TROSY-HSQC of ^{2}H ^{15}N ^{13}C-labelled CaMK1D. Approximately 80% of the 37.8kDa CaMK1D backbone was assigned using a deuterated and uniformly ^{15}N, ^{13}C-labelled sample (1.2 mM). Assignments are indicated by labelled resonances. The densely overlapped region in the centre of the spectrum is expanded by the blue box (denoted as box A). The centre of this region is further expanded by the red box (denoted as box B). A few resonances demonstrate multiple conformations in solution and some of these are highlighted by a green box. The spectrum was acquired on a 1.2 mM sample in 50 mM Na Phosphate pH 7.0, 75 mM NaCl, 0,5mM TCEP, 0.02% Na Azide at 25 °C on a Varian 800 MHZ.
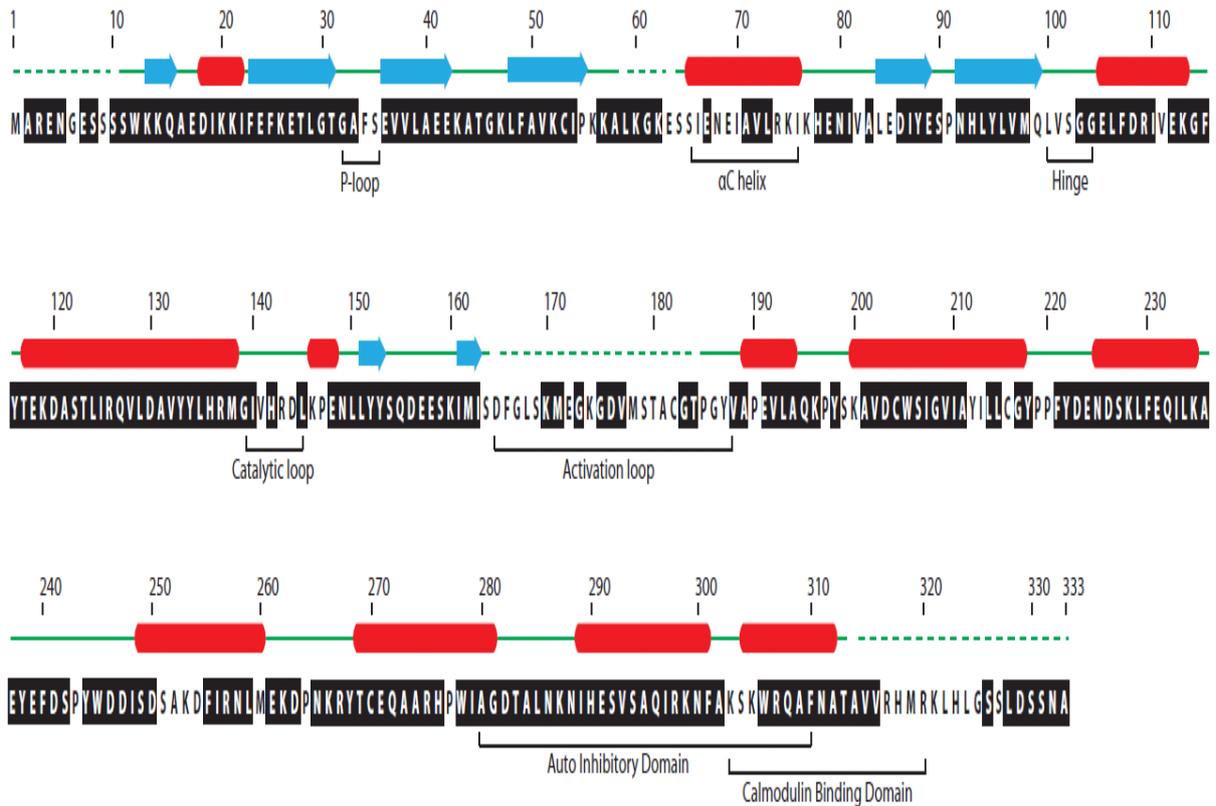
Figure 104. CaMK1D backbone assignment coverage. This schematic illustrates the extent of the backbone assignment for CaMK1D. Assigned residues in the CaMK1D sequence are highlighted in black whilst unassigned are not highlighted. The secondary structure cartoon was derived from the CaMK1D PDB: 2JC6. Regions of disorder in which the electron density could not be determined in the crystal structure are indicated by broken green lines. In total 79 % of the CaMK1D backbone was assigned using triple resonance spectra.

CaMK1D secondary structure predictions

CaMK1D secondary structure predictions beased on TALOS exhibits differences when compared to the crystal structure

With 79 % of the CaMK1D backbone amides assigned a secondary structure prediction was made using TALOS, a program which requires input of the chemical shift assignments for the backbone amides to calculate predicted dihedral angles phi and psi. Secondary structure elements predicted by TALOS can be shown as a probability (Figure 105). A probability described by a positive value predicts α-helix whereas a negative value predicts a β-strand and a probability close to 0 predicts neither thus suggesting a loop or disordered region (Shen, Delaglio et al. 2009).

A comparison of our CaMK1D secondary structure prediction by TALOS with the secondary structure defined by the CaMK1D crystal structure reveals some differences that are notable in regions that span both the N-terminal and C-terminal lobes. At the N-terminus of the sequence TALOS predicts 3 β-strands (β1, 2, 3) prior to the αC helix whereas the crystal structure defines 4 β-strands (β1, 2, 3, 4) and 1 α-helix (αB) between the first β-strand (β1) and second (β2) strand. TALOS also predicts that the start of the αC helix begins immediately after the third β-strand (β3) in the TALOS prediction although many assignments were incomplete for the αC helix region and therefore predictions for this region are interpreted with caution. The crystal structure defines the αC helix as starting after a loop that is C-terminal of β-strand (β4). The fifth α-helix (αF) in the crystal structure is absent in the TALOs prediction and at the C-terminal end the ninth (αJ) and tenth (αK) α-helices of the crystal structure are predicted to be smaller by TALOS. In addition to this the last helix (αM) in the crystal structure is missing in the TALOS prediction where it is instead predicted to be unstructured. Interestingly this last helix (αM) constitutes the latter part of the AID that also contains the overlapping CBD.

Another secondary structure prediction method we used for comparison was based on ΔCα-ΔCβ which utilises the chemical shift difference method between measured

experimental values and random coil values of α and β carbon resonances (Metzler 1993). A comparison between folded and random coil chemical shifts illustrates that α carbon resonances exhibit a propensity to shift up-field in β-strand and β-sheets, but shift downfield in α-helices. β carbon resonances exhibit an opposite pattern with chemical shifts that have a propensity to shift downfield in β-strand and β-sheets but up-field in α-helices. A subtraction between Cα and Cβ shifts will enhance the correlation between the Cα, Cβ shifts and secondary structure elements. Our ΔCα-ΔCβ data predicts the secondary structure of CaMK1D to be relatively similar to our TALOS derived secondary structure predictions because the α helix and β strand predictions occur in approximately the same positions but the boundaries of secondary structure elements predicted by TALOS are more clearly defined.
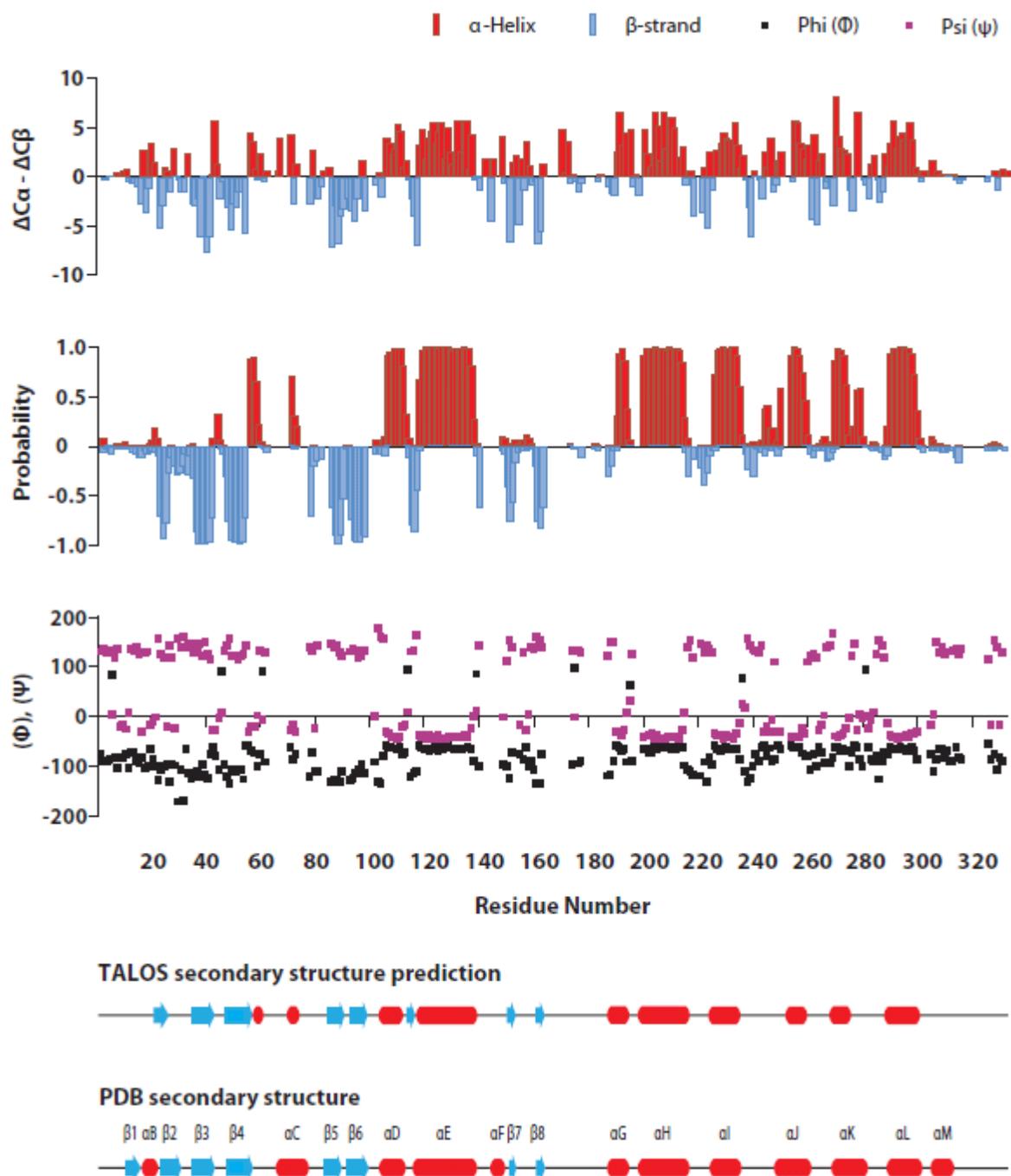
Figure 105. A summary of CaMK1D Secondary structure prediction and comparison with the PDB secondary structure. Figure A plots ΔCα-ΔCβ and predicts the propensity of the backbone to adopt an α helix (red) if > 0 or β strand (blue) if < 0. Figure B represents the propensity for α helix (red) if > 0 and β strand (blue) if < 0 as predicted by TALOS and is plotted as a probability. Figure C, is a plot of the phi (black) and psi (purple) dihedral angles calculated by TALOS. Figure D compares the TALOS derived secondary structure with that of the CaMK1D crystal structure and indicates that there are some differences.

CaMK1D pH screening

Assignment of backbone amides could not be extended beyond 79 % coverage due to weak or unobservable α and β carbon resonance intensities for the few remaining unassigned backbone amides in the spectrum. A lack of connectivities between the carbonyl, α and β carbon resonances also suggested that the number of backbone amide resonances counted in the spectrum was less than the total expected number. These missing backbone amides resonances thus presented a major issue in achieving a complete backbone assignment.

CaMK1D exhibits pH dependent CSPs

In order to improve the intensity and increase the number of peaks (observable backbone amide resonances) in the ^{1}H, ^{15}N TROSY-HSQC the pH of the buffer was lowered gradually from 7.0 to 6.7, 6.5, and 6.0 (Figure 106) because at lower pH the rate of the amide ^{1}H chemically exchanging with the bulk solvent is reduced therefore the lifetime of the protonated amide should be prolonged and consequently increase the intensity of weak and unobserved peaks. However additional peaks were not observed but the intensity of some peaks increased thus improving the signal to noise ratio of these resonances. At pH 6.0 CaMK1D also had a tendency to precipitate more after acquiring the ^{1}H, ^{15}N TROSY-HSQC at 25 °C most likely because CaMK1D has a predicted pI of 6.1.

Furthermore a few notable pH dependent CSPs were observed. Gly 114, Phe 115, Tyr 116, Ser 292, Ser 294, and Ala 295 in particular exhibited large chemical shift perturbations. Activity assays were also conducted to investigate the pH effect on CaMK1D activity. We found that activity decreased with lower pH and postulate that it the CaMK1D is likely becoming more unstable at lower pH because the pI is predicted to be 6.1. CaMK1D also exhibited a tendency to precipitate during pH screening by NMR when below pH 7.0. Although pH screening was conducted with the autoinhibited form, most of the peaks that display pH dependent CSPs appear in the same position with respect to the spectrum of the activated CaMK1D.
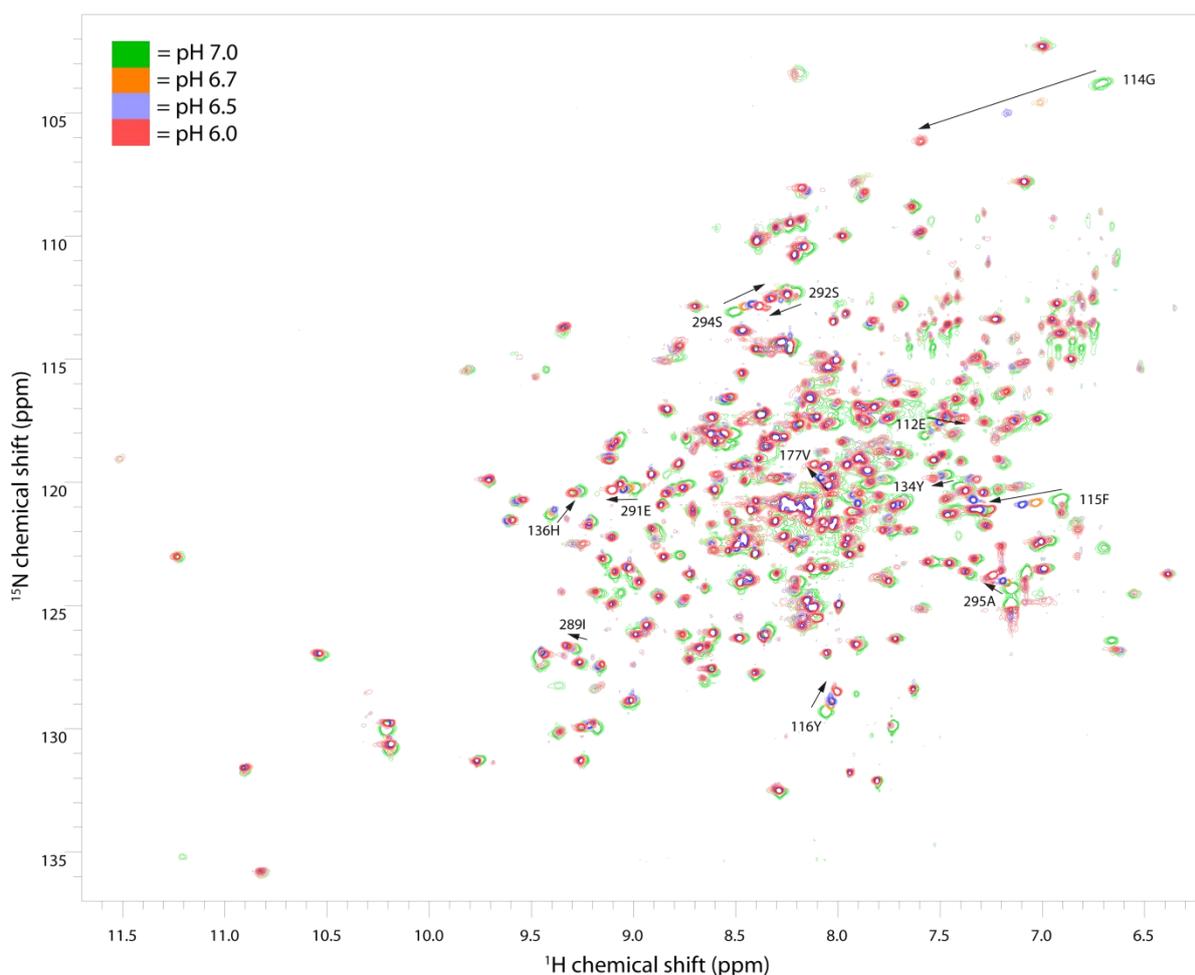
Figure 106. ^{1}H, ^{15}N TROSY HSQC based screening of optimal pH for CaMK1D. The pH of the NMR buffer was lowered from 7.0 to 6.7, 6.5 and 6.0 in order to improve the peak intensity and increase the backbone amide peak count in the ^{1}H, ^{15}N TROSY-HSQC. This was necessary because assignment of backbone amides could not be extended beyond 79 % coverage due to peaks that broadened out or exhibited weak intensity in the H, ^{15}N TROSY-HSQC. However an increase in peak count was not observed but the intensity of some peaks increased. pH dependent CSPs were also evident. Spectra were acquired with 0.4 mM ^{2}H, [15]N CaMK1D in 50 mM HEPES pH 7.5, 6.7, 6.5, and 6.0, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide at 25 °C on a Varian 800 MHZ.

The pH dependent chemical shift perturbations of some residues is due to their proximity to His residues

A histogram of the pH dependent CSPs reveals that there are four main regions significantly perturbed (Figure 107). The first region is near E112, K113, G114, F115, Y116. These residues span the C-terminal end of helix αD and the adjacent connecting loop that is between the αD and αE helices. The histogram also shows that Gly 114 exhibits the largest CSP out of all the significant CSPs ($\geq$ 1 S.D). The second region occurs around H136, which is located towards the C-terminal end of the αE helix. A single CSP occurs in the third region that constitutes the activation loop and this is attributed to V177. The fourth region spans part of the RD where S294 is significantly perturbed.

In order to understand the basis of the pH dependent CSPs we mapped the CSPs onto our apo autoinhibited model of CaMK1D generated by structural alignment with rat CaMK1A and substation of their respective RDs (Figure 108 A,B). Mapping illustrates that E112, K113, G114, F115, and Y116 are in close proximity to the side chain of H290. Likewise S294 is also positioned close to the H290 side chain. L135 and H136 are shown to be near the side chains of H78 and H142. V177 was not observed in the crystal structure because the activation loop was not determined but we postulate that V177 must be spatially close to at least one of three His side chains that reside near the active site cleft because this is the conformational space that the activation loop will typically occupy in protein kinases. The three His side chains located near the active site are H78, H136, and H142 as seen from the structure. Notably H142 is part of the catalytic loop which interacts closely with the activation loop in protein kinases.

We propose that these CSPs occur when the pH is lowered because the His side chain imidazole group has more tendency to become protonated and become positively charged (+1) and therefore the local chemical and electronic environment of His changes due this charge. Consequently the chemical shift of nearby residues is altered due to the positively charged His side chain imidazole group. Some of these nearby residues experienced larger CSPs (i.e. Gly 114) because they were closer in proximity to effects of the charge. The pKa of the imidazole group in an isolated His is 6.8, although the His residues in CaMK1D are not isolated we

presume that some of them will have a side chain pKa that is very similar to an isolated His. H290 for example induces large changes in the chemical shift of G114 when small changes in pH are made from pH 7.0. This would suggest that the pKa of H290 is around 6.8 because when the pH is close to the pKa then small changes in pH will cause large changes in the population of protonated and deprotonated forms of the His side chain imidazole group.
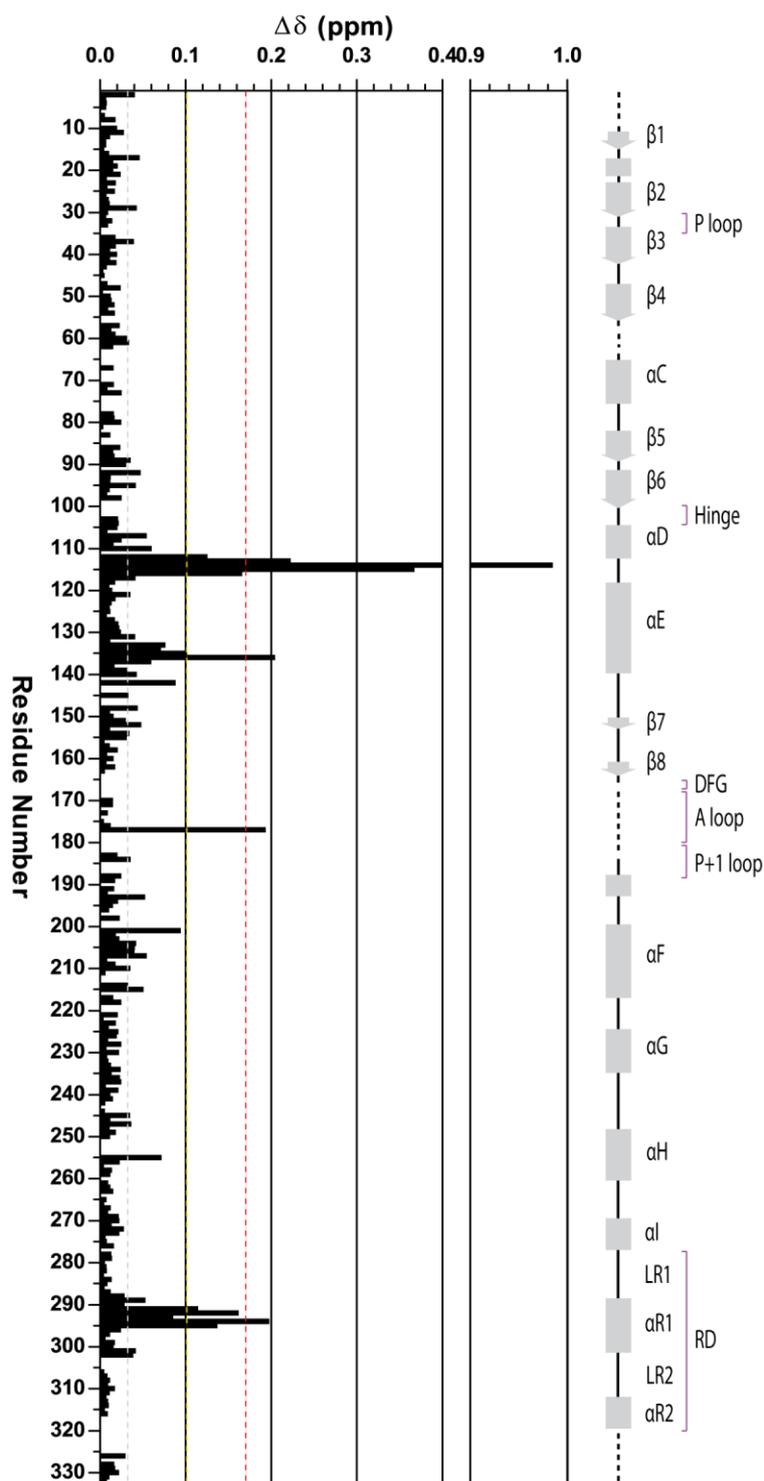
Figure 107. Histogram of pH dependent CSPs along the CaMK1D backbone. The average CSPs is shown indicated by the gray line. CSPs that are ≥ 1 or 2 S.D are indicated by the yellow and red line respectively.
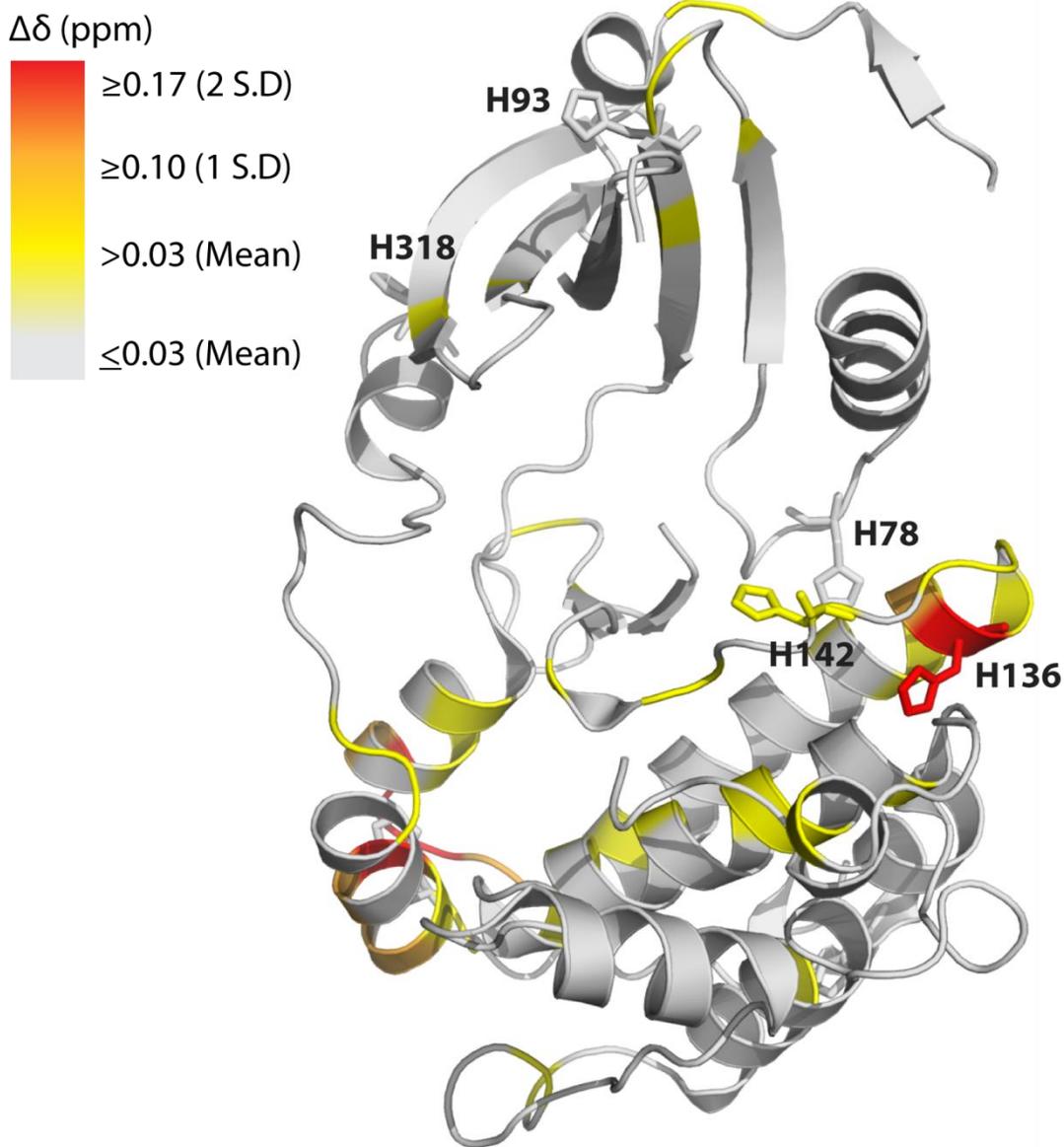
Figure 108 A. CaMK1D pH dependent CSP mapping. A cartoon representation of the apo autoinhibited CaMK1D model structure with residues colored according to the extent of CSP due to a change in the buffer pH from 7.0 to 6.0.
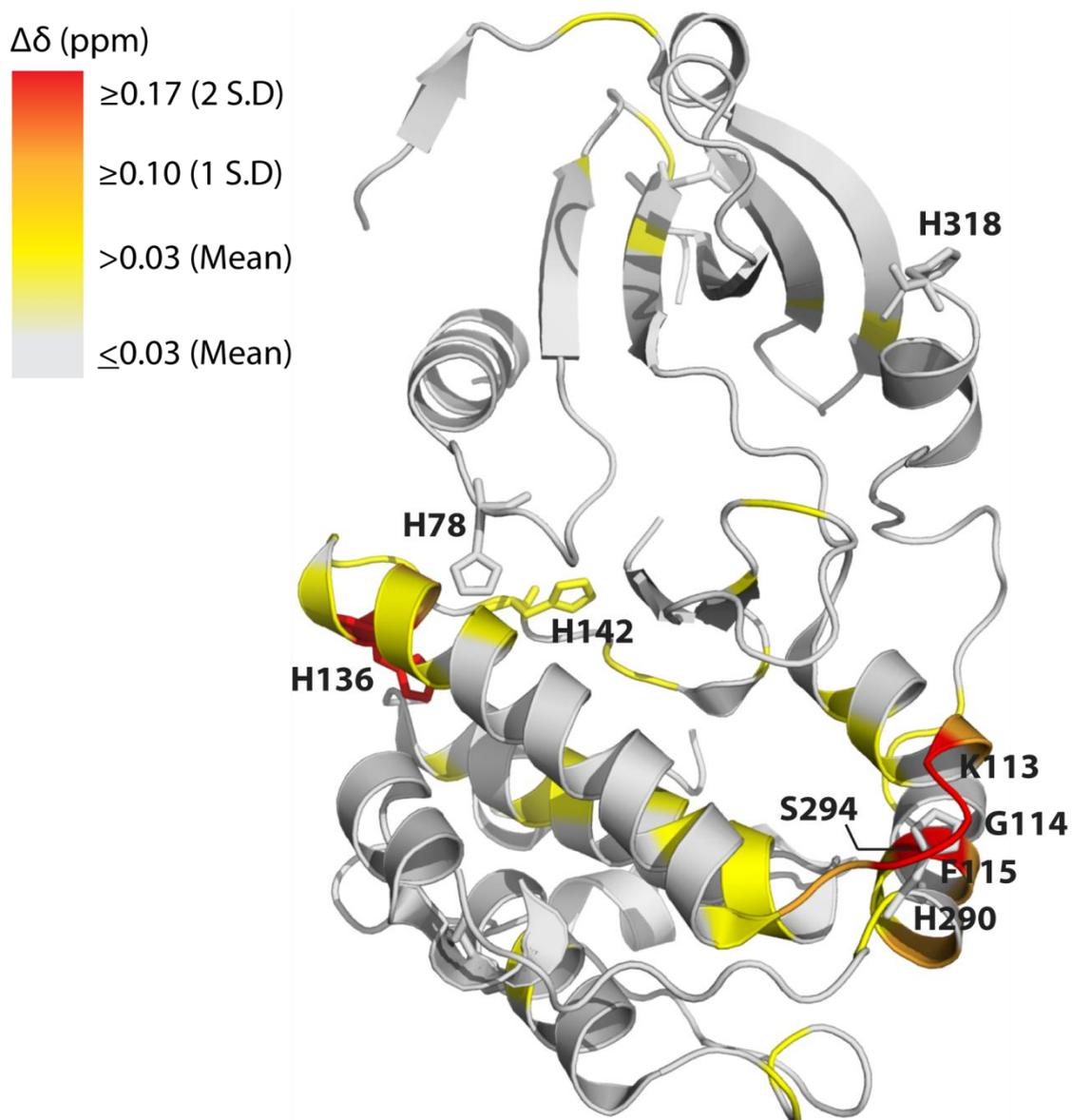
Figure 108 B. CaMK1D pH dependent CSP mapping A 180° rotated view of the previous figure illustrating a cartoon representation of the apo autoinhibited CaMK1D model structure with residues colored according to the extent of CSP due to a change in the buffer pH from 7.0 to 6.0.

CaMK1D crystal structure

CaMK1D-GSK-3 XIII co-crystallised as a dimer and is structurally similar to rat CaMK1A

Firist human CaMKI isoform structure Our collaborators the SGC (knapp et al) have co-crystallised the CaMK1D construct with the ATP competitive inhibitor Glycogen synthase kinase-3 (GSK-3) XIII inhibitor ((5-Methyl-1H-pyrazol-3-yl)-(2-phenylquinazolin-4-yl) amine) and thus determined its structure by X-ray crystallography at a resolution of 2.3 angstroms. The structure revealed that CaMK1D crystallised as a dimer (figure 113). The tertiary structure of CaMK1D adopts a canonical kinase bilobal fold consisting of a small N-terminal β sheet ATP binding lobe and a large C-terminal α helical substrate binding lobe that includes a regulatory AID and CBD. The GSK-3 XIII inhibitor is shown bound to the ATP binding site of the inactive CaMK1D. Functional elements such as the hinge, p-loop, catalytic loop and regulatory αC helix are well resolved in the structure. The AID is shown consisting of 2 helices (denoted as αR1 helix and αR2 helix) connected by a short loop with the N-terminal portion of the AID comprised of αR1 helix folding over the substrate binding site whilst making multiple contacts that are mediated by a pseudosubstrate mechanism.

The αR2 helix which comprises the C-terminal portion of the AID and the overlapping CBD makes contact with the small N-terminal lobe but residues downstream from Thr 313 are not evident in the structure presumably due to disorder giving rise to weak electron density. In addition to this, a segment of the N-termini, a loop between β4 strand and αC helix, and the activation loop are also missing from the CaMK1D crystal structure. It is apparent from the structure that dimerisation of CaMK1D appears to be mediated by the N-terminal β1 strand which forms a dimer interface. Interestingly our CaMK1D secondary structure prediction based on TALOS lacks the β1 strand and αB helix suggesting that these secondary structure elements defined in the CaMK1D crystal structure are artefacts of dimerisation driven by crystallisation. Moreover we did observe any evidence of dimerisation by NMR or size exclusion chromatography.

Large C-terminal
α helical lobe
(substrate binding)

Small N-terminal
β sheet lobe
(ATP binding)

Small N-terminal
β sheet lobe
(ATP binding)

αC helix

p-loop

hinge

catalytic loop

Large C-terminal
α helical lobe
(substrate binding)

AID/CBD
αR2 helix

AID αR1 helix
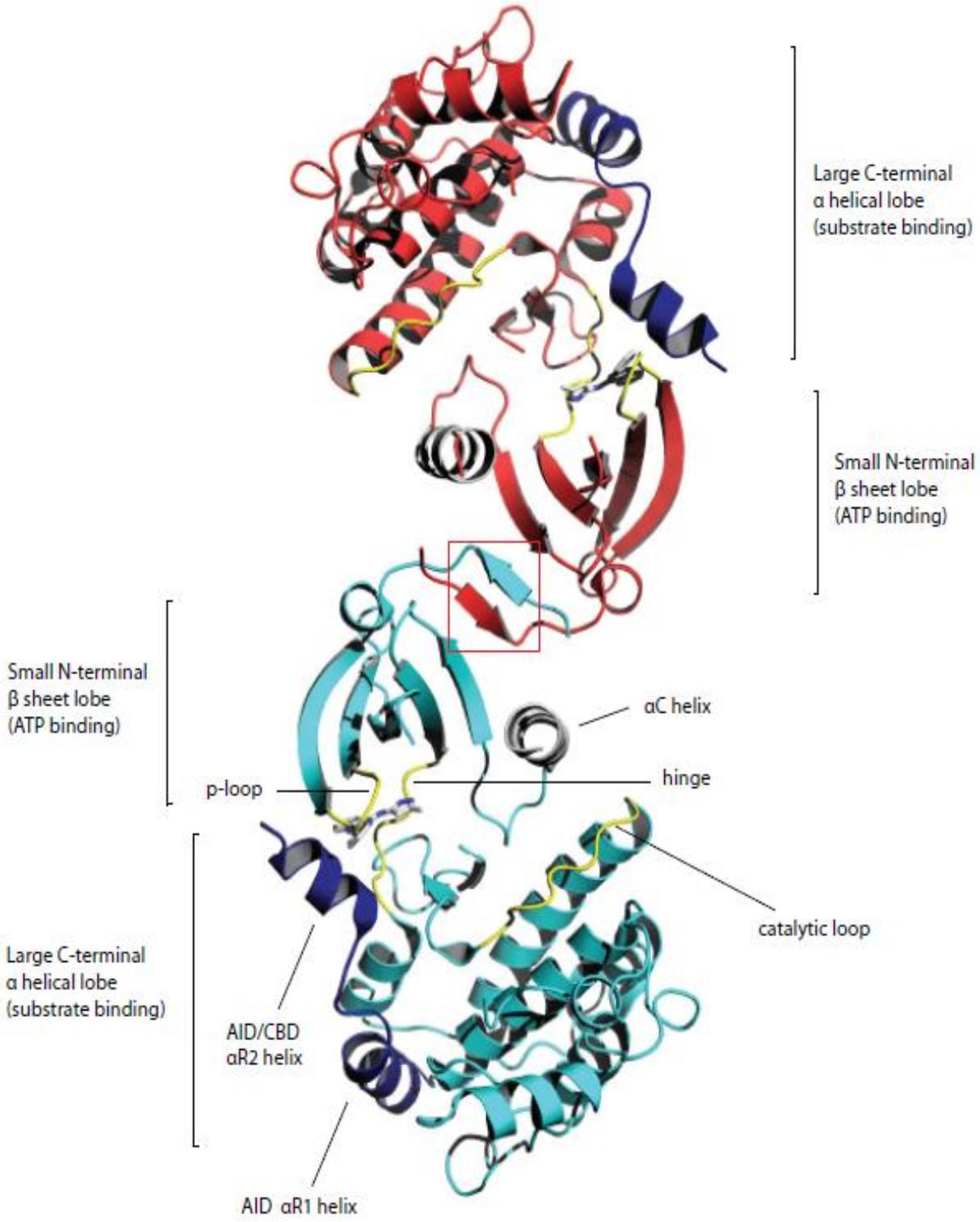
Figure 113. Crystal structure of the CaMK1D-GSK-3 XIII inhibitor complex dimer. CaMK1D was co-crystallised with a GSK-3 XIII inhibitor by the SGC (Knapp et al) and subsequent structure determination at 2.3 angstroms revealed that the CaMK1D-GSK-3 XIII inhibitor complex (PDB: 2JC6) crystallised as a dimer via an N-terminal β strand that forms a dimer interface (highlighted by the red box). The CaMK1D structure reveals a canonical bilobal kinase fold (cyan and red) and a C-terminal regulatory AID (blue) that interacts with the C-terminal substrate binding site and the small N-terminal lobe. Interactions between the AID and substrate binding site are driven by a pseudosubstrate mechanism which inhibits the kinase. Other functional elements which are highlighted in the structure include the regulatory C helix (gray), the hinge, catalytic loop, and p-loop (all shown in yellow). Notably the activation loop (residues 164-185) was not determined due to weak electron density. The ATP competitive inhibitor GSK-3 XIII was shown to bind to the ATP binding site. CaMK1D (10mg/ml) was mixed with 3-fold excess of GSK-3 XIII inhibitor and crystals were obtained by the SGC at 4 °C in sitting drops using the vapour diffusion method by mixing 75 nl of the concentrated protein (10 mg/ml in 25 mM HEPES pH 7.5, 150 mM NaCl) with 75 nl of the well solution (0.1 M citrate pH 5.6, 20 % isopropanol, 20% PEG 4K).

CaMK1D hybrid NMR/crystal structure model

Refinement of the CaMK1D crystal structure and modelling of missing loops using TALOS data

The previously described CaMK1D crystal structure contained 4 regions of disorder that resulted in missing electron density. These regions correlated to residues 1-10 (N-termini), 59-63 (loop between β4 strand and αC helix), 164-185 (activation loop), and 314-333 (C-termini). Furthermore our comparison of the secondary structure defined by the crystal structure and TALOS data highlighted some discrepancies and that some secondary structure elements in the crystal structure maybe due to crystallisation artefacts. We thus decided to refine the crystal structure as a monomer and model the disordered regions by using ψ and φ dihedral angle restraints predicted by TALOS which utilises chemical shift data from our CaMK1D backbone resonance assignments.

The refined CaMK1D models were subsequently generated by CNS and 5 of these models were superimposed (figure 114) to reveal that residues at the N-termini and C-termini exhibited a large degree of flexibility and dynamics which was not surprising because they are part of the flexible and disordered termini. Sequence analysis based on Ca^{2+}/CaM and rat CaMK1A CBD binding studies (Matsushita and Nairn 1998) indicates that the CaMK1D CBD spans from residue 303-320 but in the crystal structure the electron density of residues beyond Thr 313 were not determined presumably due to disorder and dynamics causing weak electron density. In our model residues 314-320 which comprise the remainder of the CBD appear as flexible random coil thus reinforcing the idea that this part of the CBD is unstructured and dynamic. Similarly the activation loop also exhibits a considerable degree of flexibility and dynamics but samples a smaller conformational space. Residues 59-63 which comprise part of a small loop displayed minimal mobility in the model.
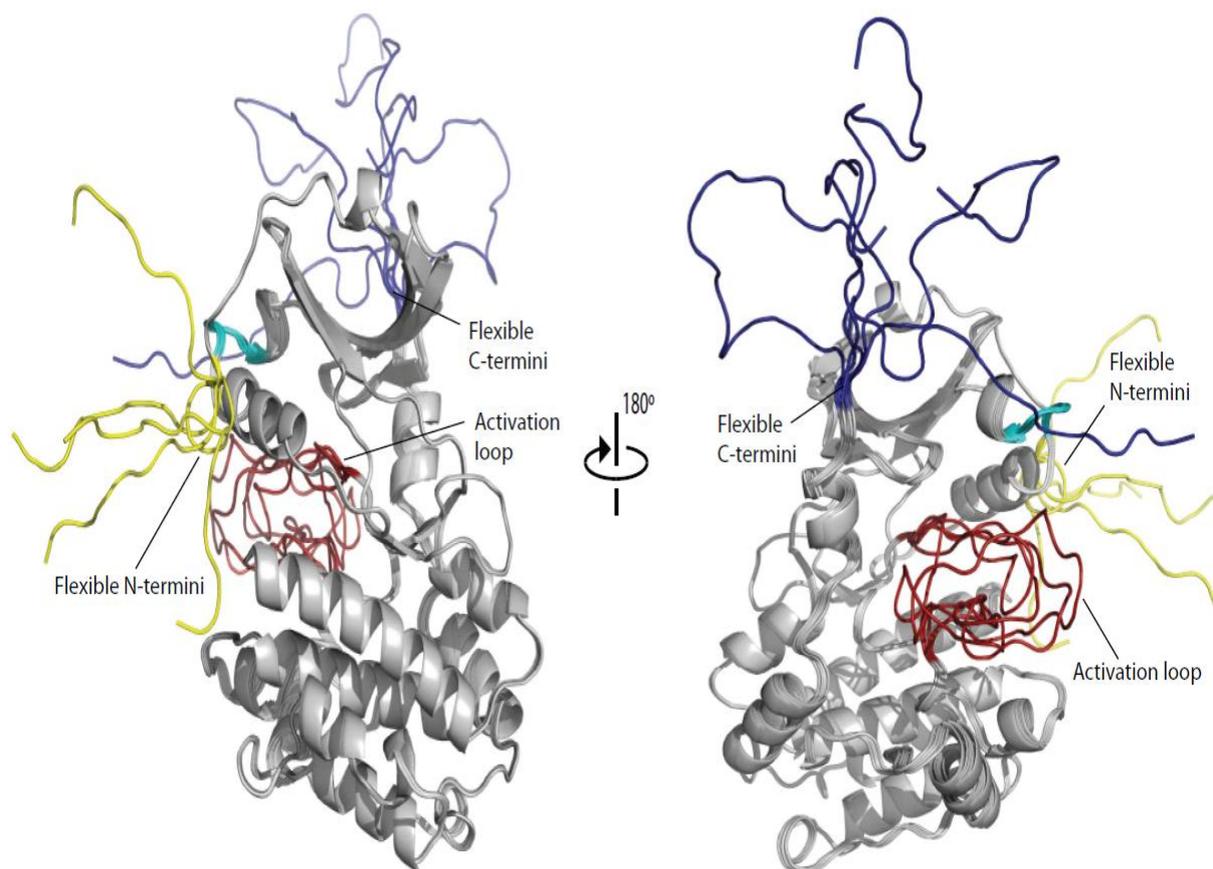
Figure 114. CaMK1D crystal structure and NMR hybrid model. The crystal structure of CaMK1D and its missing loops was refined and modelled as a monomer using ψ and φ dihedral angle restraints predicted by TALOS using chemical shifts from our backbone resonance assignments. Five resulting structures were then generated from CNS and superimposed via the Cα carbons of the CaMK1D structure (gray) using molmol. Modelling of the disordered loops are highlighted and include the N-termini (yellow), C-termini (blue), the loop between β4 strand and αC helix (cyan), and the activation loop (red). These disordered loops exhibit flexibility and dynamics.

Characterisation of the CaMK1D-Ca^{2+}/CaM complex by NMR, ITC, and thermofluor^{TM}

CaMK1D-Ca^{2+}/CaM complex by NMR suggests that Ca^{2+}/CaM binds tightly to the CBD

Most members of the CaMK family are regulated by a C-terminal AID and require Ca^{2+}/CaM binding to the CBD to relieve autoinhibition but the mechanism of activation is not fully understood (Hook and Means 2001; Soderling and Stull 2001; Swulius and Waxham 2008). Hence we decided to investigate the interaction between CaMK1D and Ca^{2+}/CaM from the perspective of CaMK1D by using NMR in order to gain insight into the solution structure of the complex, mechanism of activation, and binding mode. This involved titrating unlabelled Ca^{2+}/CaM into ^{15}N labelled CaMK1D with subsequent conformational changes in CaMK1D monitored via ^{1}H, ^{15}N HSQC TROSY experiments. However, large protein complexes are associated with slow and aniostropic tumbling in solution that causes an increase in T2 relaxation due to enhanced ^{1}H-^{1}H dipolar relaxation and ^{1}H-^{1}H scalar coupling. Formation of the CaMK1D-Ca^{2+}/CaM complex results in a combined molecular weight of 54.6 kDa, therefore deuterated ^{15}N labelled CaMK1D and deuterated CaM were used in order to suppress ^{1}H-^{1}H dipolar relaxation and ^{1}H-^{1}H scalar contributions to ^{1}H lineshape broadening. Furthermore deuteration decreases the amide ^{1}H transverse relaxation rate due to the fact that ^{2}H has a lower gyromagnetic ratio compared to ^{1}H (Sattler and Fesik 1996; Gardner and Kay 1998; Foster, McElroy et al. 2007).

Titrations were also made in increments of 0.25, 0.5, 1.0, 1.5 and 2.0 equivalents of Ca^{2+}/CaM (Figure 116) in order to validate saturation of CaMK1D and to follow the CSPs which were revealed display line shape reduction and slow exchange phenomena thus suggesting an intermediate and high affinity interaction. At 2.0 equivalents of Ca^{2+}/CaM, CaMK1D is saturated, peak intensity is reduced and many broaden out completely whilst. Presumbably because new peaks occur clustered around the centre whilst others broaden out this suggests a large complex formation and anisotropic tumbling in solution.

Previous reports estimate the dissociation constant of rat CaMK1A and Ca^{2+}/CaM by Trp fluorescence to be about 30 nM (Clapperton, Martin et al. 2002). The tracking of peaks that display CSPs on a slow exchange time scale is a lot more difficult compared to fast exchange and requires reassignment of the complex. Thus the majority of the additional peaks observed in the CaMK1D-Ca^{2+}/CaM spectrum cannot be identified without reassignment of the complex. Furthermore due to the significant overlap of peaks seen in the centre of the CaMK1D-Ca^{2+}/CaM spectrum this further complicates the tracking and identification of peaks experiencing CSPs. Mapping of these peak intensity changes to the structure indicates that they correlate predominantly to the AID and the overlapping CBD (figure 116). The large peak intensity reduction for peaks spanning this region even at 0.25 eq of Ca^{2+}/CaM suggests that these residues are directly involved in the interaction with Ca^{2+}/CaM.

Peak intensity reductions were also evident at the base of the C-terminal lobe suggesting weak interactions with Ca^{2+}/CaM upon disengagement of the RD as illustrated by our SAXS CaMK1D-Ca^{2+}/CaM complex model (figure 111). The other notable region displaying peak broadening was the p-loop which has been shown to make contact with the AID via the overlapping CBD (αR2 helix) in the crystal structure of CaMK1D. Peak intensity reduction within and around the vicinity of the p-loop suggests that Ca^{2+}/CaM binding spans the entire CBD (αR2 helix) and part of the p-loop. Further evidence in support of this has been demonstrated by the recent DAPK-Ca^{2+}/CaM structure which indicates that Ca^{2+}/CaM binding spans near this region and slightly alters the interaction between the CDB and p-loop (de Diego, Kuper et al. 2010).

No CSP or significant line broadening was observed for Trp 306 side chain εH

A key residue implicated in the CaMK1D-Ca^{2+}/CaM interaction is reported to be Trp 306 (Goldberg, Nairn et al. 1996; Matsushita and Nairn 1998; Clapperton, Martin et al. 2002). In the bottom left corner of the spectrum we observe 5 peaks that corresponds to 5 Trp side chain amide εH but none of these undergo any significant CSPs or line shape broadening in our spectrum. We thus postulate that these 5

peaks that are observed do not correspond to each of the 5 individual Trp side chain amides εH including Trp 306 suggesting that some of the other Trp side chain amides εH maybe existing in more than 1 conformation and that we cannot observe the Trp 306 side chain amide εH because it is intermediate exchange. Alternatively the side chain of Trp 306 in CaMK1D may not be implicated in the interaction as is believed but we think this is an unlikely scenario.

Additional peaks in the CaMK1D-Ca^{2+}/CaM ^{1}H ^{15}N TROSY-HSQC suggests random coil and/or increased helical content

The appearance of additional peaks in the centre of the spectrum and apparent reduction in backbone amide resonance dispersion in the first increment of the ^{1}H, ^{15}N TROSY HSQC (data not shown) initially suggested that there was an unfolding mechanism occurring but based on the crystal structures of CaMK AID peptide-Ca^{2+}/CaM (Clapperton, Martin et al. 2002; de Diego, Kuper et al. 2010) and DAPK-Ca^{2+}/CaM complex an unfolding mechanism does not correlate with these structures. In contrast our SAXS model of the CaMK1D-Ca^{2+}/CaM complex suggests that there is unfolding in the RD upon disengagement, perhaps facilitating its release from the substrate binding site.

Limited CSPs are seen in the functional elements upon Ca^{2+}/CaM binding

A protein kinase that switches from an inactive to active state is also often associated with rearrangement of the p-loop, αC helix, catalytic lysine, catalytic loop and stabilization of the activation loop. These changes are concomitant with formation of the R-spine and completion of the C-spine upon ATP binding which reflects an active and catalytically competent state (Taylor and Kornev 2010). All these events bring about the orientation of key catalytic residues into a catalytically competent configuration for ATP binding and phosphoryl transfer (Adams 2001). In our CaMK1D-Ca^{2+}/CaM ^{1}H, ^{15}N TROSY HSQC the peaks that correlate to the backbone amides of residues within key functional elements such as the activation loop, catalytic loop, hinge, catalytic lysine (Lys 52 in CaMK1D), and regulatory αC helix exhibit peak intensity reduction above the mean suggesting small

conformational changes. This suggests that binding of Ca^{2+}/CaM reconfigures the conformation of these elements into a catalytically competent state. For the appearance of new peaks in slow exchange at saturation only reassignment of the complex would likely confirm their chemical shift.

A loop preceding the AID and substrate binding site only has small peak intensity redcution

Other regions that we postulate to be implicated in the CaMK1D-Ca^{2+}/CaM interaction include the loop that precedes the AID and the substrate binding site however these do not exhibit significantly large peak intensity reduction, leaving ambigiuity in the state of the complex that is observed in solution. We expect structural changes in this loop region on the basis of our proposed SAXS CaMK1D-Ca^{2+}/CaM complex model (figure 111) because the swing out motion of the AID bound to Ca^{2+}/CaM via the CBD causes rearrangement of the AID and bound Ca^{2+}/CaM to occupy a position near the base of the CaMK1D C-terminal lobe resulting in the opening of the substrate binding site. This conformational change would require rearrangement of the AID and the loop preceding the AID but if the pivot point of the AID αR1 helix is closer to the N-terminal end of the αR1 helix then perhaps small changes and thus smaller peak intensity reductions occur or perhaps there chemical environment and dynamic is not significantly altered in autoinhibited and disengaged states. The latter scenario could explain the lack of CSPs and peak intensity reduction associated with residues from the loop.

Ca^{2+} is required for CaM to interact with CaMK1D

CaM contains 4 E-F hand motifs with each CaM molecule known to bind a total of 4 Ca^{2+} ions. Therefore CaM with respect to Ca^{2+} has a binding stoichiometry of 1:4. It has been reported that CaMK1D and Ca^{2+}/CaM can interact even at substoichiometric concentrations of Ca^{2+} and postulated that CaM can interact weakly with CaMK members independent of Ca^{2+} (Hook and Means 2001). We decided to investigate the interaction of CaM with CaMK1D in the absence of Ca^{2+} and at substoichiometric concentrations of Ca^{2+} to characterise its binding mode in

comparison to Ca^{2+} saturated CaM. Consequently our data demonstrates that CaM cannot interact in the absence of CaM even weakly when up to 2 equivalents of CaM was added due to the absence of CSPs and additional peaks (figure 117). However an interaction between CaMK1D and CaM at a substiochimetric concentration of 1:2 (CaM to Ca^{2+}) was evident due to the appearance of additional peaks. At a stoichiometry of 1:4, 1:8, and 1:16 (CaM to Ca^{2+}) the spectra of the CaMK1D-Ca^{2+}/CaM complex ^1H, ^{15}N TROSY HSQC appeared to resemble the saturated form of the CaMK1D-Ca^{2+}/CaM complex (Figure 118, 119) with evidence of additional peaks in slow exchange indicative of tight binding as previously observed. Overlay of the spectra at different concentrations of Ca^{2+} also indicates that the binding affinity appears to be tight even at a substiochimetric concentration of 1:2 (CaM to Ca^{2+}) because the peaks in slow exchange superimpose well and increase in peak intensity when the Ca^{2+} concentration was increased (Figure 118).
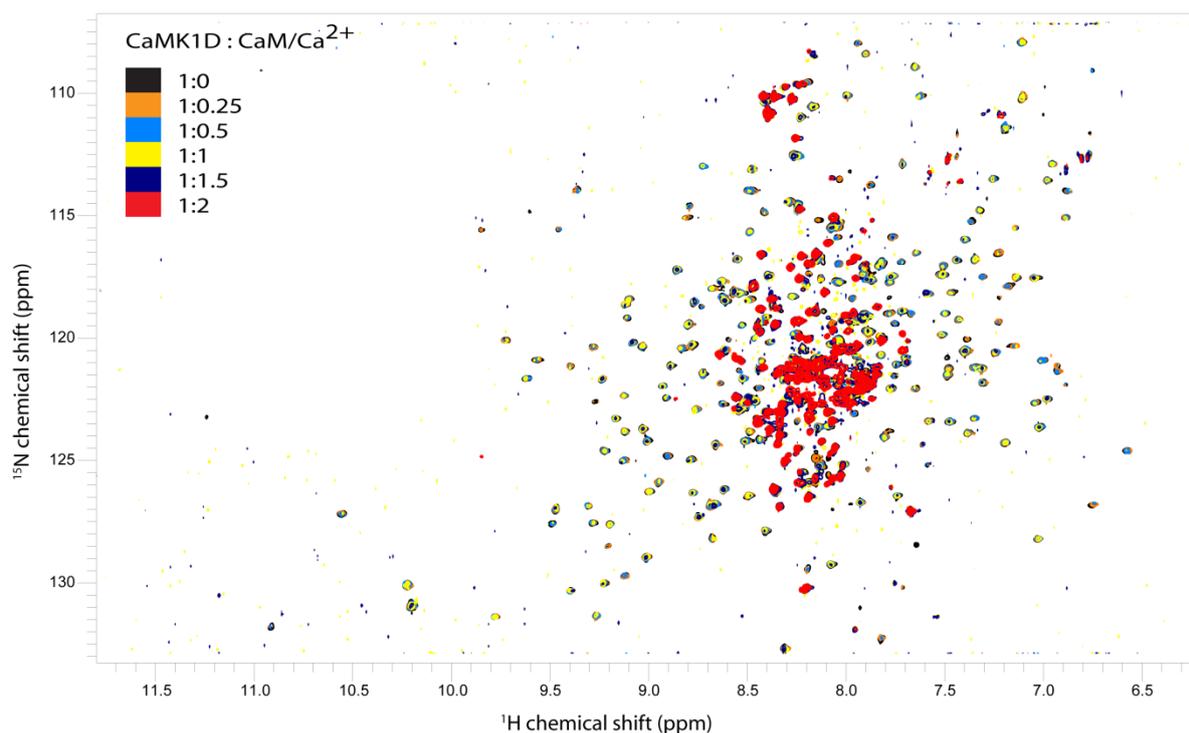
Figure 115. CaMK1D-Ca²⁺/CaM interaction shown by ¹H, ¹⁵N TROSY HSQC. The interaction between apo deuterated ¹⁵N CaMK1D and deuterated Ca²⁺/CaM was monitored by ¹H, ¹⁵N TROSY HSQC. The superimposed spectra indicates that the addition of Ca²⁺/CaM to apo CaMK1D (black spectrum) produces a reduction in peak intensity whilst new peak appear indicating slow exchange timescale of binding suggesting that conformational changes are occurring in the backbone upon a high affinity interaction between CaMK1D and Ca²⁺/CaM.. Spectra were acquired with 0.2 mM of ²H, ¹⁵N CaMK1D in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, CaCl₂, 0.02 % Na Azide on a Varian 800 MHz at 25 °C.
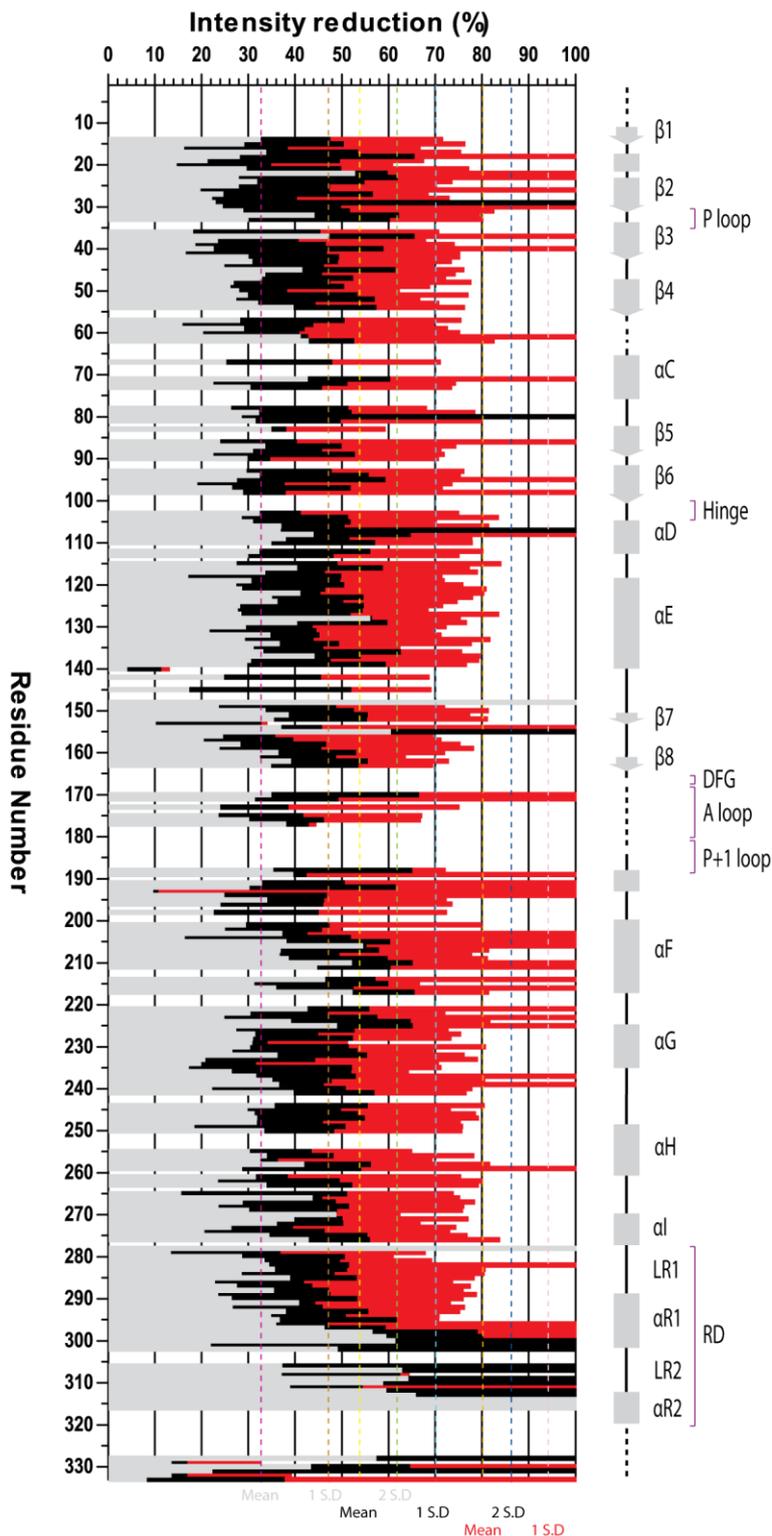
Figure 116. A histogram of CaMK1D amide peak intensity reduction induced by Ca2+/CaM. The mean and and S.D are shown for titration of 0.25, 0.5 and 1,0 eq of Ca2+/CaM in gray, black, and red respectively.
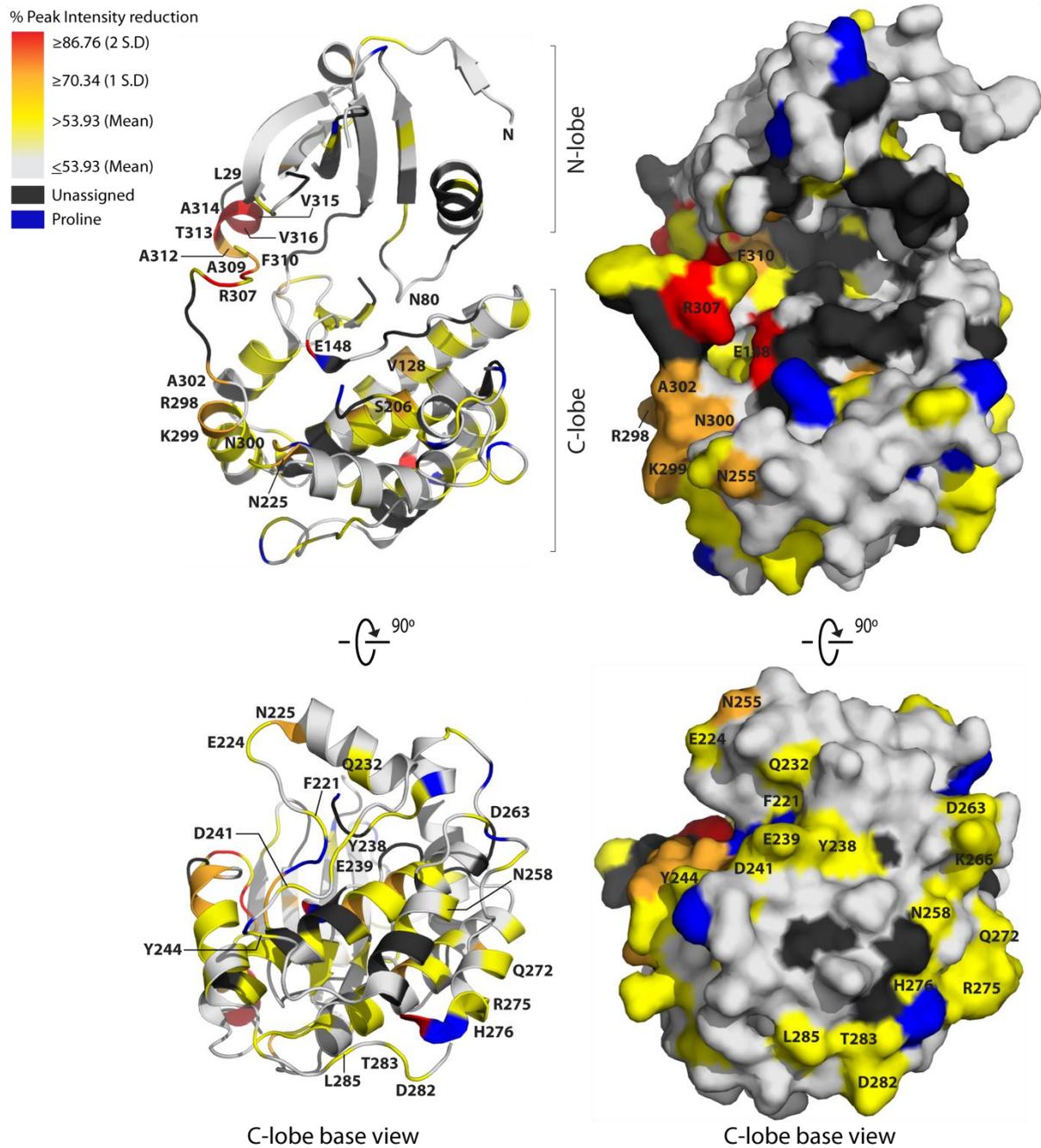
Figure 116A. Peak intensity reduction mapping upon binding 0.25eq Ca2+/CaM. Peak intensity reductions occur around the RD, predominantly the near CBD where Ca^{2+}/CaM is known to bind. The base of the C-lobe is shown to illustrate long range peak intensity reduction suggesting Ca2+/CaM may disengage the Rd and interact weakly with this region.
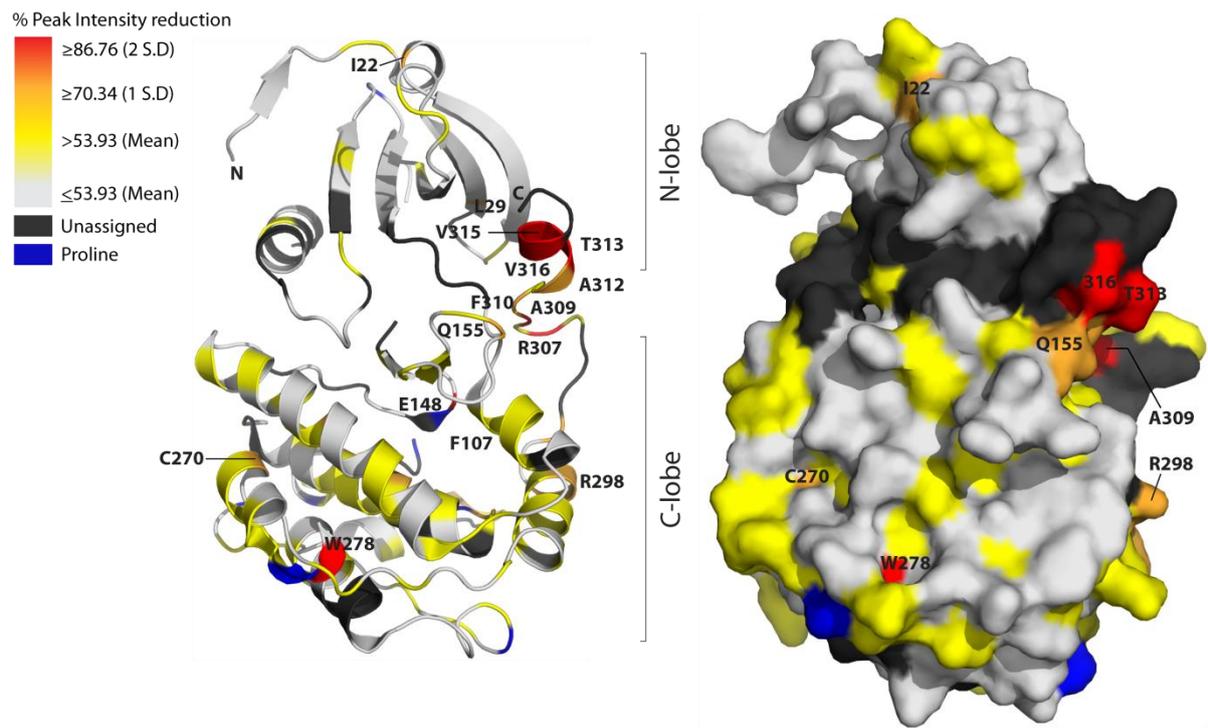
Figure 116B. Peak intensity reduction mapping upon binding 0.25eq Ca2+/CaM. A 180° view of the previous figure
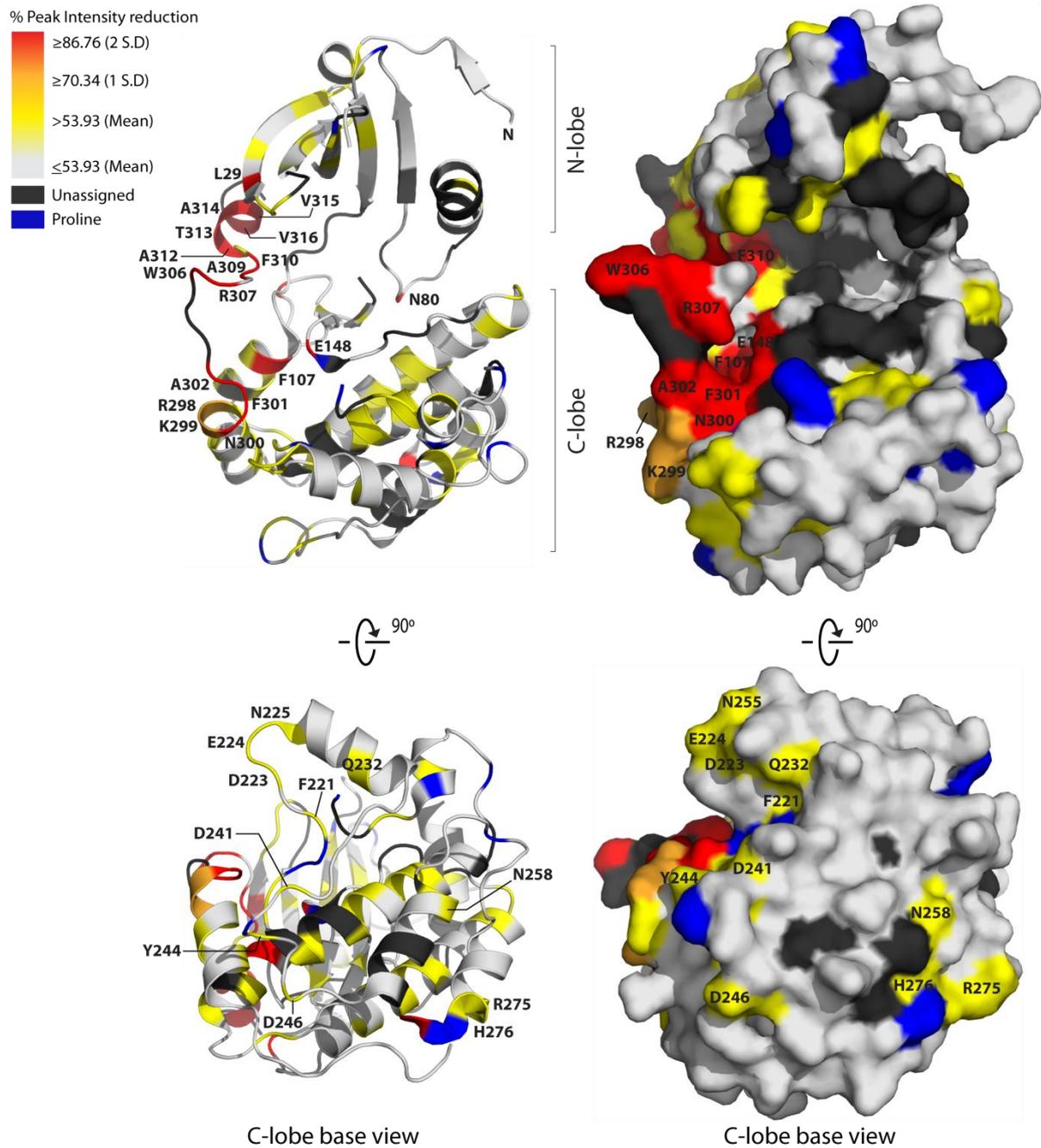
Figure 116B. Peak intensity reduction mapping upon binding 0.5eq Ca2+/CaM. Peak intensity reductions occur around the RD, predominantly near the CBD where Ca^{2+}/CaM is known to bind. The base of the C-lobe is shown to illustrate long range peak intensity reduction suggesting Ca2+/CaM may disengage the Rd and interact weakly with this region.
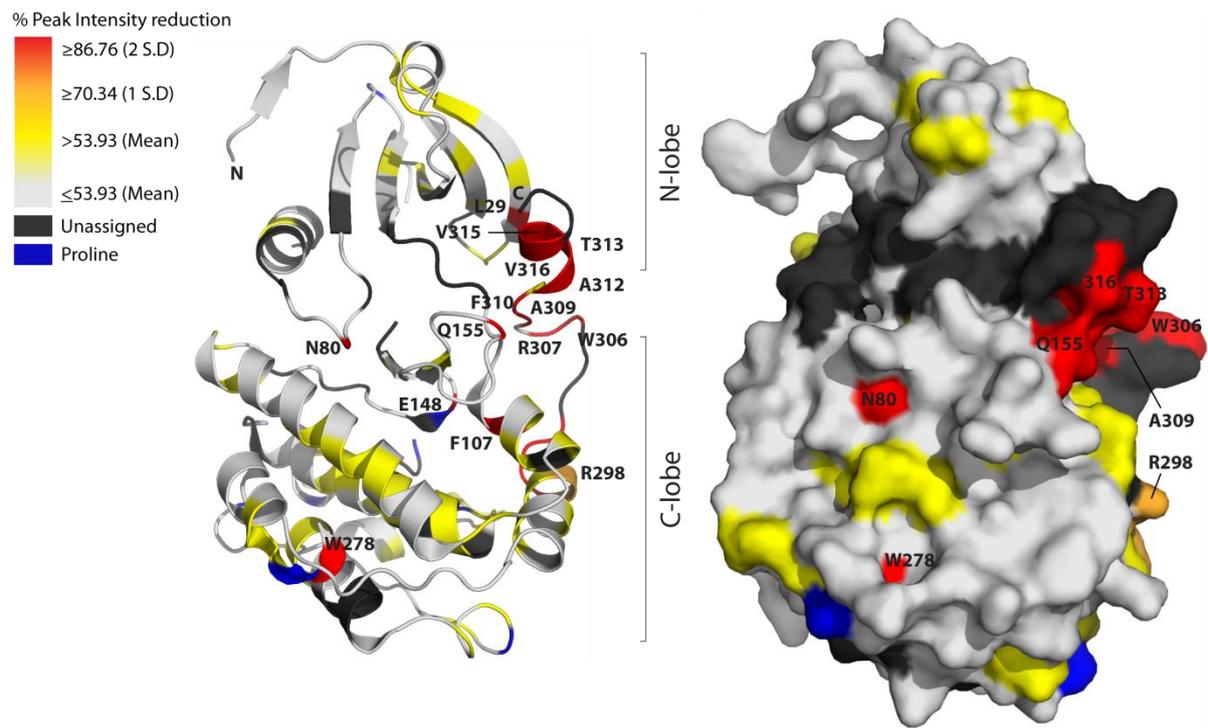
Figure 116B. Peak intensity reduction mapping upon binding 0.5eq Ca2+/CaM.

A 180° view of the previous figure
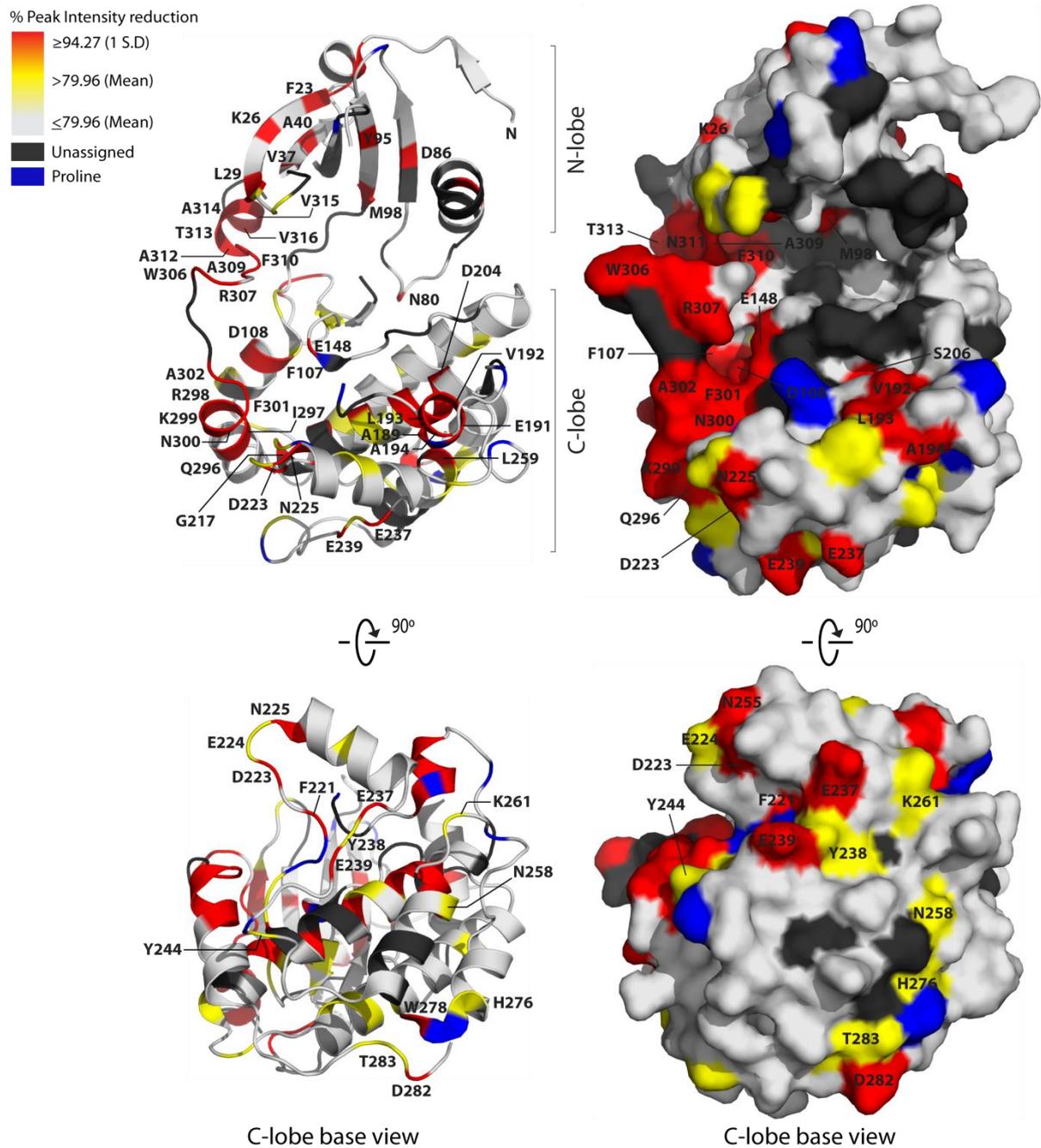
Figure 116C. Peak intensity reduction mapping upon binding 1eq Ca2+/CaM. Peak intensity reductions occur around the RD, predominantly the near CBD where Ca^{2+}/CaM is known to bind. The base of the C-lobe is shown to illustrate long range peak intensity reduction suggesting Ca2+/CaM may disengage the Rd and interact weakly with this region.

.
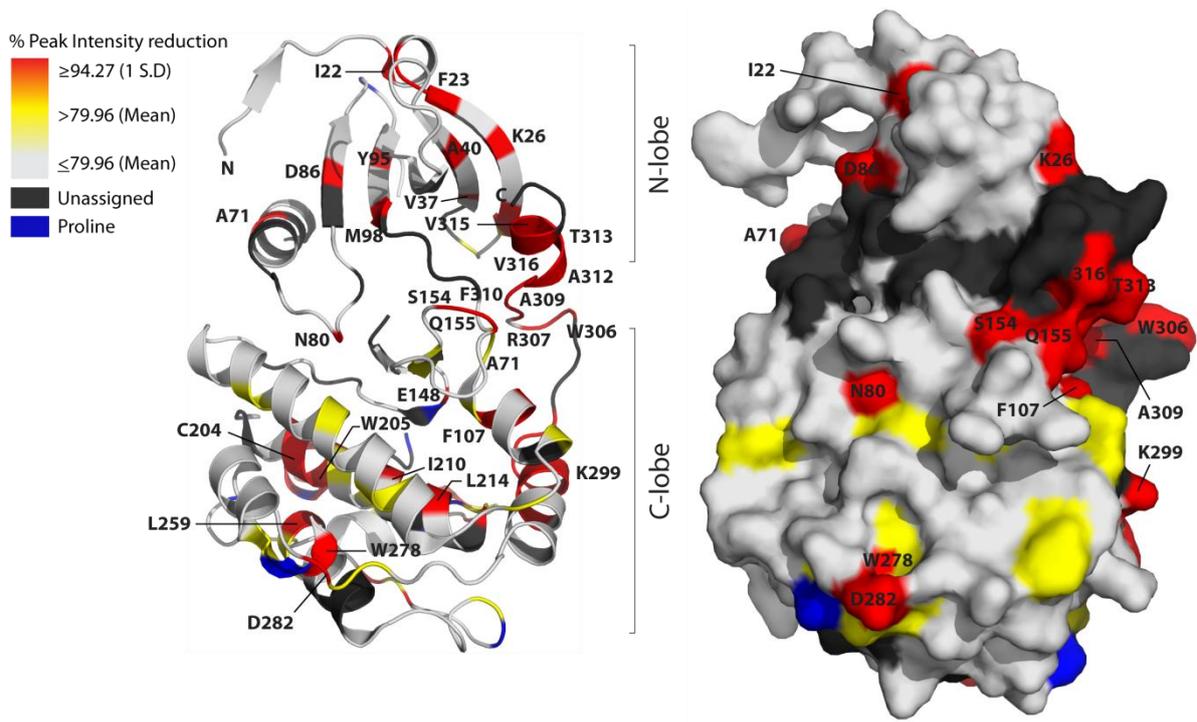
Figure 116D. Peak intensity reduction mapping upon binding 1eq Ca2+/CaM. A 180° view of the previous figure.
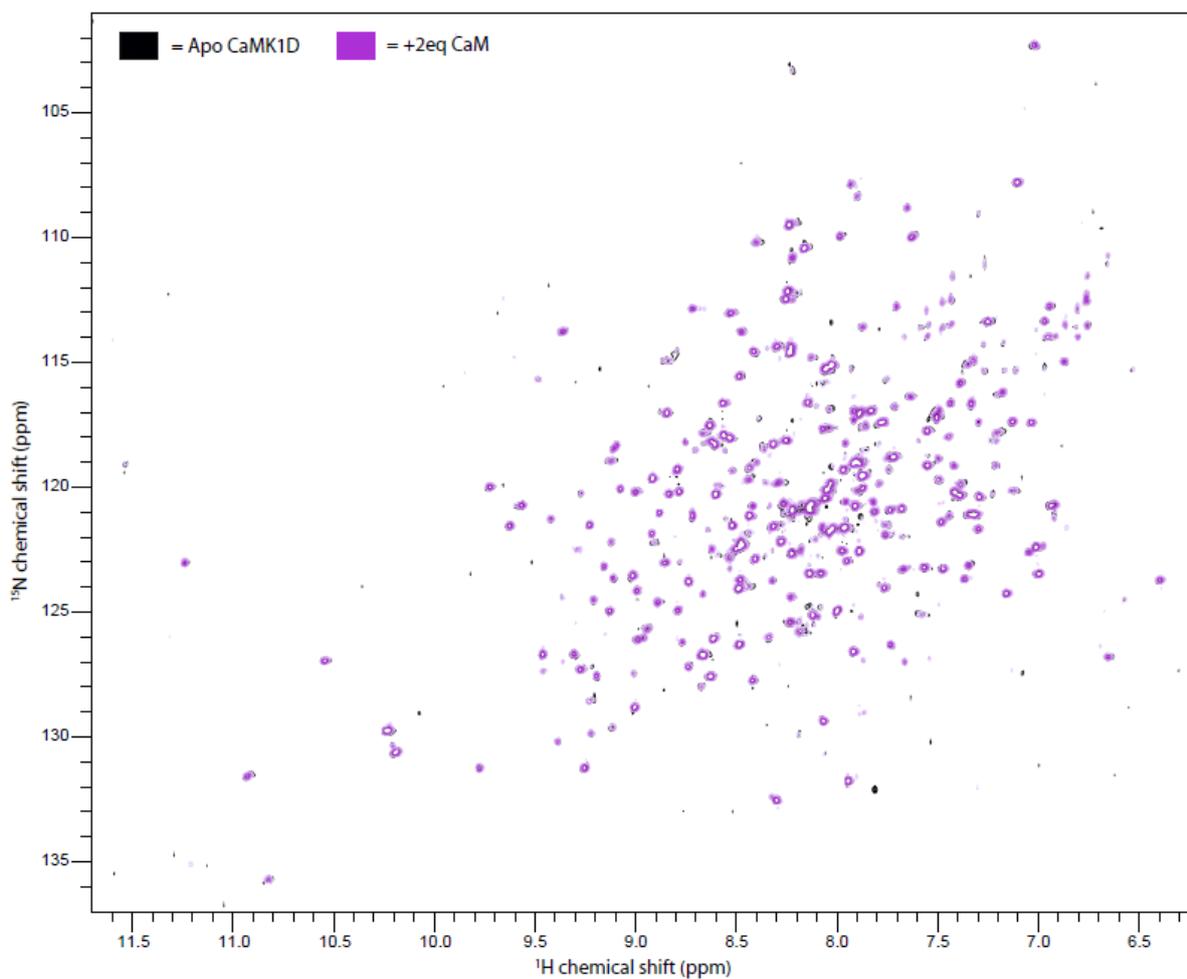
Figure 117. ^{1}H, ^{15}N TROSY HSQC of CaMK1D and CaM. A titration of CaMK1D with two equivalents of CaM in the absence of Ca^{2+} indicates that there is no interaction because no CSPs were evident. Spectra were acquired with 0.185 mM of ^{2}H, ^{15}N CaMK1D in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, CaCl$_2$, 0.02 % Na Azide on a Varian 800 MHz at 25 °C.
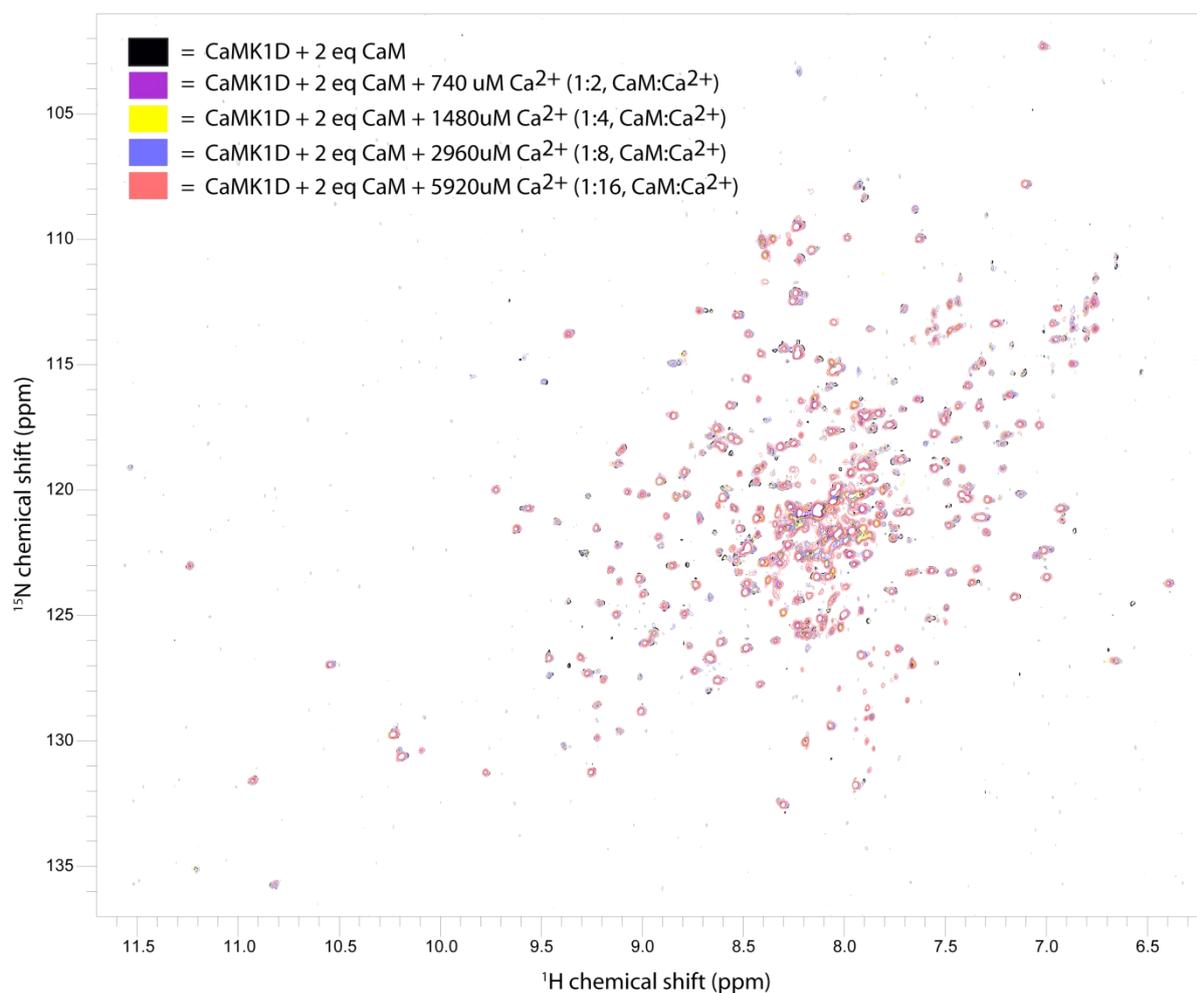
Figure 118. ^{1}H, ^{15}N TROSY HSQC of CaMK1D and CaM titration with Ca^{2+}. CaMK1D in the presence of two equivalents of CaM (black) was titrated with increasing concentrations of Ca^{2+}. Upon addition of 740 uM Ca^{2+} additional peaks appear in slow exchange indicative of tight binding affinity. With each increasing addition of Ca^{2+} the intensity of the additional peaks in slow exchange increases further suggesting an increase in affinity. Spectra were acquired with 0.185 mM of ^{2}H, ^{15}N CaMK1D in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, CaCl$_2$, 0.02 % Na Azide on a Varian 800 MHz at 25 °C.
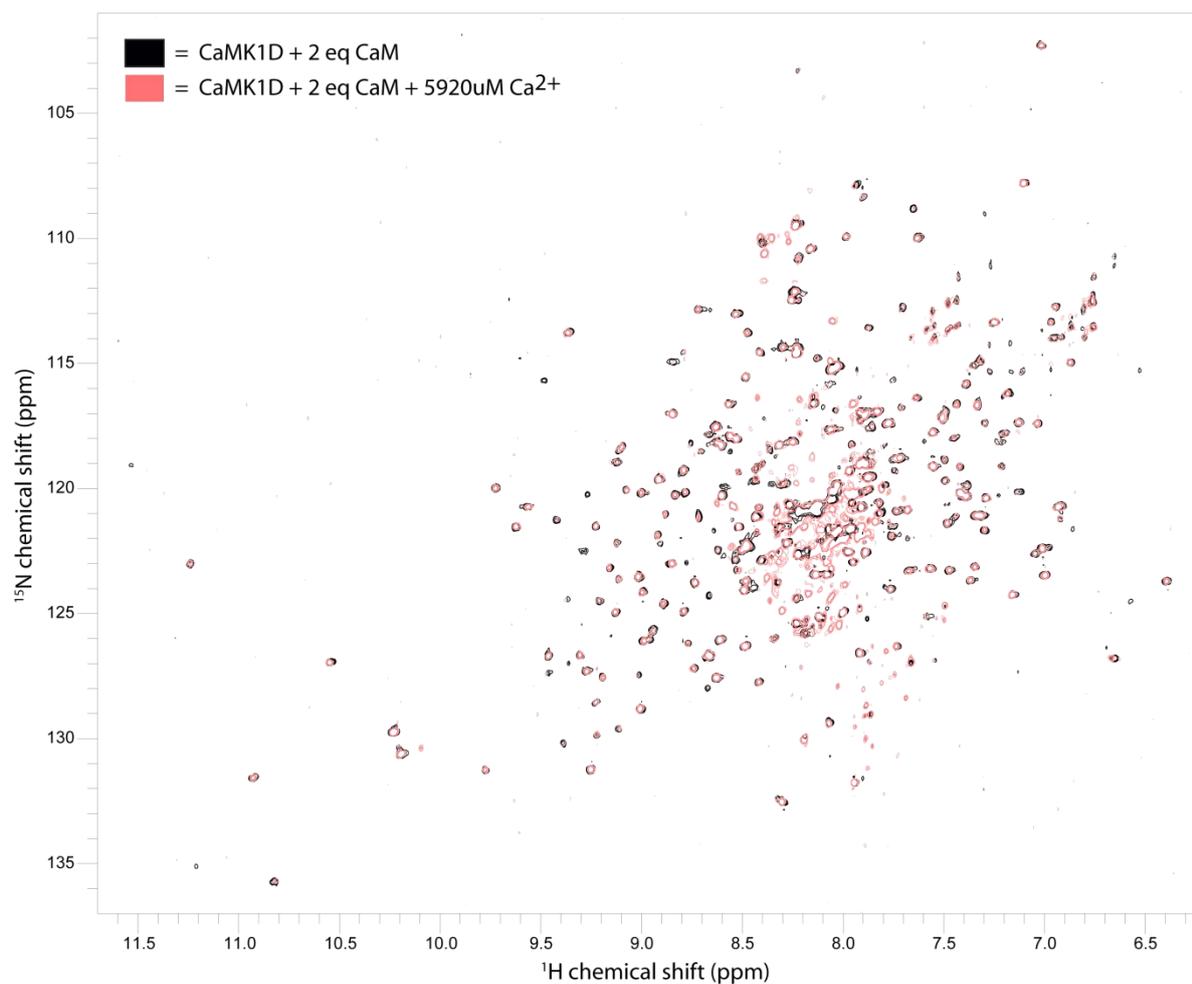
Figure 119. CaMK1D and Ca²⁺/CaM endpoint titration with 5920 uM Ca²⁺ by ¹H, ¹⁵N TROSY HSQC. For simplification of the previous spectra this spectra only illustrates CaMK1D in the presence of two equivalents of CaM (black) being titrated with 5920 uM of Ca²⁺.

ITC confirms the high affinity interaction between CaMK1D and Ca^{2+}/CaM

We previously inferred from our ^{1}H, ^{15}N TROSY HSQC spectra showing the titration of deuterated Ca^{2+}/CaM with deuterated ^{15}N CaM1D (figure 115) that there was a high binding affinity between the proteins due to evidence of CSPs in slow exchange. Isothermal titration calorimetry (ITC) was therefore used to determine the binding affinity between CaMK1D and Ca^{2+}/CaM (figure) and we found an average dissociation constant (K_d) of 28.5 nM (deduced from 2 separate ITC experiments which produced K_d values of 38$\pm$12.9 nM and 19$\pm$7.7 nM respectively). These results reinforce the high affinity interaction suggested by our NMR data. Our determined K_d value for the CaMK1D interaction with Ca^{2+}/CaM was comparable with the reported K_d value for the interaction between rat CaMK1A and Ca^{2+}/CaM determined by Trp fluorescence (Clapperton, Martin et al. 2002). Likewise Knapp et al report similar binding affinities for CaMK proteins by ITC (Knapp 2010).

The CaMK1D and Ca^{2+}/CaM interaction is enthapically driven and forms a 1:1 complex

The heat change associated with the interaction between Ca^{2+}/CaM (ligand) and CaMK1D (macromolecule) shown in figure, indicates an exothermic reaction in the system and corresponds to a ΔH (enthalpy change) of -5431$\pm$171.5 cal mol^{-1}, whilst the associated ΔS (entropy change) was 15.7 cal mol^{-1} K^{-1}. The thermodynamic parameter ΔH reflects the strength of non-covalent interactions between Ca^{2+}/CaM and CaMK1D relative to those existing with the solvent whereas ΔS reflects entropic contributions from changes in solvation entropy and changes in conformational entropy. Both ΔH and ΔS are determinants of ΔG (free energy of binding) which is also related to the association constant (K_a) because $\Delta G = -RTlnK_a = \Delta H - T\Delta S$. Therefore our ΔH and ΔS values produce a large negative ΔG indicating that the binding of Ca^{2+}/CaM to CaMK1D is tight and a spontaneous process. Furthermore the interaction appears to be enthalpically driven is because the ΔH value was negative and of greater magnitude than the ΔS contribution. The stoichiometry of the Ca^{2+}/CaM and CaMK1D interaction was expected to be 1:1 based evidence of rat CaMK1A and other members of the CaMK family () interacting with Ca^{2+}/CaM. Our ITC data produced an N value (stoichiometry) of 0.95+0.0187 (figure) which

approximates to a 1:1 stoichiometry between CaMK1D and Ca^{2+}/CaM. The data for these thermodynamic parameters represents the results from one of two ITC experiments and was comparable to results produced from the second experiment (data for the repeated experiment not shown).

The panel annotations read:

CaMK1D-Ca^{2+}/CaM
Chi^2/DoF = 1.372E5
N   0.950 ±0.0187
K   2.62E7 ±1.36E7
ΔH   - 5341 ±171.5
ΔS   15.7
Kd   38 nM

Kd average = 28.5 nM (deduced as average from 2 separate ITC CaMK1D + Ca^{2+}/CaM experiemnts)
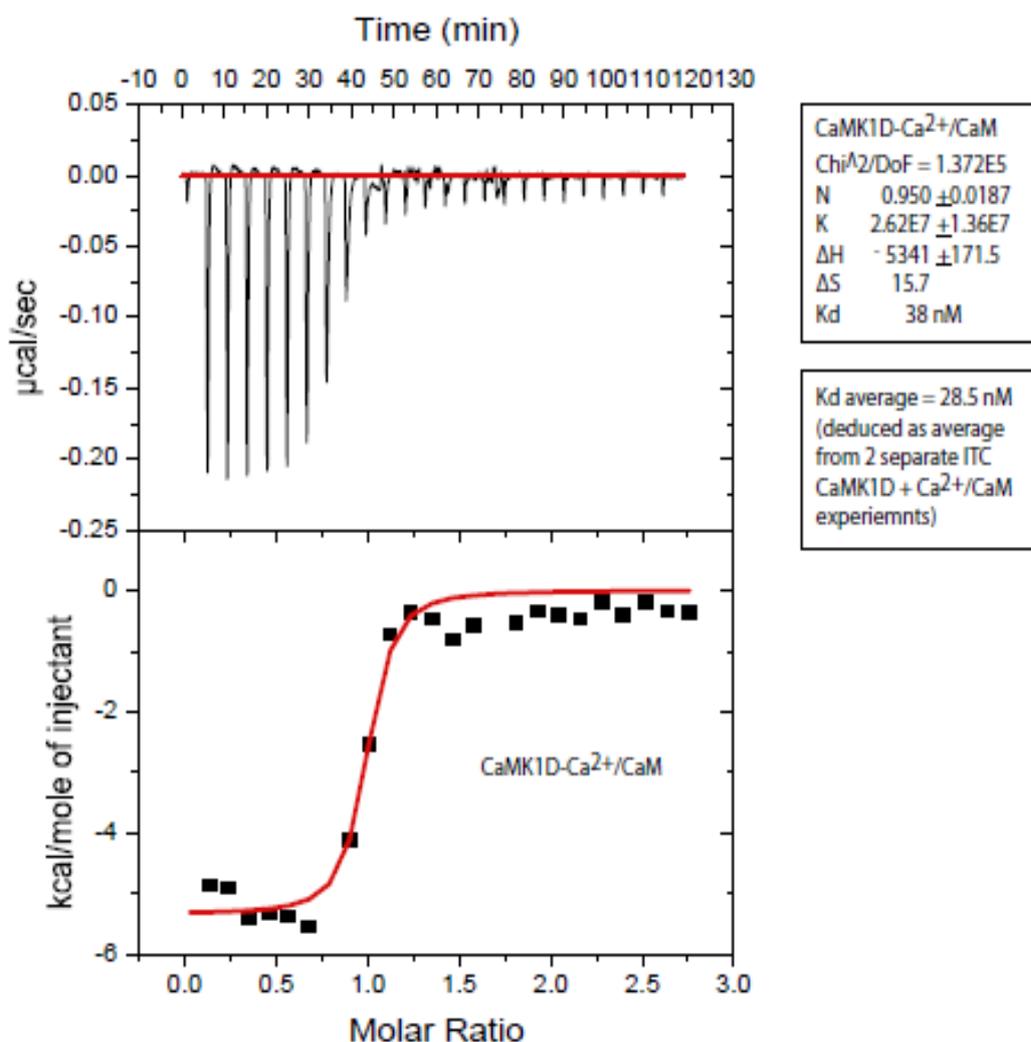
CaMK1D-Ca^{2+}/CaM

Figure 121. Characterisation of the CaMK1D and Ca^{2+}/CaM interaction by ITC. The figure shows the raw exothermic heat changes associated with each injection (upper panel) and the binding isotherm which illustrates the integrated binding enthalpy changes (lower panel). The K_d determined for this experiment and a repeat experiment was 38 nM and 19 nM respectively thus giving an average K_d of 28.5 nM. The large negative ΔH and small positive ΔS values contribute to a negative and thus favourable ΔG of binding indicative of a spontaneous and high affinity interaction. For this particular ITC experiment, the protein concentrations used were 12 µM CaMK1D (macromolecule) and 120 µM Ca^{2+}/CaM (ligand). All CaMK1D and Ca^{2+}/CaM experiments were conducted in 50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM CaCl$_2$ at 25 °C.

The CBD mutant CaMK1D_W306S exhibits weak binding affinity with Ca^{2+}/CaM

A key CaMK1D residue that is postulated to be important in mediating the tight binding interaction between CaMK1D and Ca^{2+}/CaM is Trp 306. This residue is located in the CBD and its side chain is solvent exposed in the rat CaMK1A structure (Goldberg, Nairn et al. 1996). It is postulated that this Trp side chain (Trp 303 in rat CaMK1A) can initiate the interaction with Ca^{2+}/CaM by binding into a hydrophobic pocket within one of the Ca^{2+}/CaM globular domains (Clapperton, Martin et al. 2002). Studies by Matsushita et al have demonstrated that mutation of rat CaMK1A Trp 303 to a Ser increased the K_d for the rat CaMK1A interaction with Ca^{2+}/CaM to about 100 µM. Analysis of the CaMK1D structure however indicates that Trp 306 is not solvent exposed but instead appears to be buried unlike in rat CaMK1A where its Trp 303 is solvent exposed (Figure 122). Assuming Trp 306 in CaMK1D is indeed solvent exposed then the CBD conformation in the CaMK1D crystal structure is not representative of its native conformation in solution but is likely an artefact of crystallisation or due to the presence of GSK-3 XIII. We predict that mutation of CaMK1D Trp 306 will reduce the binding affinity with Ca^{2+}/CaM without significantly affecting the global backbone conformation i.e. not because of unfolding but because of a localised side chain mutation.

Alternatively we also hypothesise that if Trp 306 is indeed buried then a non conservative mutation to a Ser will cause drastic changes to the hydrophobic interior of the protein and cause significant backbone conformational changes whereas a more conservative mutation to a Phe would result in subtle changes to the backbone conformation. To address these hypothesises we produced a CBD mutant termed CaMK1D_W306S and first used ITC in order to determine the K_d and associated thermodynamics for its interaction with Ca^{2+}/CaM. Production of a more conservative mutant, CaMK1D_W_306F was also attempted but could not be produced successfully due to expression issues. Our CaMK1D_W306S ITC data (figure 123) indicates that the binding isotherm was very shallow and that the K_d and thermodynamic parameters could not be determined quantatively. The shallow binding isotherm suggests that there was insufficient heat change associated with

each injection because either or both the affinity of the interaction was too weak or the CaMK1D_W306S concentration in the sample cell was too low.

Based on results demonstrated by the rat CaMK1A Trp mutant we predicted the K_d for CaMK1D_W306S binding to Ca^{2+}/CaM to be around 100 µM and therefore required a high CaMK1D_W306S concentration in order to produce a sufficient heat change and a reasonable c value ($c = NK_a[M]_T$) for a sigmoidal binding isotherm. Despite using 250 µM CaMK1D_W306S and 2.5 mM Ca^{2+}/CaM to produce a c value of 2.5, the heat change was insufficient and thus below the detection limits of the ITC instrument. The range of K_d that can be determined by the ITC instrument is from estimated to be from 0.01-100 µM thus implying that the K_d for the binding interaction between CaMK1D_W306S and Ca^{2+}/CaM was at least greater than 100 µM.
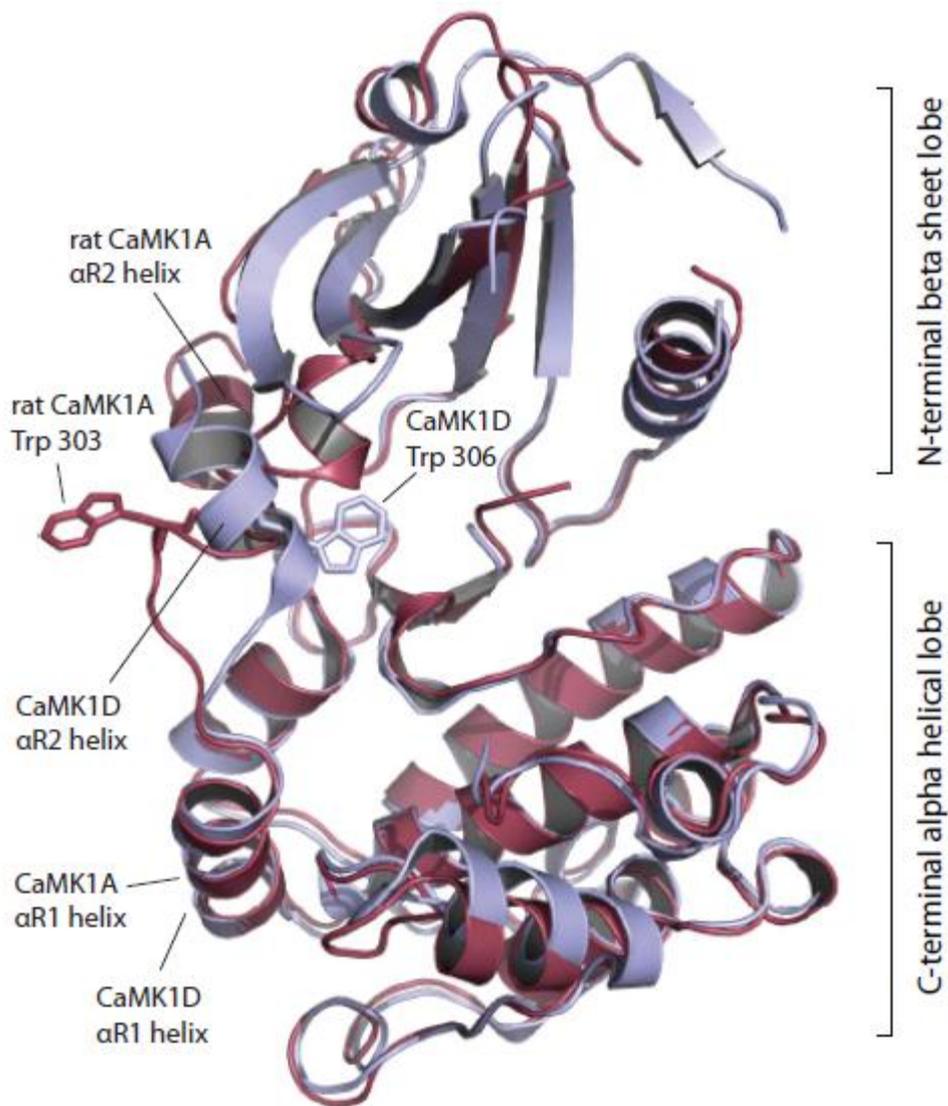
Figure 122. Human CaMK1D and rat CaMK1A structural alignment. A Structural alignment of CaMK1D (light blue) and rat CaMK1A (purple) was conducted to examine potential differences in structure. It is apparent from the alignment that structural differences arise in the CBD region defined by the αR2 helix. Specifically highlighted is the conformation of CaMK1D Trp 306, which is buried whereas rat CaMK1A Trp 303 is exposed to solvent. These residues are implicated in mediating interactions with Ca^{2+}/CaM. Given that CaMK1D Trp 306 appears to be buried implies that it might have a different binding mode to Ca^{2+}/CaM or it is the result of a crystallisation artefact and is a non-native conformation. This structural alignment was created using Pymol and produced a RMS value of 0.835 indicative of a conserved fold.
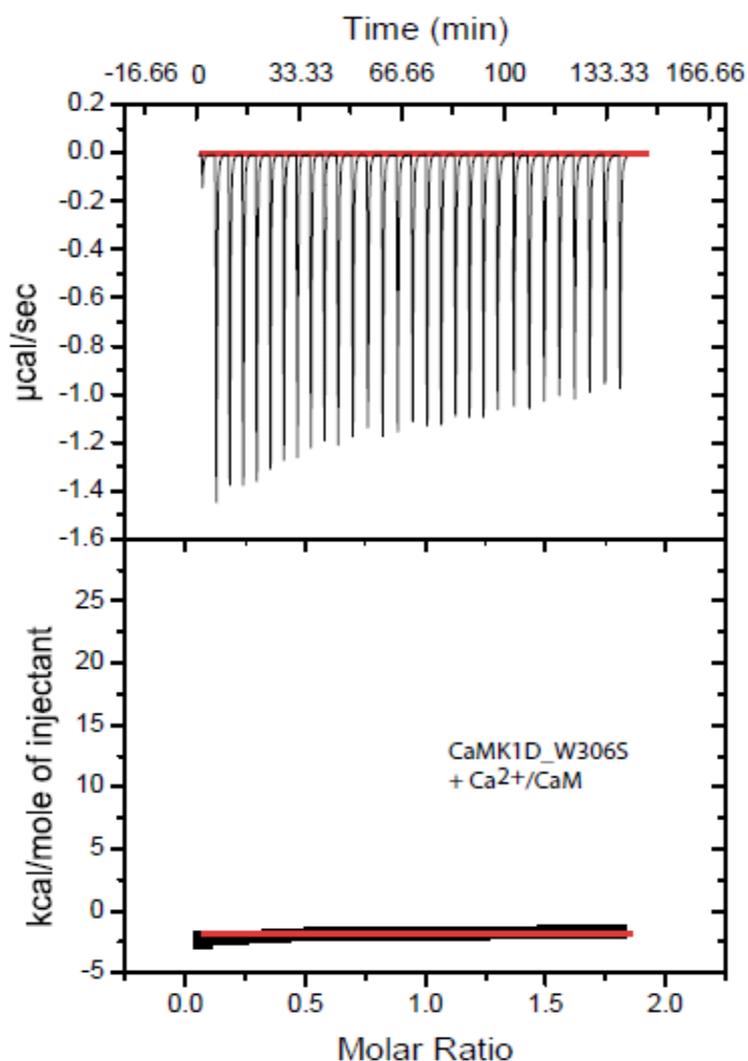
Figure 123. Characterisation of the CaMK1D_W306S and Ca²⁺/CaM interaction by ITC. The shape of the binding isotherm for the titration of Ca^{2+}/CaM with CaMK1D_W306S appears very shallow indicating that the binding affinity is very weak beyond the K_d sensitivity limits of 0.01 – 100 uM. Therefore we can deduce from the binding isotherm that the K_d is greater than 100 uM . For this particular ITC experiment, the protein concentrations used were 250 μM CaMK1D (macromolecule) and 2500 μM Ca^{2+}/CaM (ligand). All CaMK1D and Ca^{2+}/CaM experiments were conducted in 50 mM Hepes pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 10 mM $CaCl_2$ at 25 °C.

CaMK1D_W306S is folded and suggests that W306 is in intermediate exchange

Structural alignment of CaMK1D and rat CaMK1A revealed an RMSD of 0.835 indicating a highly conserved fold. However a notable structural difference occurs in the CBD region in which the relative orientation of Trp 306 is different. In CaMK1D it appears buried but in rat CaMK1A it is solvent exposed. To confirm that if Trp 306 was solvent exposed or buried we acquired a HQSC TROSY (figure 124) and anticipated that the protein would still be folded suggesting that Trp306 is not buried. Consequently the data indicated the protein to be folded. This confirmed that the side chain is not buried but curiously five Trp side chain amides still appear indicating that Trp306 in CaMK1D is in intermediate exchange and does not correspond to anyone of these five Trp. DNA sequencing was validated to confirm that the mutation was present. This finding reinforces the weak binding (> 100 uM) kd observed by ITC for the CaMK1D_W306S interaction with Ca^{2+}/CaM. CaMK1D_W306S and Ca^{2+}/CaM by ITC were due to the mutation of Trp 306 to Ser and not because of the protein unfolding.
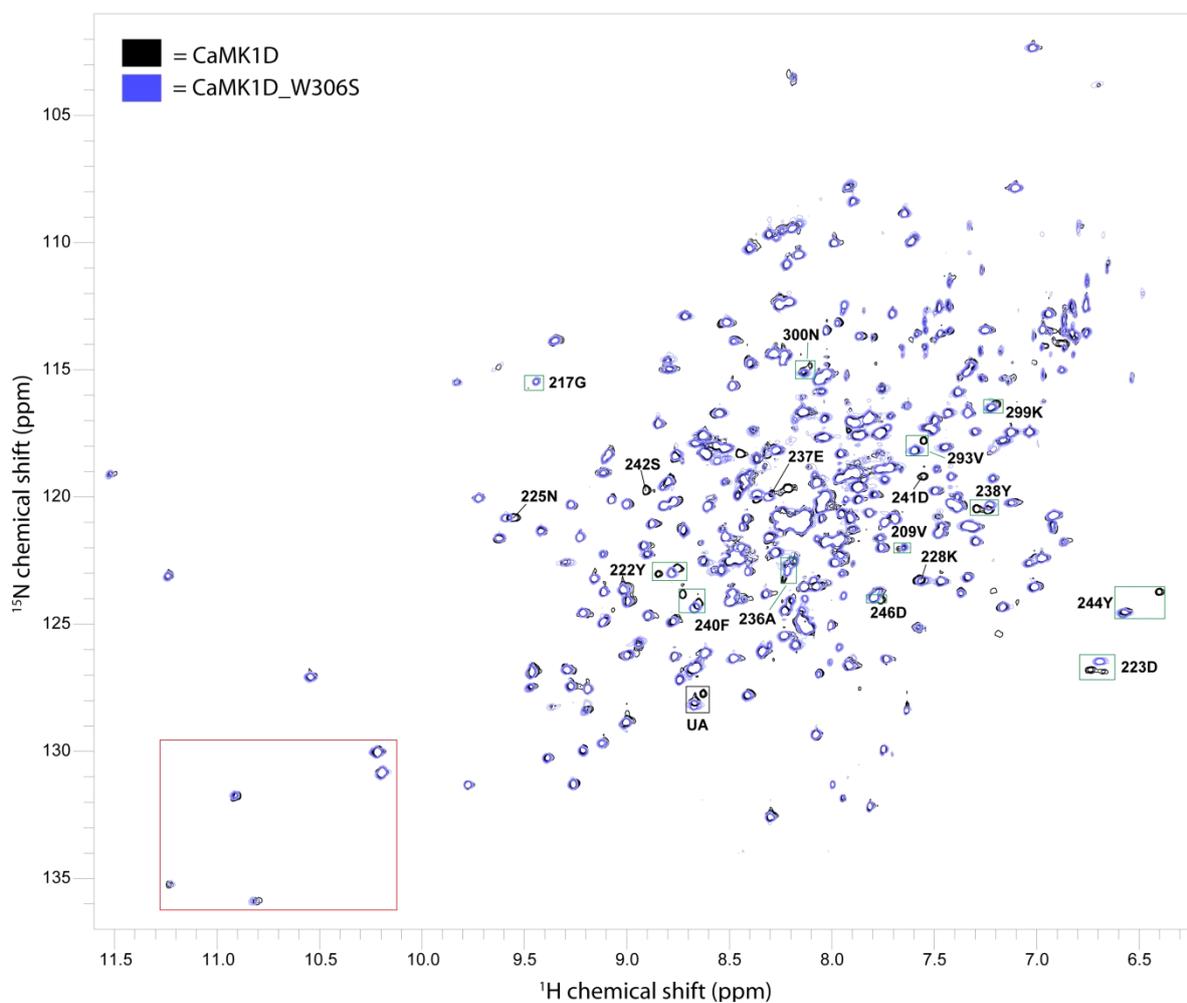
Figure 124. A comparison of the CaMK1D_W306S backbone structure with CaMK1D. The TROSY HSQC spectrum of CaMK1DW306S (blue) was superimposed with the TROSY HSQC spectrum of CaMK1D and shows that CaMK1D_W306S was still folded and 5 Trp side chain Hε are still apparent.
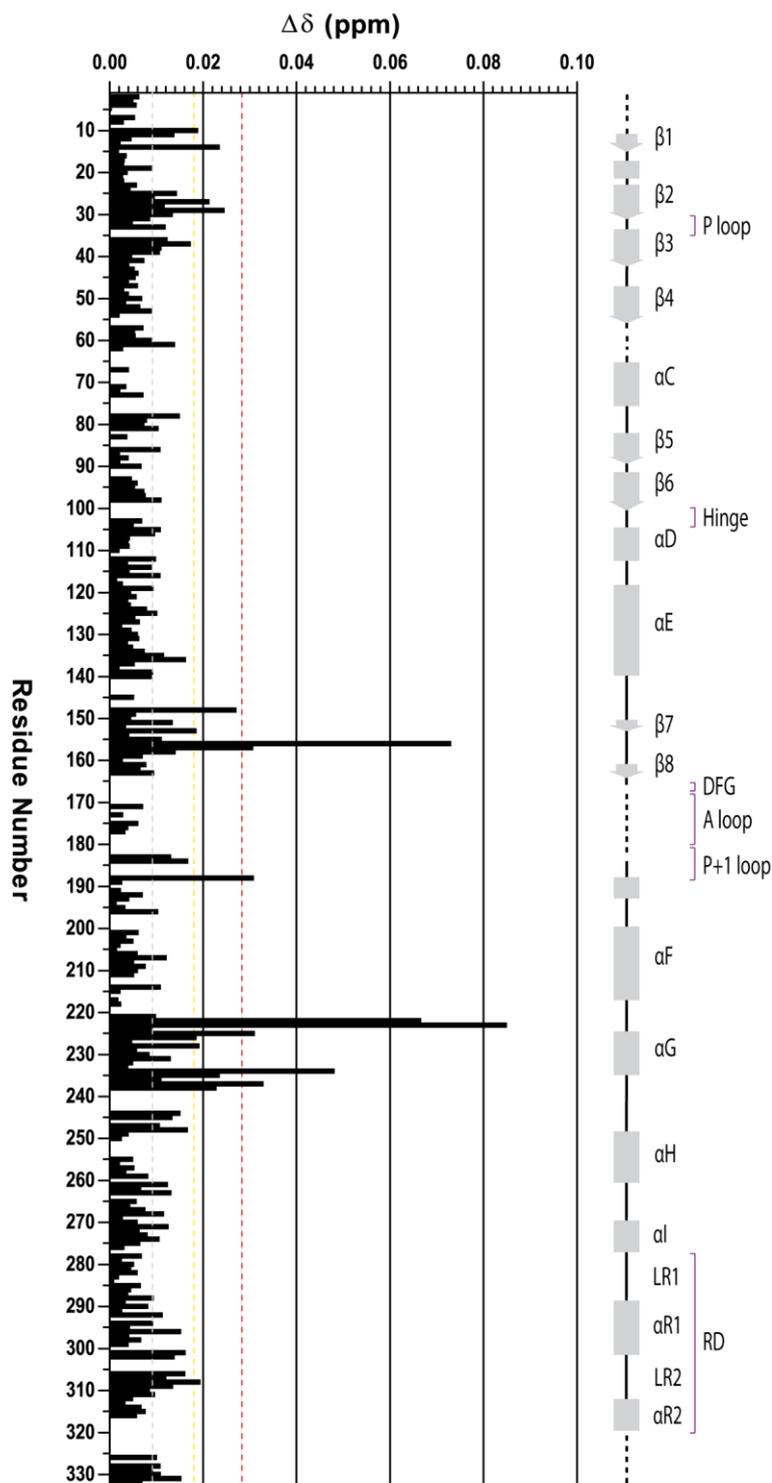
Figure 124. Histogram showing the CS difference between the backbone of CaMK1D and CaMK1D_W306S. Residues with the most pronounced CSPs ($\geq$ 2 or 1S.D) are indicated by the yellow and red line. The mean CSP is indicated by the gray line.

Thermofluor^{TM} suggests the CaMK1D-Ca^{2+}/CaM complex has a lower stability compared to apo CaMK1D

Our previous NMR and ITC data on the CaMK1D-Ca^{2+}/CaM complex suggested that the interaction between CaMK1D and Ca^{2+}/CaM was tight binding. Therefore we decided to characterise the thermal stability of the complex presuming that a tight binding complex with a Kd of 38 nM would result in an increase in the thermal stability compared to CaMK1D. However thermofluor^{TM} results (Figure 125) indicate that this was not the case because the complex exhibited a Tm of 49.3 °C whereas the Tm of CaMK1D was 51.2 °C. This indicates a slight decrease in Tm upon CaMK1D-Ca^{2+}/CaM complex formation and suggests that there is destabilization when Ca^{2+}/CaM binds and disengages the AID. Notably the Tm of apo CaMK1D is also slightly lower than its Tm of reported in its default NMR buffer
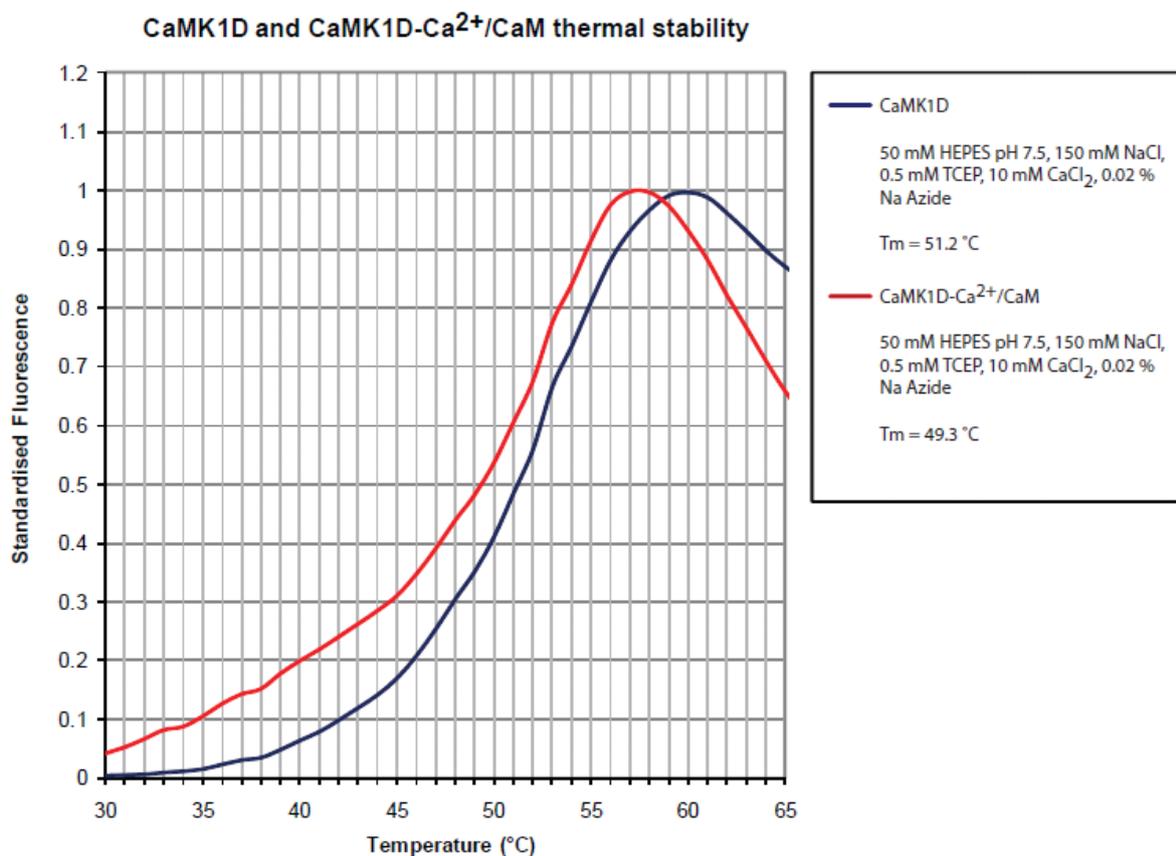
Figure 125. Thermal stability of CaMK1D and CaMK1D-Ca^{2+}/CaM. Thermofluor^{TM} was used to compare the stability of apo CaMK1D and the CaMK1D-Ca^{2+}/CaM complex. Our results indicate that the complex has a lower stability with respect to apo CaMK1D. The ΔTm between apo CaMK1D and the complex was estimated to be 1.9 °C. We postulate that the complex has lower stability due to the removal of stabilizing interactions imparted by the pseudosubstrate AID upon Ca^{2+}/CaM binding.

SAXS solution structures of CaMK1D, Ca^{2+}/CaM, and CaMK1D-Ca^{2+}/CaM complex

SAXS model of CaMK1D is monomeric and adopts a bilobal shape

In order to further understand the mechanism of CaMK1D activation by Ca^{2+}/CaM we attempted to determine the global solution structure of the CaMK1D-Ca^{2+}/CaM complex, CaMK1D and Ca^{2+}/CaM by small angle x-ray scattering (SAXS). Solution structure determination of the individual proteins using SAXS was necessary to compare global conformational changes that might occur upon complex formation and guide the construction of a SAXS derived complex. All data collected were free of aggregation as indicated by the scattering curves and Rg values from the guinier plots which did not show significant deviation from the linear region suggesting negligible particle interference and aggregation (see table for SAXS statistics). Therefore we began with characterising the SAXS model of CaMK1D. Figure109 reveals that the CaMK1D SAXS model appears to be bilobal. Superimposition of the CaMK1D crystal structure in momomeric form with the SAX model indicates that only one CaMK1D protein can fit within the SAX envelope and it closely resembles the bilobal fold of the crystal structure. The SAXS model of CaMK1D therefore indicates that CaMK1D is a monomer in solution and validates the suggestion that CaMK1D elutes as a 37.8 kDa monomer by size exclusion. Furthermore it reinforces the notion that the CaMK1D dimer evident in the crystal structure is indeed an artefact of crystallisation. Most of the core secondary structural elements from both of the lobes are enclosed within the envelope except for a small loop within the N-terminal lobe which protrudes the SAXS envelope slightly. Overall the CaMK1D crystal structure closely resembles the CaMK1D SAXS solution structure.

SAXS model of Ca^{2+}/CaM is monomeric and adopts a dumbbell shape

Structure determination of Ca^{2+}/CaM by SAXS (figure 110) reveals that the SAXS model of Ca^{2+}/CaM adopts a dumbbell shaped conformation in solution. Consequently superimposition of the Ca^{2+}/CaM crystal structure with the SAX model illustrates that they conform. Only one Ca^{2+}/CaM protein can fit into the SAX model

of Ca^{2+}/CaM which confirms that Ca^{2+}/CaM is monomeric in solution. However the SAX envelope does not completely envelope the Ca^{2+}/CaM crystal structure. In comparison to the CaMK1D SAXS structure there are more secondary structure elements in Ca^{2+}/CaM that protrude the SAX envelope. Notably these include α helices from the E-F hand motif.

SAXS model of the CaMK1D-Ca^{2+}/CaM complex adopts an elongated shape

Activation of CaMK1D requires binding of Ca^{2+}/CaM to the C-terminal CBD of CaMK1D but the exact binding mode and conformation of the active complex is not fully understood. To address this we first determined the SAXS model of the CaMK1D-Ca^{2+}/CaM complex then constructed the complex by combining SAXS models with the crystal structures of CaMK1D and Ca^{2+}/CaM. Our SAXS model of the CaMK1D-Ca^{2+}/CaM complex is shown to adopt an elongated conformation (Figure 111) thus indicating a global conformational change when compared to the SAXS models of CaMK1D and Ca^{2+}/CaM. Subsequently the SAXS model of the complex was used to guide the construction of the final complex. by manually fitting a modified crystal structure of CaMK1D lacking the AID-CBD and Ca2+/CaM in complex with the .

This structure of CaMK1D is termed CaMK1D1-288 due to removal of the C-terminal Δ289-313 segment that includes the AID and CBD. With CaMK1D1-288 we constructed the final complex using the crystal structure of Ca^{2+}/CaM in complex with an AID-CBD peptide. Manual fitting of the two crystal structures to the SAXS complex envelope reveals a model whereby Ca^{2+}/CaM-AID-CBD is positioned at the base of the CaMK1D C-terminal substrate binding lobe. This model proposes the active conformation of a CaMK protein when in complex with Ca^{2+}/CaM. Interestingly this binding mode is different from the DAPK-Ca^{2+}/CaM crystal structure which demonstrates Ca^{2+}/CaM bound to the AID-CBD but not disengaged from the substrate binding site (de Diego, Kuper et al. 2010). It is also evident in our final model that both proteins bind to each other in a one to one stoichiometry which

conforms to our ITC data and both proteins are almost completely enclosed within the SAXS envelope.

	Rg (nM)	Dmax (nM)	NSD	Chi value
CaMK1D	2.42	7.3	1.713	1.189
Ca^{2+}/CaM	2.14	6.4	1.487	2.210
CaMK1D-Ca^{2+}/CaM	3.04	10.2	1.567	1.360

Table 4. SAXS statistics
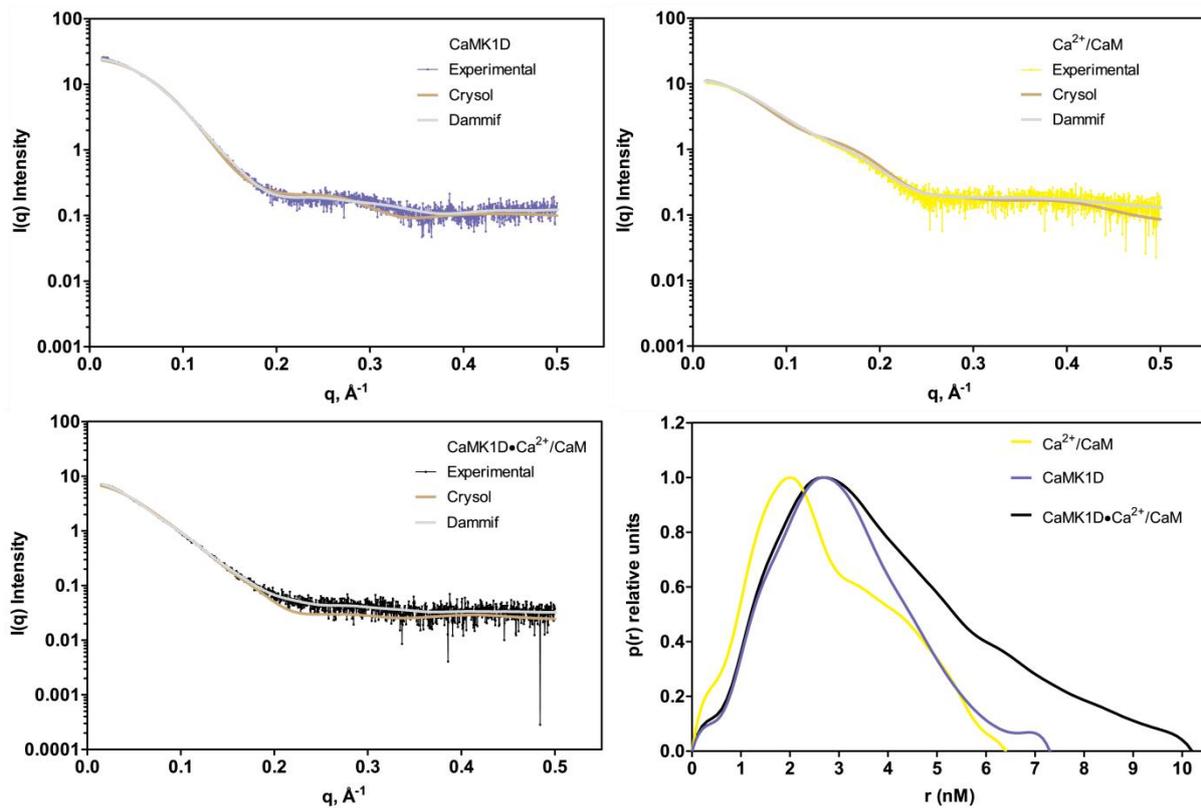
Figure 109A. SAXS scattering curves and the P(r) distribution plot for CaMK1D, Ca²⁺/CaM, and the CaMK1D-Ca²⁺/CaM complex. The scattering curves indicate that samples are not affected by aggregation or interparticle interference when q approaches 0. Furthermore all scattering curves superimpose well with their crysol and dammif derived scattering curves suggesting that the solution and crystal structure conformations are similar. The P(r) distribution plot indicates the Dmax values of each sample and as expected the complex has the largest Dmax.
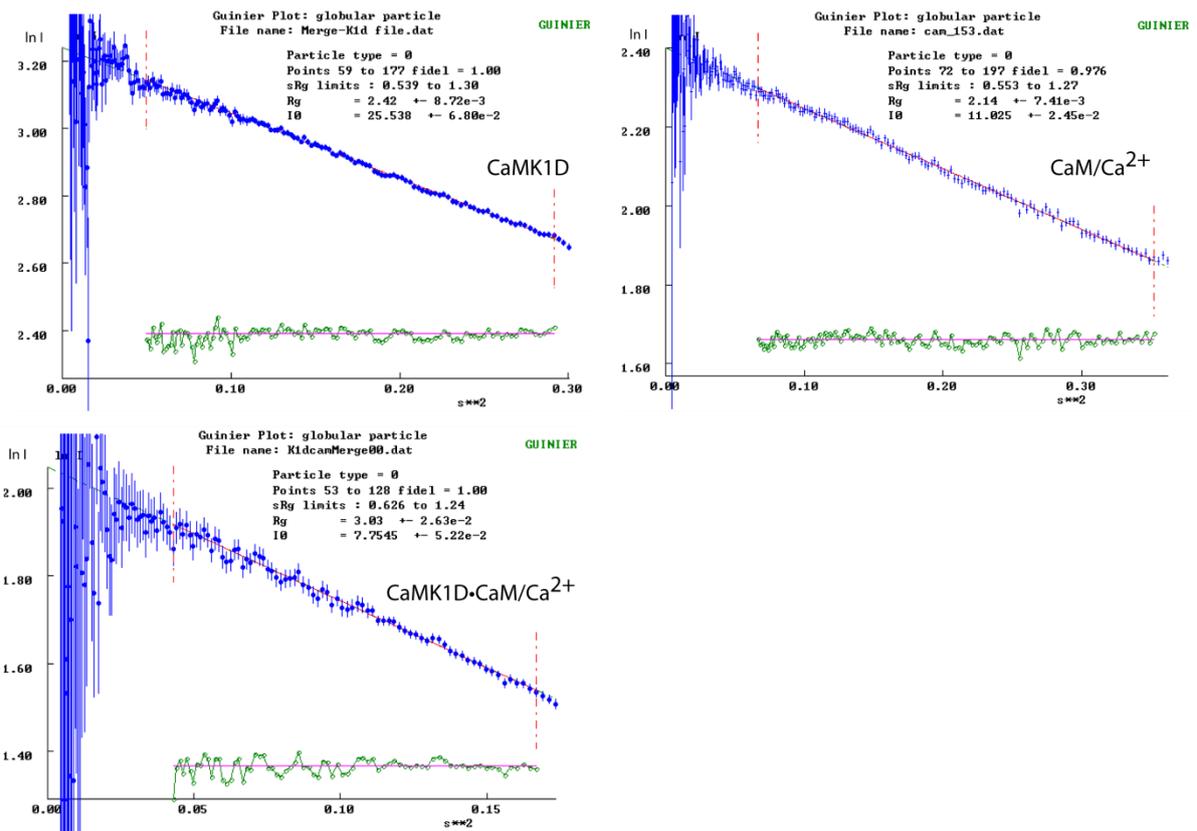
Figure 109B. Guinier plots of CaMK1D, Ca^{2+}/CaM, and the CaMK1D-Ca^{2+}/CaM complex. The guinier plots for all samples appear linear in the regions shown indicating that the samples are free of aggregration and interparticle interference aswell as producing reasonable estimates of the Rg values for each protein.
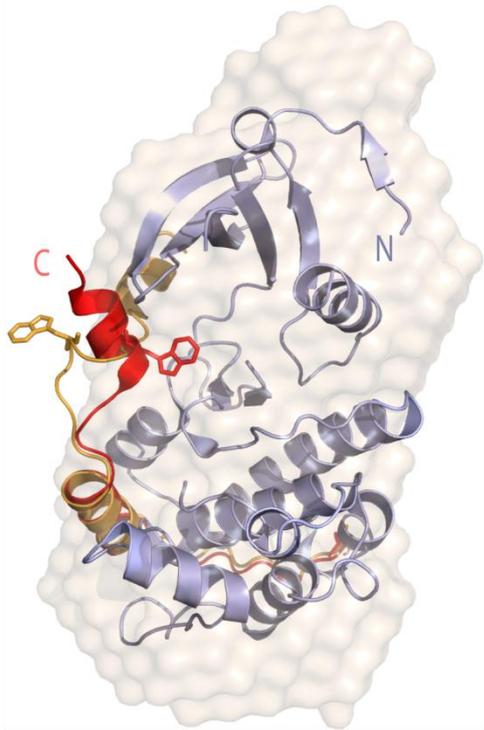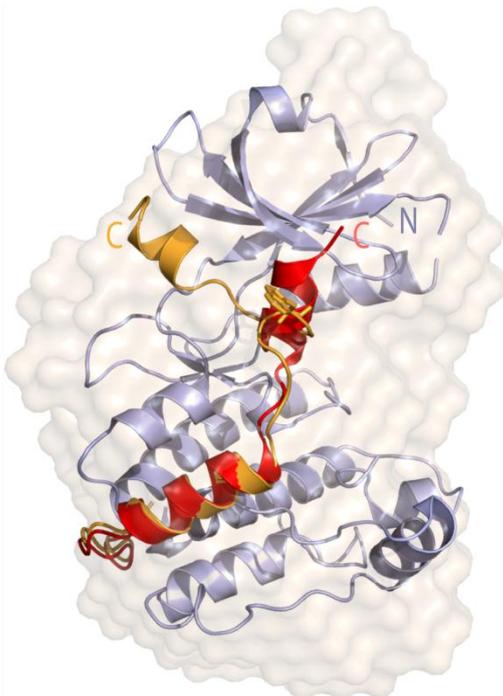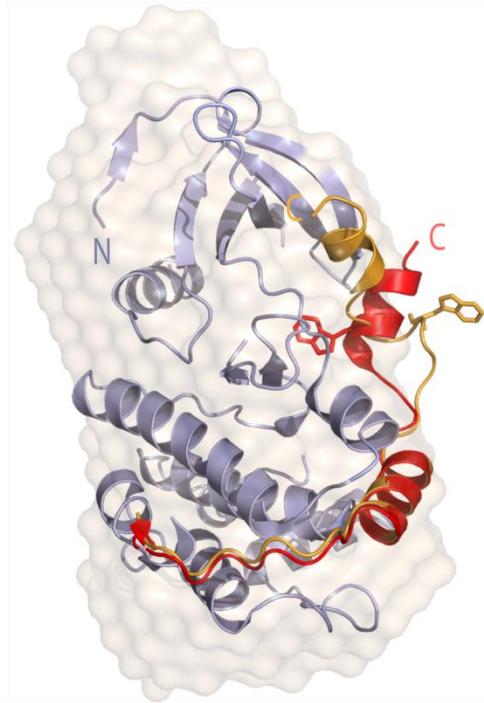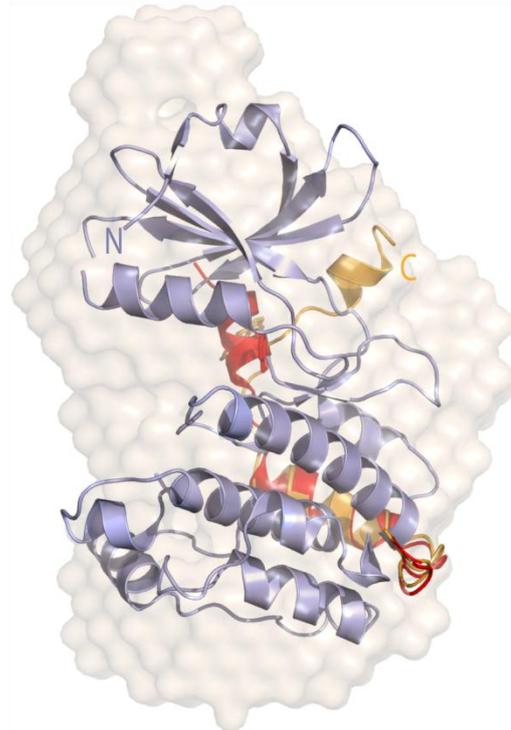
Figure 109C. CaMK1D SAXS envelope model fitted with the CaMK1D crystal structure. The CaMK1D SAXS envelope is shown as a surface representation and reveals a bilobal shaped conformation. The small N-terminal lobe (N-lobe) and large C-terminal lobe (C-lobe) of the CaMK1D (cyan) crystal structure (PDB: 2JC6) fits favourably within the SAXS envelope. This indicates that the SAXS CaMK1D solution conformation closely resembles the crystal structure conformation. The SAXS model of CaMK1D also implies that in solution CaMK1D exists as a monomer rather than the crystallised strand exchanged dimer. Similarly the RD of rat CaMK1A modelled into our CaMK1D structure by structural alignment also conforms to the SAXS envelope.
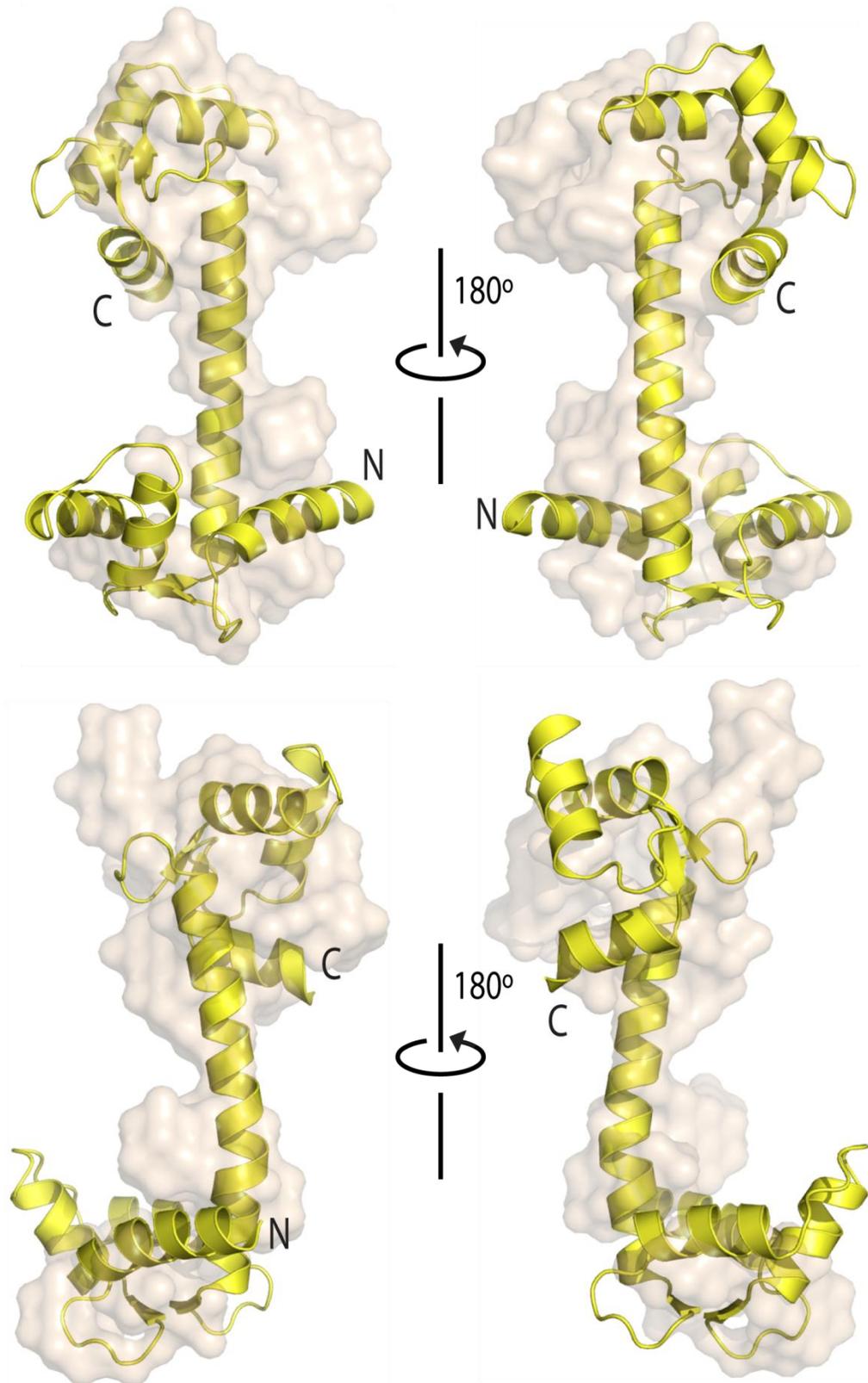
Figure 110. Ca^{2+}/CaM SAXS envelope model fitted with the Ca^{2+}/CaM crystal structure. The SAXS envelop of Ca^{2+}/CaM is shown as a surface representation and reveals a dumbbell shaped conformation that is similar to the Ca^{2+}/CaM (yellow) crystal structure (3CLN). Fitting of the Ca^{2+}/CaM crystal structure into the SAXS envelope demonstrates that both the globular domains and the flexible linker are accommodated into the envelope. Not all of the local secondary structure elements of Ca^{2+}/CaM are completely enclosed within the SAXS envelope but their domain orientations conform to the SAXS envelope and indicate that in solution Ca^{2+}/CaM exists as a monomer in an extended dumbbell conformation.
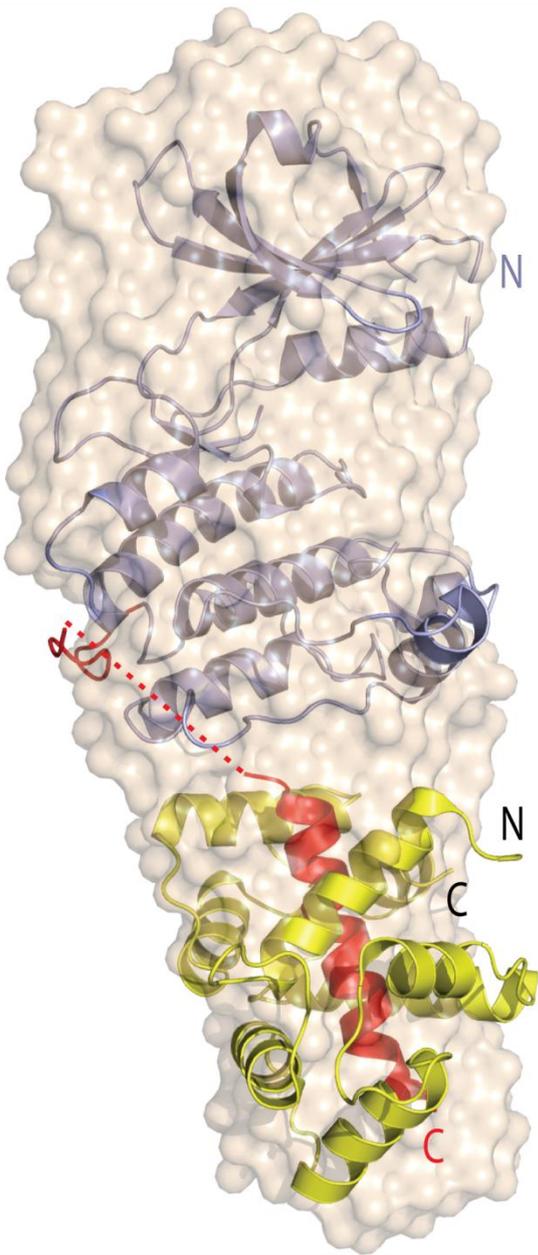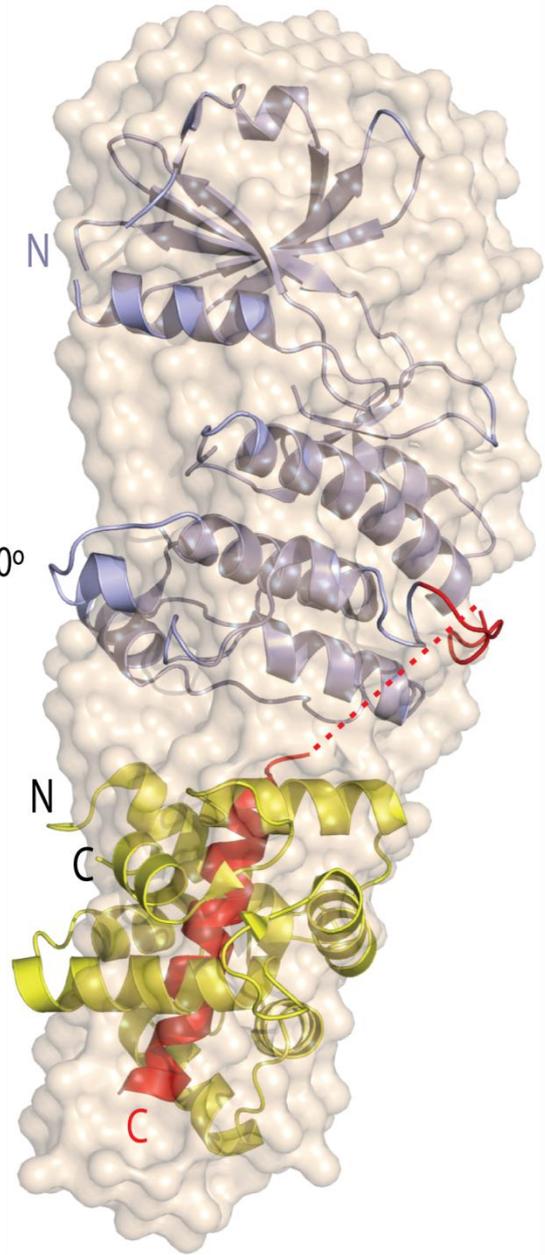
Figure 111. CaMK1D-Ca^{2+}/CaM complex reconstructed from SAXS and crystal structures. The SAXS envelope model of the complex is illustrated as a surface representation and adopts an elongated shape. The CaMK1D-Ca^{2+}/CaM complex was reconstructed using a modified version of the CaMK1D crystal structure (2JC6) in which the co-ordinates for a portion of the regulatory AID-CDB segment Δ289-313 which presumably is rearranged was deleted for simplification. The crystal structure of CaMK1D1-288 kinase domain (blue) was fitted into the top portion of the envelope whilst the crystal structure of Ca^{2+}/CaM (yellow) bound to the peptide sequence of the AID-CBD (red) (1MXE) was fitted into the lower portion of the envelope. The front, back, top and base perspectives show that the majority of the secondary structure elements and loops fit within the SAXS envelope. The dotted line is a simplified extrapolation of the AID loop to the loop of the AID-CBD peptide. The AID-CBD peptide sequence is taken from the rat homolog of CaMK1A and corresponds to: IKKNFAKSKWKQAFNATAVVRHMRK (residues 294 -318) (Clapperton, Martin et al. 2002).
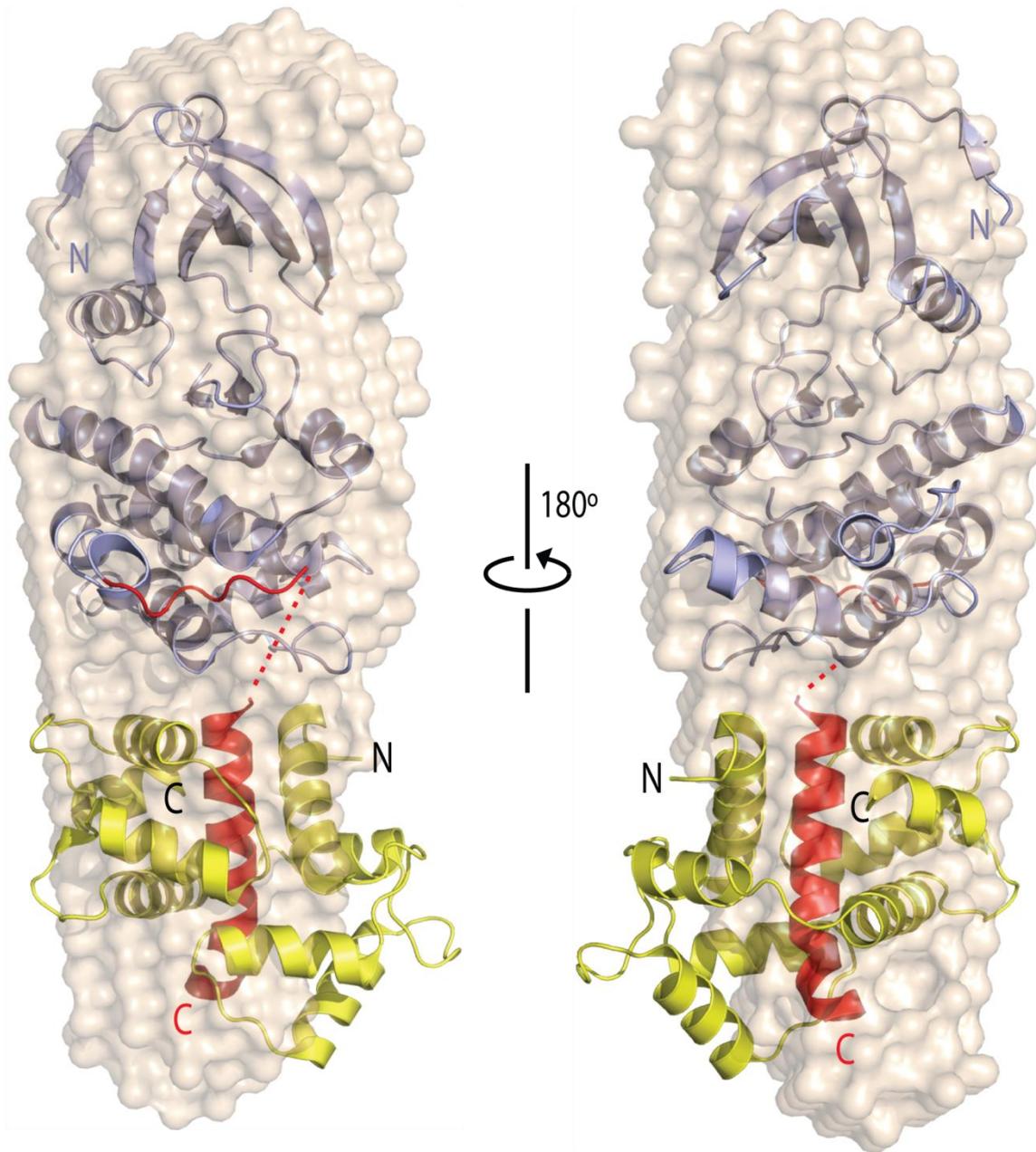
Figure 112. CaMK1D-Ca^{2+}/CaM complex reconstructed from SAXS and crystal structures. Model of the CaMK1D-Ca^{2+}/CaM complex calculated using the SAXS envelope and fitting of the crystal structures of CaMK1D1-288 and Ca^{2+}/CaM-AID-CBD as described and shown in another perspective similar to those shown in figure. The CaMK1D1-288 (blue) and Ca^{2+}/CaM-AID-CBD crystal structures (yellow and red) was fitted into the top and lower portions of the SAXS envelope, respectively.

CaMK1D and small molecule ligand interactions by NMR

Apo autoinhibited CaMK1D binds to ATP with low affinity

The crystal structure of rat CaMK1A determined by Goldberg and co-workers first revealed that the canonical kinase bilobal fold was autoinhibited by the AID via a pseudosubstrate mechanism that occluded the substrate binding site and disrupted the conformation of the N-terminal ATP binding lobe (Goldberg, Nairn et al. 1996). However the possibility of ATP binding cannot be ruled out because of the open conformation adopted by rat CaMK1A which results in solvent exposure of the active site cleft (Goldberg, Nairn et al. 1996). Structural alignment of CaMK1D with rat CaMK1A indicates that the active site cleft of autoinhibited CaMK1D also adopts a similar conformation. We postulated that although CaMK1D is autoinhibited it could still bind ATP but with reduced affinity compared to active CaMK1D-Ca^{2+}/CaM. Furthermore the fact that CaMK1D was co-crystallised with GSK-3 XIII in the active site cleft indicated that the site was accessible and solvent exposed.

We used an NMR based approach to study the interaction of AMPPNP (a non-hydrolysable ATP analog) with CaMK1D by monitoring the CSPs of peaks in the ^{1}H, ^{15}N TROSY HSQC to infer residues that interact directly or indirectly with AMPPNP. Our results (Figure 126) indicate that titration of CaMK1D with increasing concentrations of AMPPNP caused a number of peaks to become progressively perturbed in a fast exchange which suggests that the affinity of the interaction is within the weak-moderate affinity range. Due to the nature of fast exchange, peaks which experience this NMR timescale can be tracked and their chemical shift change can be determined. In order to estimate the binding affinity, titrations were conducted until saturation with the ligand in excess to determine the endpoint where there are no longer any observed chemical shift changes. Consequently the K_d for the binding interaction can be estimated from a binding curve by plotting the normalised chemical shift changes of a residue against the increasing ligand concentrations. Using this method our K_d estimate for CaMK1D binding to AMPPNP was 1.306$\pm$0.093 mM, indicative of very weak binding (figure 129).

A histogram of the CSPs induced by AMPPNP (figure 127) binding reveals the residues which undergo the most significant perturbations span predominantly across the N-terminal ATP binding lobe. These residues correlate closely in sequence to three distinct functional regions. Both Val 38 and Ala 40 are located near in sequence to the p-loop which interacts with the phosphate groups of nucleotides i.e. ATP and ADP, whereas Val 51 and Ile 54 are close to Lys 52, a highly conserved Lys that interacts with the nucleotide and is essential for catalysis of the phosphoryl transfer reaction. Glu 89, Ser 90, Leu 94, and Tyr 95 are all located N-terminal of the hinge, a loop region (residues 100-104) which is implicated in hydrogen bonding to the adenine ring of nucleotides. Binding of AMPPNP presumably causes the hinge to convey long range conformational changes to Glu 89, Ser 90, Leu 94, and Tyr 95. A few residues located outside of the N-terminal lobe also exhibit significant CSPs and include Lys 14, Ile 19, Glu 112, and Leu 150. Analysis of these CSPs mapped onto the CaMK1D structure (figure 128) indicates that Lys 14, Ile 19, and Glu 112 are perturbed indirectly via conformational changes induced by AMPPNP binding. Glu 112 in particular appears to be a long range downstream effect of the hinge. Leu 150 appears close in space to the active site cleft where AMPPNP binds.
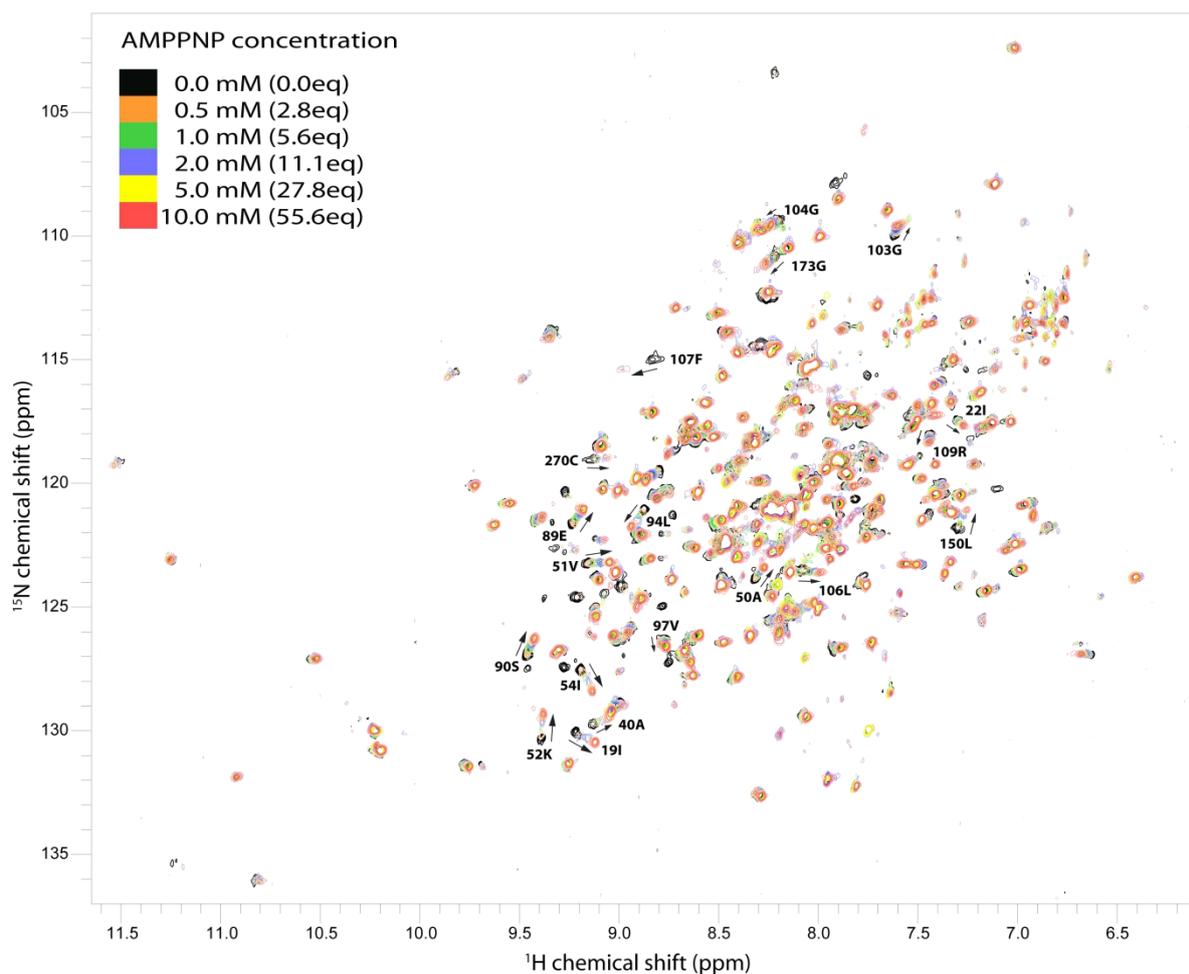
Figure 126. CaMK1D titration with AMPPNP. The superimposed ^{1}H, ^{15}N TROSY HSQC spectra shows the progressive CSPs that occurred when increasing concentrations of AMPPNP was titrated into CaMK1D. A few notable CSPs are highlighted. Titration experiments were performed using 170 uM ^{2}H, ^{15}N CaMK1D, in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide, 10 % ^{2}H$_2$O, at 25 °C

,
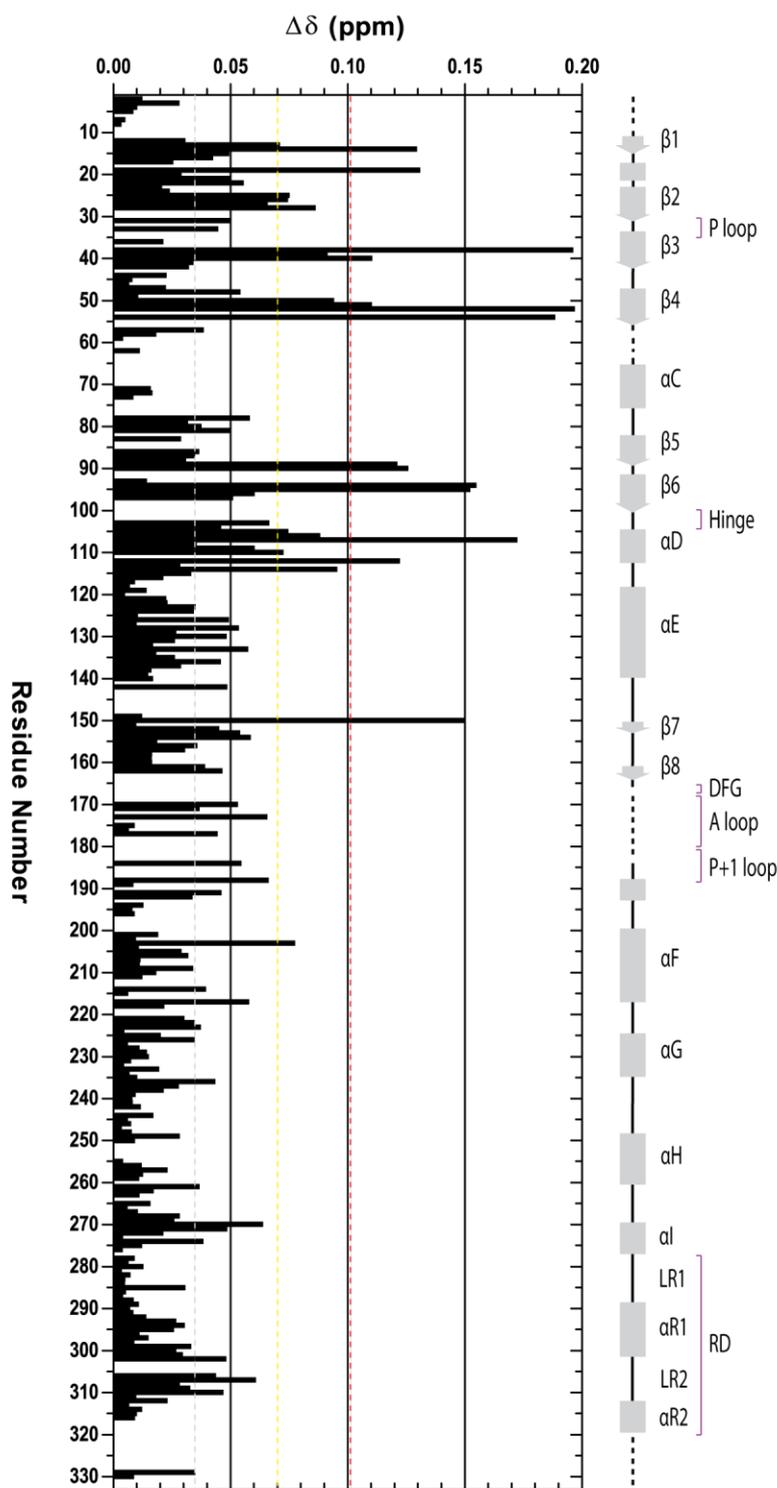
Figure 127. Histogram of the CaMK1D CSPs induced by AMPPNP. The histogram shows the residues which undergo CSPs due to AMPPNP binding. Residues with the most pronounced CSPs ($\geq$ 2 or 1S.D) are indicated by the yellow and red line. The mean CSP is indicated by the gray line.
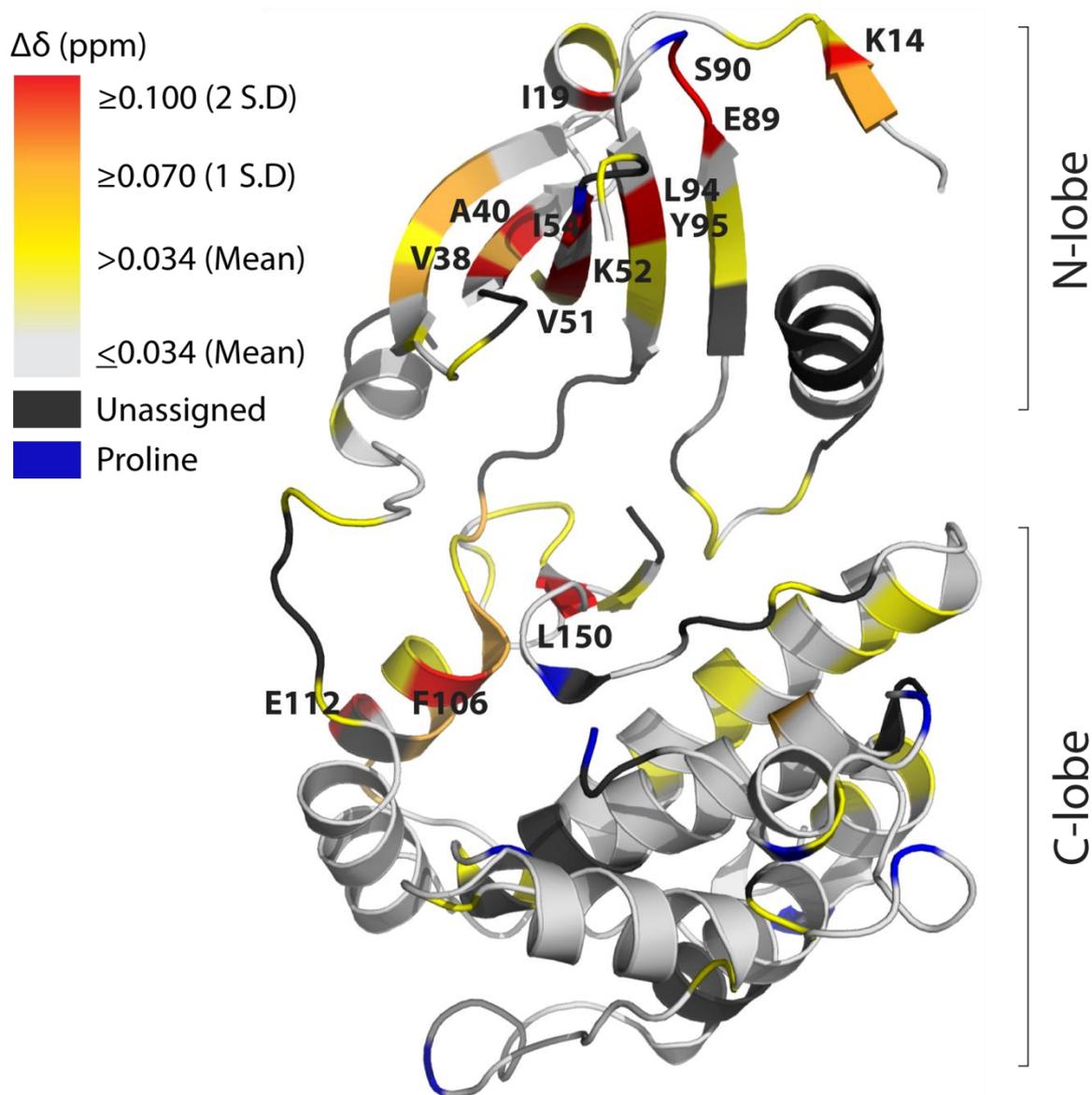
Figure 128 A. CaMK1D CSPs induced by AMPPNP. A cartoon of our CaMK1D autoinhibited model structure with CSPs induced by AMPPNP mapped on. As indicated by the gradient scale, backbone residues are coloured according to the magnitude of the normalised chemical shift changes that occurred upon titrating in AMPPNP.
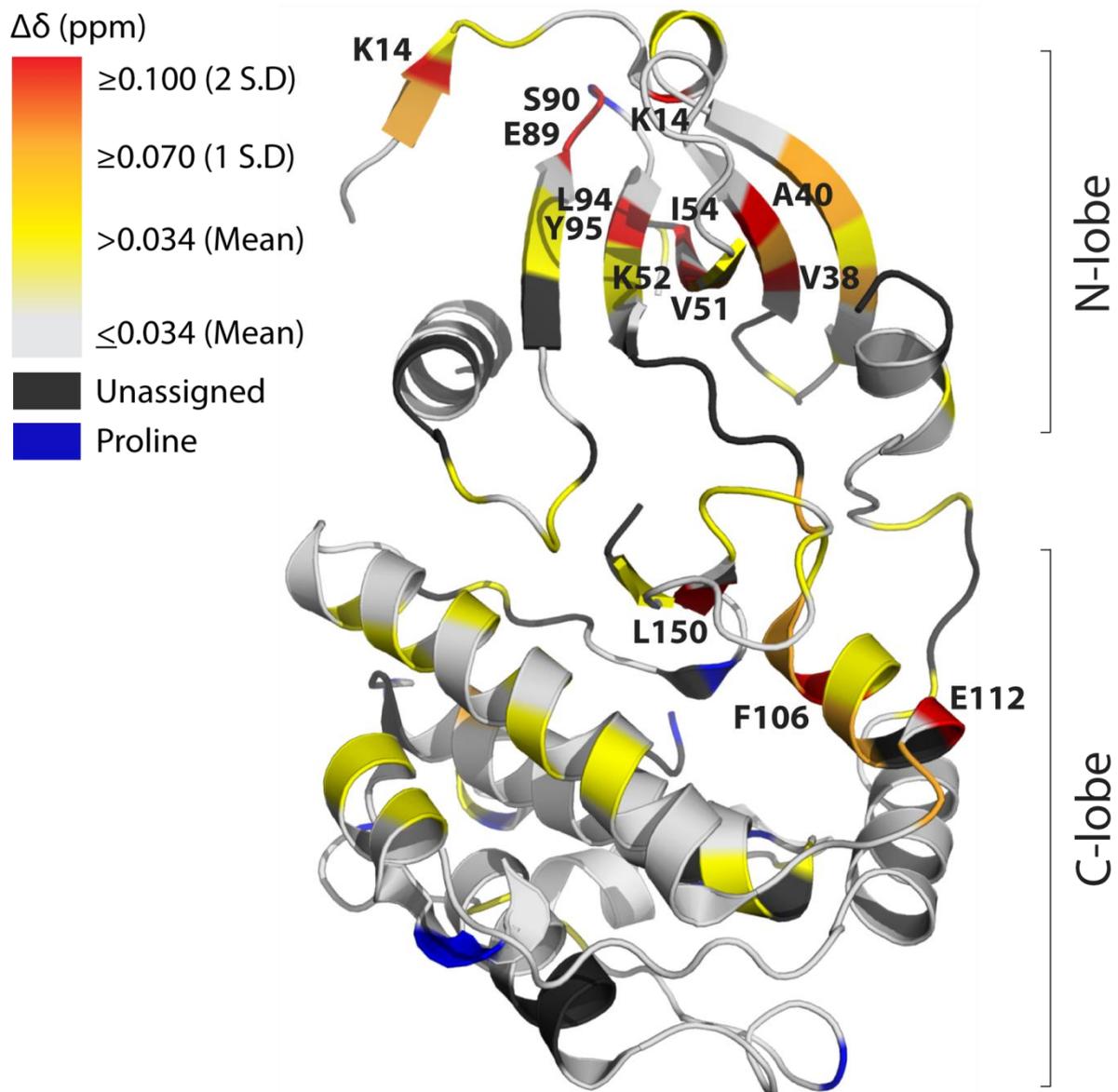
Figure 128 B. CaMK1D CSPs induced by AMPPNP. A 180° rotated view of the previous figure illustrating a cartoon of the CaMK1D crystal structure with CSPs induced by AMPPNP mapped on with levels of perturbation corresponding to the gradient scale indicated.
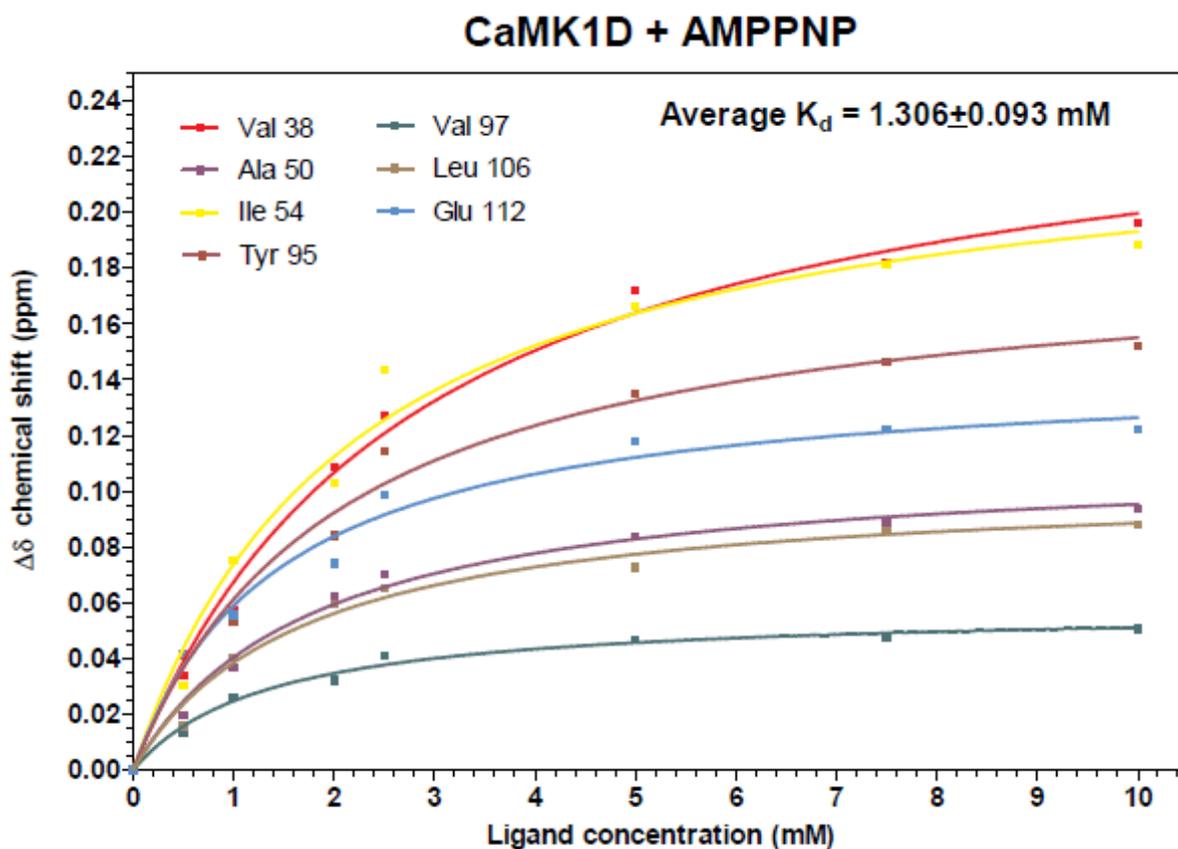
Figure 129. CaMK1D binds to AMPPNP with a K_d of 1.3 mM. The magnitude of normalised chemical shift change (y axis) for various perturbed CaMK1D residues are plotted against increasing AMPPNP concentration (x axis) to produce binding curves. K_d estimates were deduced from the binding curves and averaged to produce a weak K_d value of 1.306±0.093 mM.

CaMK1D in the presence of AMPPNP binds Mg^{2+} with a low affinity

Most kinase proteins including CaMK1D require Mg^{2+} in complex with ATP for kinase activity. In order to characterise the binding of AMPPNP-Mg^{2+} to CaMK1D we titrated $MgCl_2$ into CaMK1D saturated with AMPPNP and used the same ^1H, ^{15}N TROSY HSQC approach in our AMPPNP and CaMK1D interaction to monitor the CSPs. Our results illustrate (figure 130) that Mg^{2+} interacts with CaMK1D-AMPPNP to form CaMK1D-AMPPNP-Mg^{2+} due to evidence of CSPs. These perturbed peaks displayed progressive CSPs characteristic of fast exchange phenomena which allowed a binding curve to be plotted as previously described for our CaMK1D and AMPPNP interaction. From our binding curve (figure 133) we estimate the K_d for CaMK1D and AMPPNP-Mg^{2+} interaction to be 0.611±0.133 mM. This approximately a two-fold increase in affinity compared to the CaMK1D interaction with AMPPNP.

A comparison of the CSPs from the titration of $MgCl_2$ into CaMK1D-AMPPNP with those from the titration of AMPPNP into CaMK1D reveals a common set of peaks that are perturbed in both experiments. Some of the CaMK1D-AMPPNP CSPs induced by Mg^{2+} appear to move in an opposite direction to the CSPs induced by AMPPNP. This suggests that Mg^{2+} modulates the binding interaction of AMPPNP bound to CaMK1D by binding to AMPPNP to form the AMPPNP-Mg^{2+} complex. The Mg^{2+} induced CaMK1D-AMPPNP CSPs which move in an opposite direction to the CaMK1D CSPs induced by AMPPNP suggests that these residues in CaMK1D-AMPPNP undergo a conformational change that is similar to its conformation prior to AMPPNP binding to CaMK1D (i.e. a similar chemical environment) because the peaks move back towards the position from which they originally started from .

Peaks that display this behaviour correlate to residues Ala 40 Lys 52, Ile 54, 89E, Ser 90, 150 Leu, and Gly 173. Titration of $MgCl_2$ also caused additional peaks to undergo CSPs suggesting that the AMPPNP-Mg^{2+} complex forms additional interactions with CaMK1D which is expected because Mg^{2+} is known to coordinate further interactions between ATP and the activation loop (adams) Additional perturbed peaks include those assigned as Thr 31, Glu 79, Gly 105, Glu 112 and

Phe 115. Thr 31 and Gly 105 are located within the p-loop and hinge respectively indicating that Mg^{2+} via AMPPNP-Mg^{2+} mediates further interactions with these functional elements. Titration experiments were also conducted using up to 10 mM $MgCl_2$ in the absence of AMPPNP (data not shown) to investigate any AMPPNP independent Mg^{2+} binding sites. Consequently no CSPs were evident indicating that $MgCl_2$ does not interact with CaMK1D in the absence of AMPPNP.

Minimal CaMK1D CSPs due to buffer differences

All of our titration experiments by NMR were conducted in 50 mM HEPES buffer pH 7.0 instead of 50 mM Na Phosphate pH 7.0 because the intention was to repeat the same titration experiments on the active CaMK1D-Ca^{2+}/CaM complex which was not compatible with Na phosphate due to the need for Ca^{2+} which reacts with phosphate and precipitates. An overlay of the CaMK1D ^1H, ^{15}N TROSY HSQC spectra in both respective buffers indicated minimal CSPs.
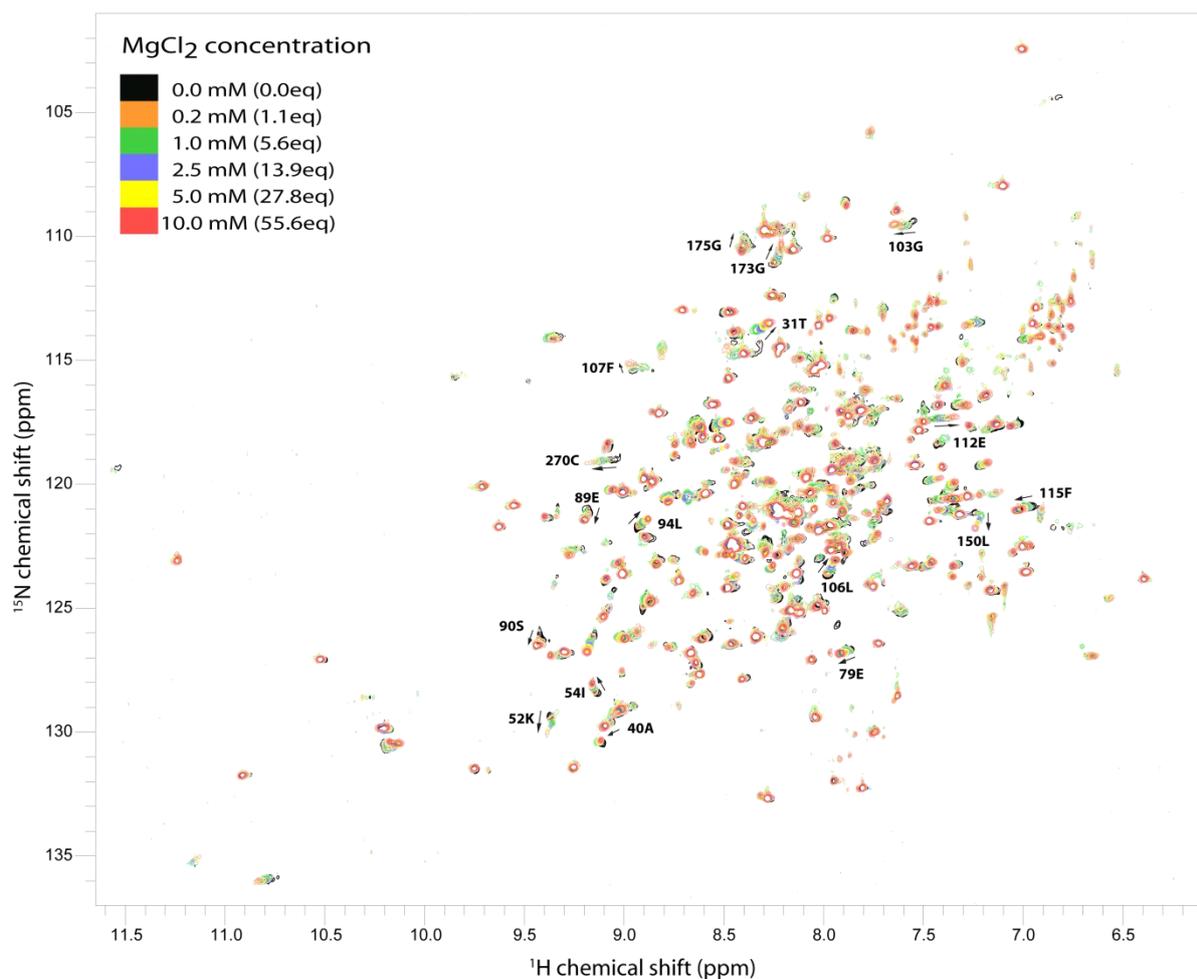
Figure 130. CaMK1D-AMPPNP titration with MgCl₂. The superimposed ^{1}H, ^{15}N TROSY HSQC spectra shows the progressive CSPs that occurred when increasing concentrations of MgCl₂ was titrated into CaMK1D saturated with 10 mM AMPPNP. A few notable CSPs are highlighted. Titration experiments were performed using 180 uM ^{2}H, ^{15}N CaMK1D, in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide, 10 % ^{2}H₂O, on a Varian 800 MHz at 25 °C
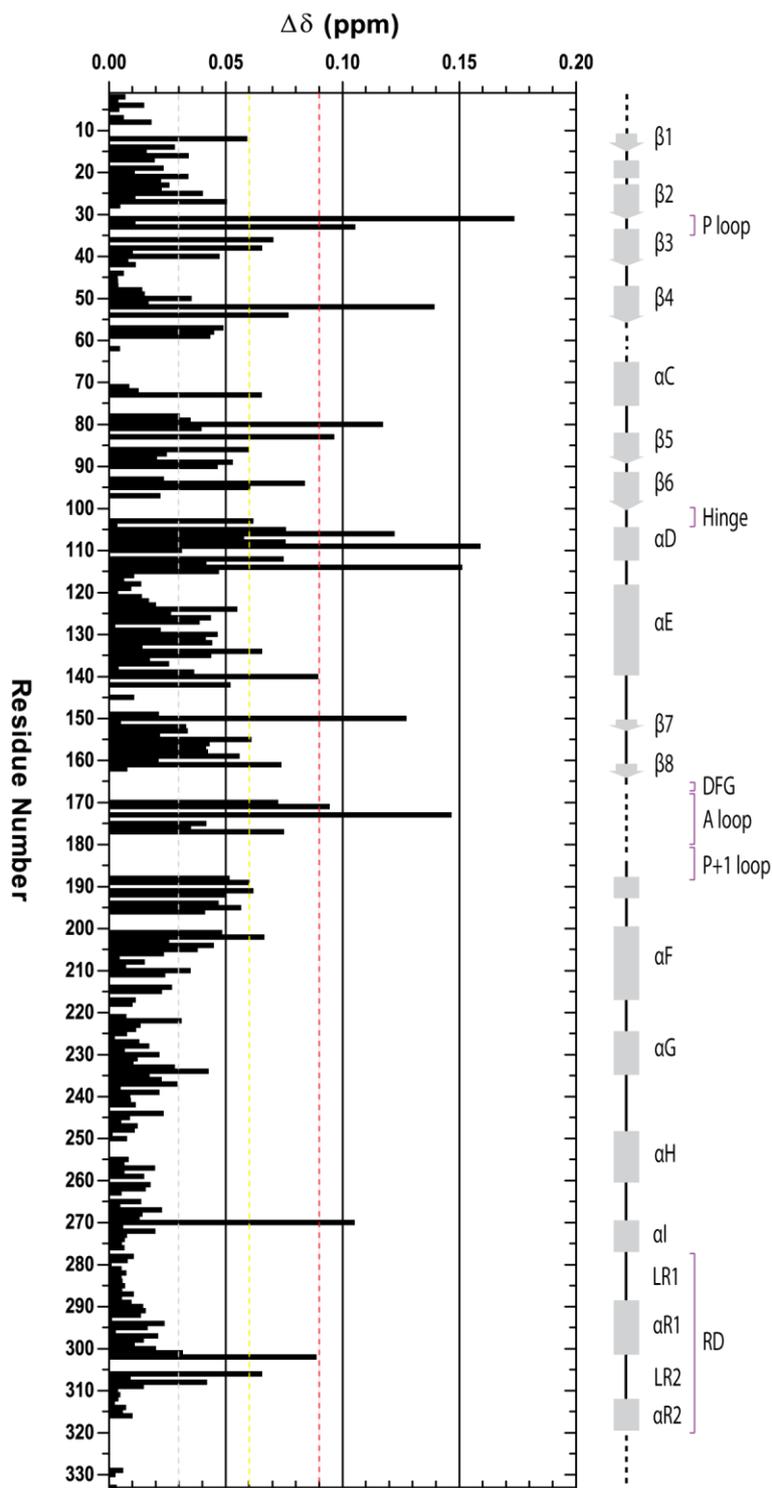
Figure 131. Histogram of CSPs for CaMK1D-AMPPNP induced by Mg^{2+}. The histogram shows the residues which undergo CSPs due to Mg^{2+} binding to CaMK1D-AMPPNP. Residues with the most pronounced CSPs ($\geq$ 2 or 1S.D) are indicated by the yellow and red line. The mean CSP is indicated by the gray line.
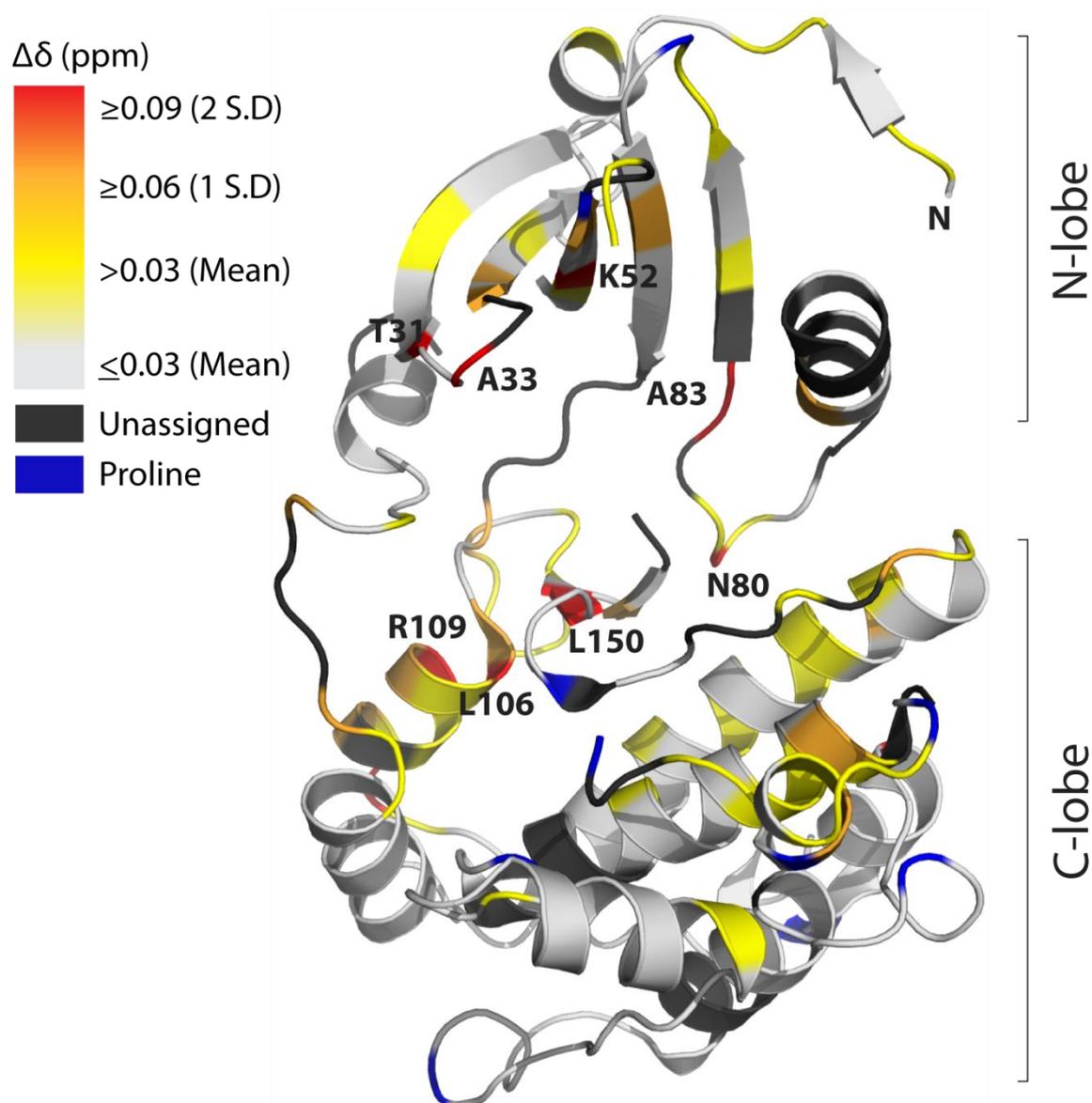
Figure 132 A. CaMK1D-AMPPNP CSPs induced by MgCl₂. A cartoon representation of the CaMK1D crystal structure with CSPs induced by Mg^{2+} mapped on. As indicated by the gradient scale, backbone residues are coloured according to the magnitude of the normalised chemical shift changes that occurred upon titrating in $MgCl_2$.
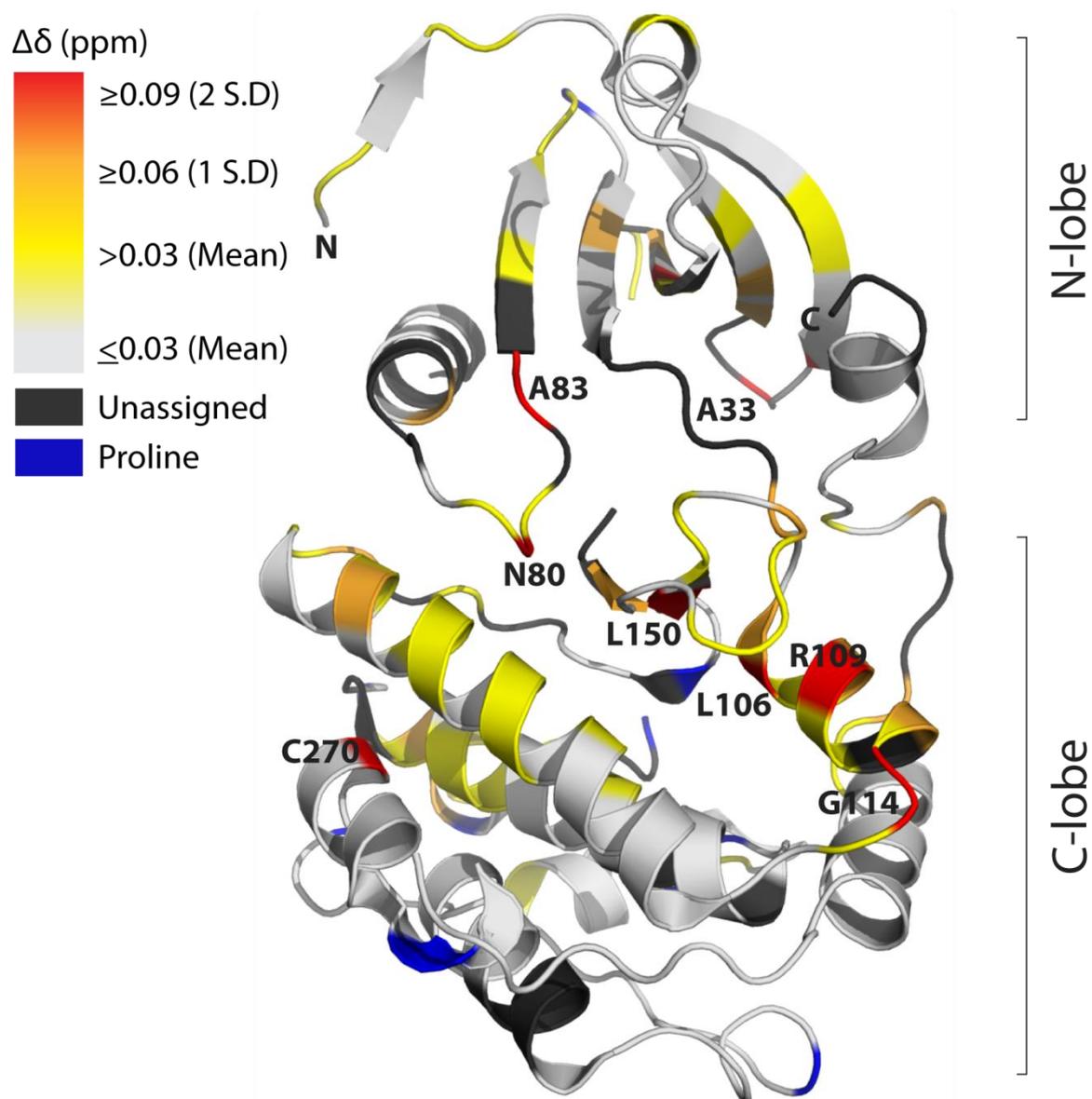
Figure 132 B. CaMK1D-AMPPNP CSPs induced by MgCl₂. A 180° rotated view of the previous figure illustrating a cartoon of the CaMK1D crystal structure with CSPs induced by Mg²⁺ mapped on.
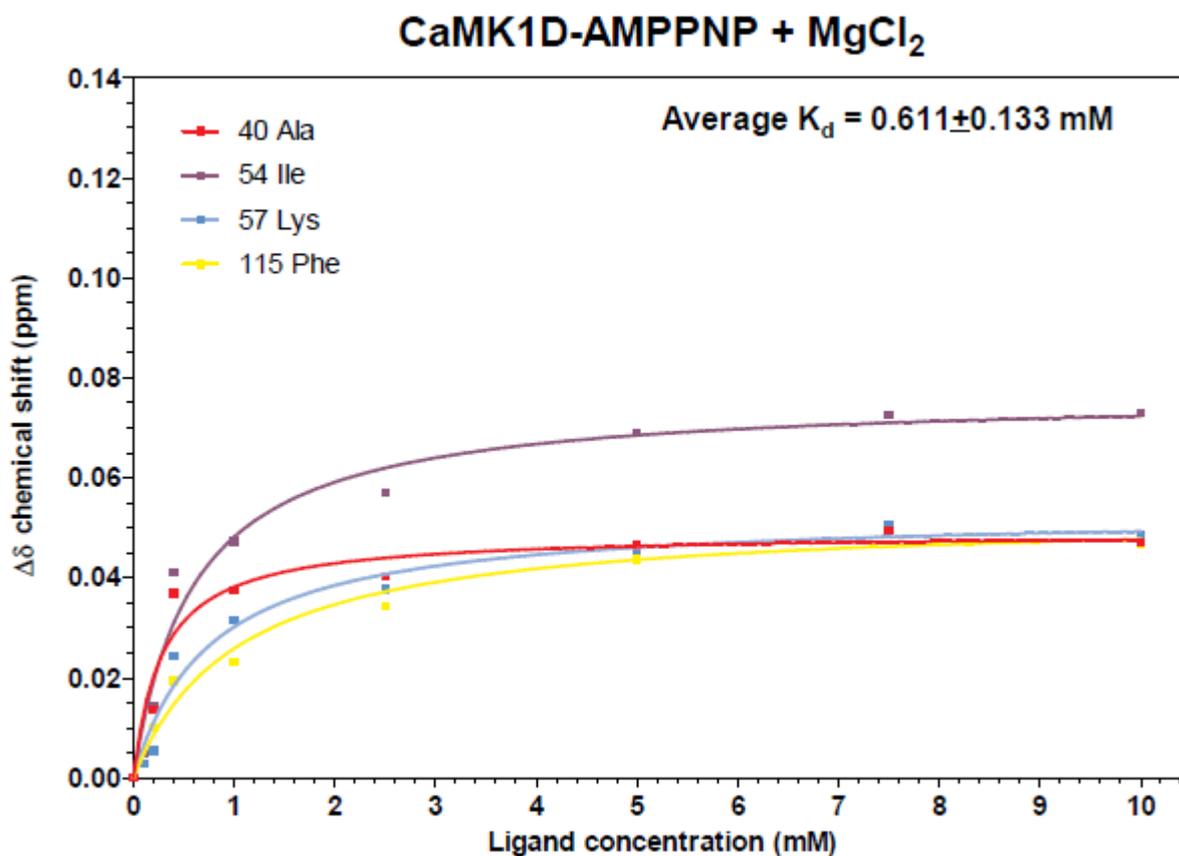
Figure 133. CaMK1D binds AMPPNP-MgCl$_2$ with a K$_d$ of 0.611 mM. The magnitude of normalised chemical shift change (y axis) for various perturbed CaMK1D residues are plotted against increasing MgCl$_2$ concentration (x axis) to produce binding curves. K$_d$ estimates were deduced from the binding curves and averaged to produce a weak K$_d$ value of 0.611±0.133 mM for the interaction between CaMK1D and AMPPNP-Mg^{2+}.

CaMK1D binds several kinase inhibitors by thermofluor^{TM}

In addition to providing information about the thermal stability of proteins, thermofluor^{TM} can also indicate binding to ligands because often ligand binding results in an increase in the thermal stability of a protein depending on its affinity. In collaboration with the SGC, CaMK1D was screened against a library of kinase inhibitors using thermofluor^{TM} and several hits were discovered as shown in table 5. The data indicates that the active site cleft is accessible to various ATP analog inhibitors. Inhibitors that produced larger changes in Tm suggested tighter binding.

5 % DMSO lowers the thermal stability of CaMK1D and causes a few CSPs

Most of the inhibitors (table) that were suggested to bind to CaMK1D by thermofluor^{TM} required dimethlysulfoxide (DMSO) for solubility. Therefore prior to characterisation of such ligands by NMR we decided to deduce an optimal DMSO % (v/v) for CaMK1D stability and established this to be 5.0 % by thermofluor^{TM}. Our results indicated a 5.0 °C decrease in the Tm of CaMK1D when in the presence of 5.0 % DMSO (Figure 134). At 10.0 % DMSO, sigmoidal melting curves could not be obtained due to precipitation of CaMK1D (data not shown). CaMK1D stability in the presence of 5.0 % DMSO was further characterised by NMR in which CSPs were monitored by titrating into CaMK1D 0, 1.0, 2.5, and 5.0 % DMSO (figure 135). Results from this titration experiment indicate that CaMK1D is still folded in 5.0 % DMSO. Small progressive CSPs in fast exchange were also evident. Tracking these CSPs allowed us to account for the CSPs induced by 5.0 % DMSO when conducting titrations with DMSO soluble inhibitors by NMR in our CaMK1D. Mapping of the DMSO interaction sites on CaMK1D suggests that DMSO binds to two main patches one of which is located near Thr 28 and Ala 33 near the C-terminal portion of the R2 helix of the AID/CBD and at the base of the C-terminal lobe (figure 138). We postulate that at 10.0 % DMSO or above, there is disruption and destabilisation of the local structure of these two patches caused by DMSO binding. Destabilisation of the AID/CBD αR2 helix in particular will impact the overall stability of CaMK1D because this helix contributes towards stabilizing contacts with the CaMK1D catalytic domain.
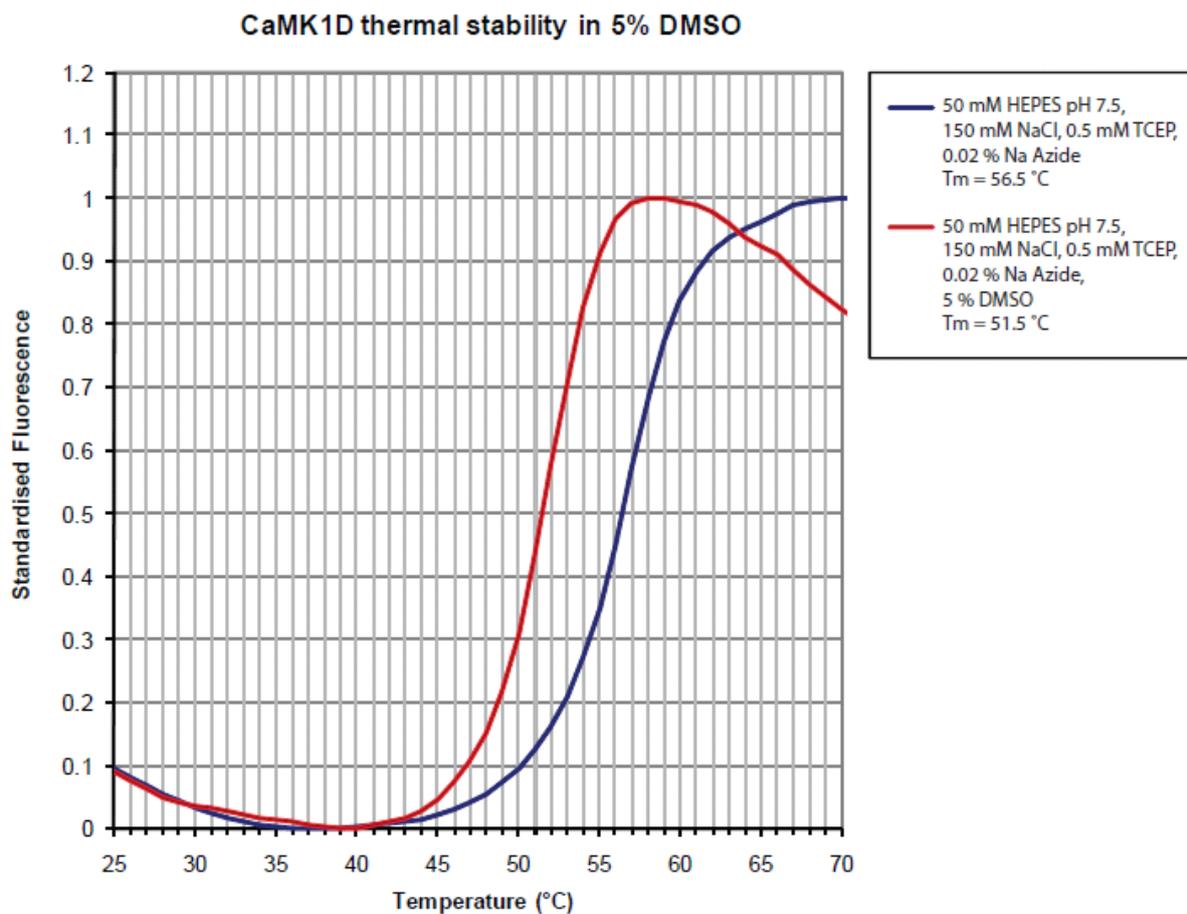
Figure 134. CaMK1D thermal melting curves in the presence of 5 % DMSO. The thermal stability of CaMK1D was assessed by thermofluor^{TM} in the presence of 5 % DMSO and found to decrease by 5.0 °C.

Compound Description	Compound Screen: Tm Shift (C)
Staurosporine	11
Syk Inhibitor II	5.3
SB 218078	4.6
IKK Inhibitor VII	4.3
K252c; Staurosporine Aglycone	4.2
JAK Inhibitor I	4.2
K252a	3.8
GSK inhibitor XIII	3.5
IKK Inhibitor VI	3.4
Cdk4 Inhibitor	3.4
Quercetin	3.3
Cdk1/2 Inhibitor III	3.2

Table 5. Compounds that bind to CaMK1D. Our collaborators the SGC conducted a high throughput screen of known kinase inhibitors against CaMK1D by thermofluor^{TM} and discovered several interacting compounds. Changes in the Tm of CaMK1D induced by each compound are listed to illustrate their effect on CaMK1D stability.
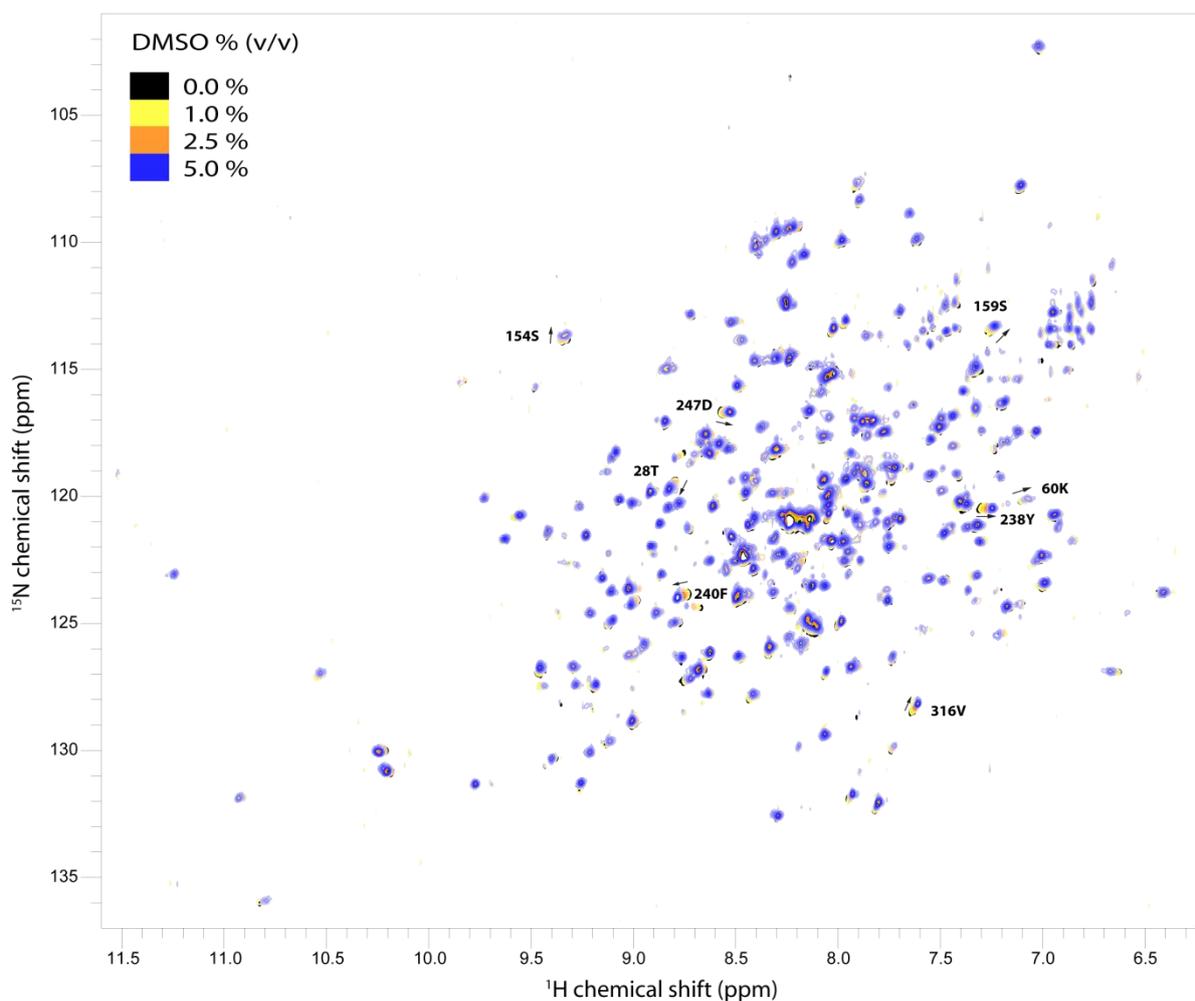
Figure 135. CaMK1D titration with DMSO. Increasing concentrations of DMSO from 0-5 % (v/v) was titrated into CaMK1D and CSPs were monitored by ^{1}H, ^{15}N TROSY HSQC. A few peaks display progressive perturbations in fast exchange. At DMSO concentrations beyond 5 %, CaMK1D precipitation accelerated greatly and compromised the quality of the spectrum. Spectra were acquired on a Varian 800 MHz at 25 °C using 180 μM ^{2}H, ^{15}N CaMK1D, in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide with 1, 2.5, and % 5 DMSO titrated in.
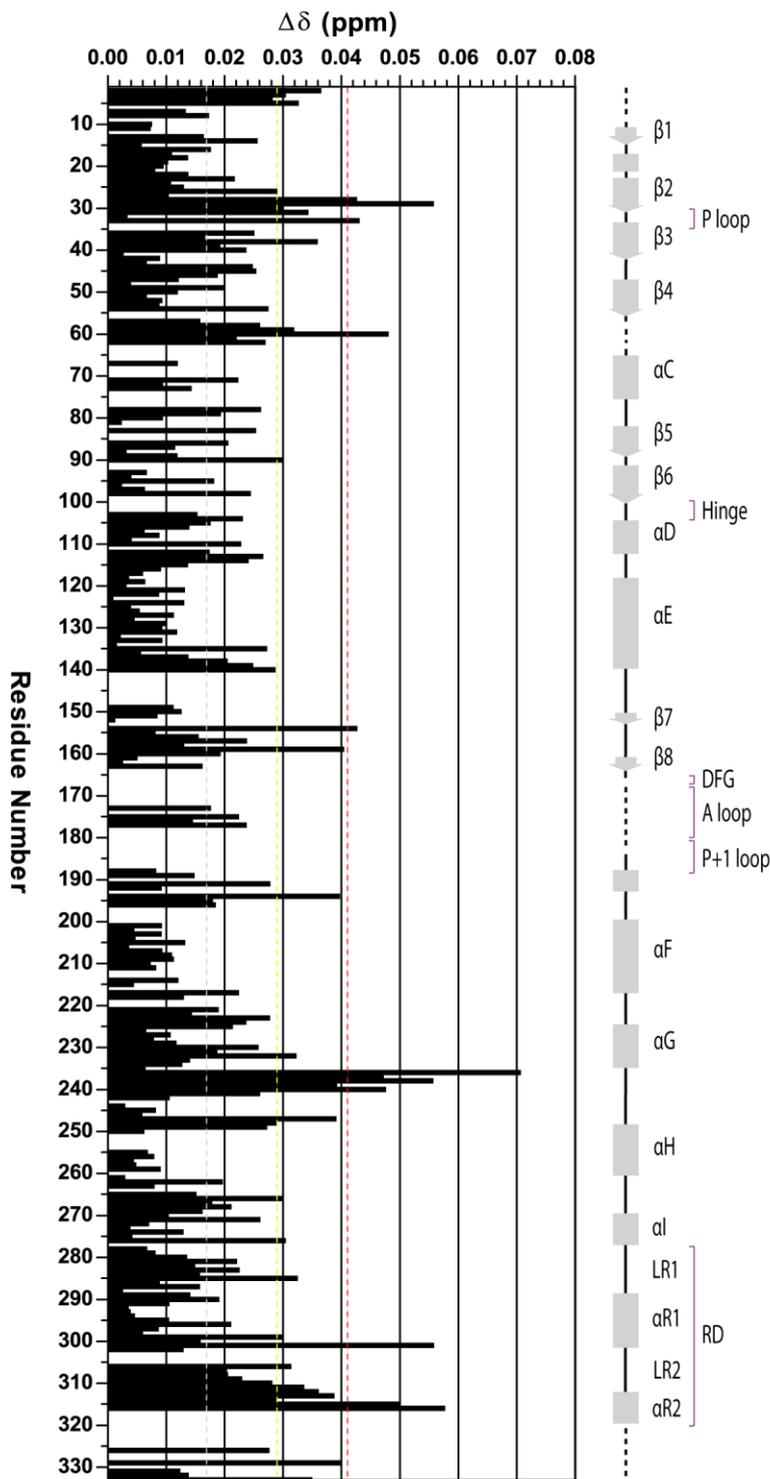
Figure 136. Histogram of CaMK1D CSPs induced by 5 % DMSO. The average peak intensity reducion is shown indicated by the gray line. Those that are ≥1 or 2 S.D are indicated by the yellow and red line respectively.
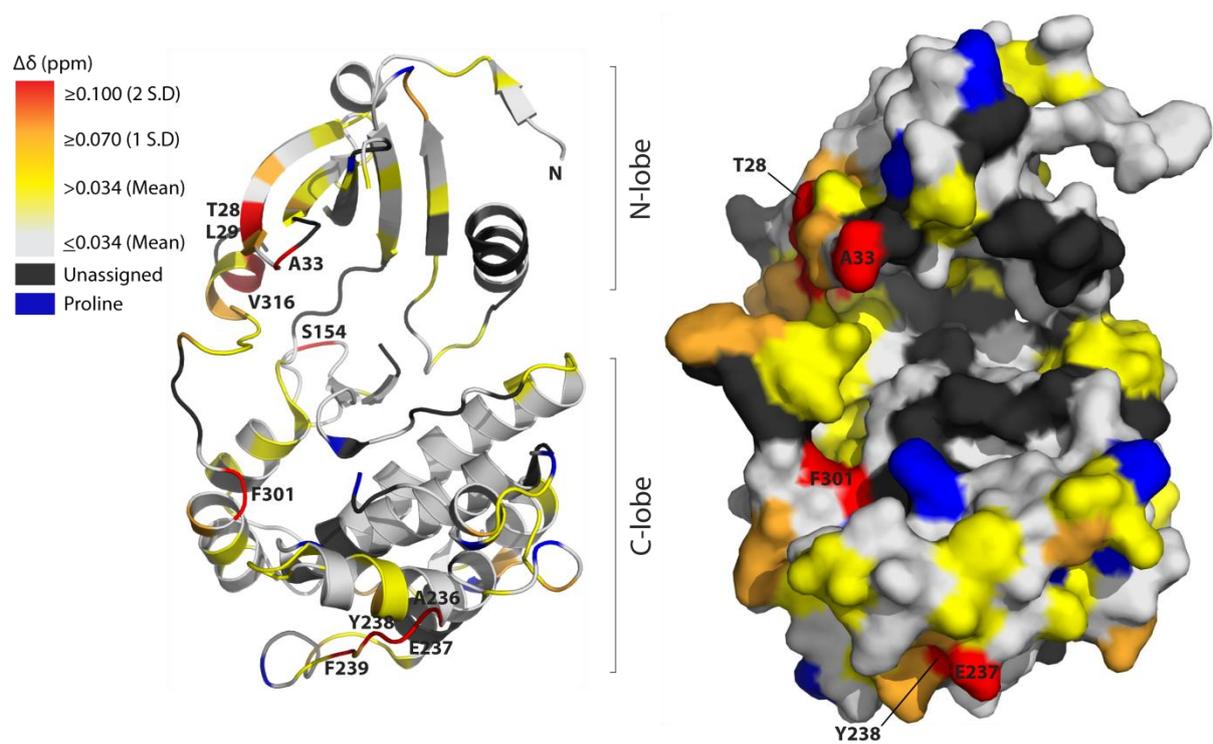
Figure 137 A. CaMK1D CSPs induced by DMSO. A cartoon representation of the CaMK1D crystal structure with CSPs induced by DMSO mapped on. As indicated by the gradient scale, backbone residues are coloured according to the magnitude of the normalised chemical shift changes that occurred upon titrating in DMSO from 0-5 % (v/v).
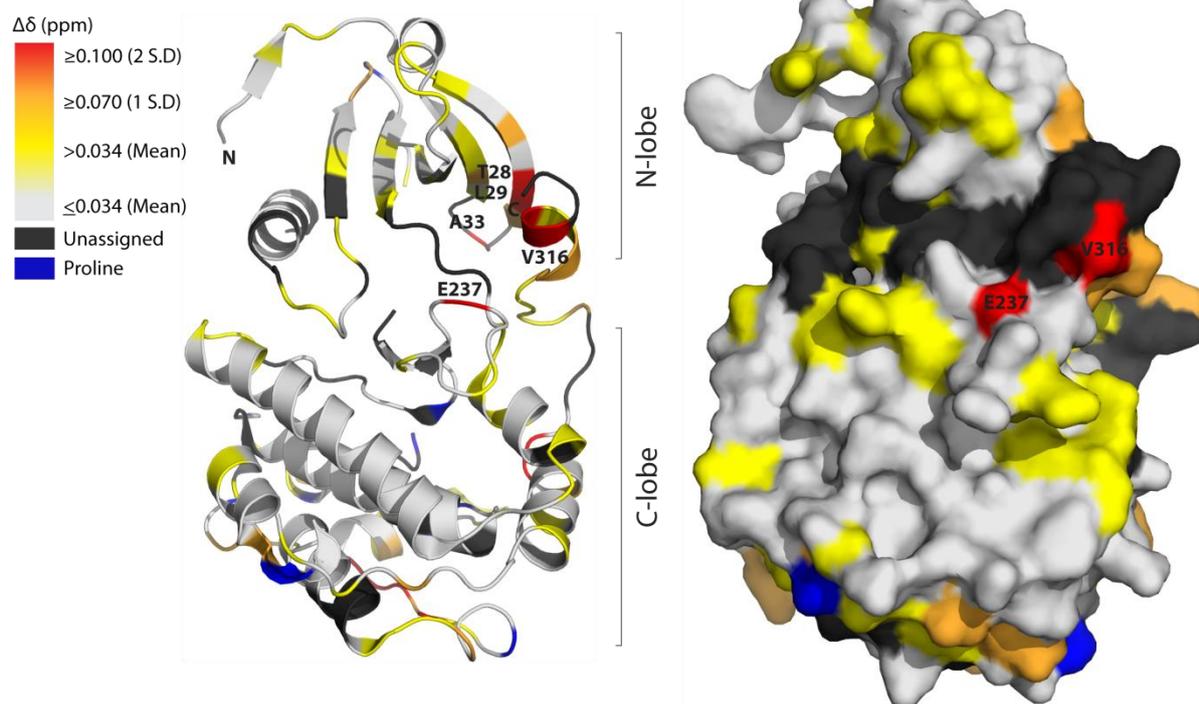
Figure 137 A. CaMK1D CSPs induced by DMSO. A cartoon representation of the CaMK1D crystal structure with CSPs induced by DMSO mapped on. As indicated by the gradient scale, backbone residues are coloured according to the magnitude of the normalised chemical shift changes that occurred upon titrating in DMSO from 0-5 % (v/v).

GSK-3 XIII inhibitor binds into the CaMK1D active site cleft and interacts with the AID/CBD

The binding mechanism of the ATP analog GSK-3 XIII to CaMK1D was elucidated by comparing X-ray crystallography and solution state NMR data, revealing the local and unprecedented long range interactions. Most protein kinases are successfully co-crystallised in the presence of an inhibitor because the apo form is often dynamic and samples multiple conformations (Vogtherr, Saxena et al. 2006)Federov 2007). Likewise the structure of CaMK1D was determined by the SGC as a CaMK1D-GSK-3 XIII inhibitor complex. This ligand was discovered in a HTS using thermofluor[TM] (table 4). The complex reveals that the ATP competitive inhibitor binds to the active site cleft of autoinhibited CaMK1D and makes several contacts (Figure 139). Two notable hydrogen bonds are evident between the GSK-3 XIII inhibitor and Leu 29 and Val 101. A hydrogen bond interaction between the inhibitor and Leu 29 is mediated via a water molecule. Other residues shown in close proximity to GSK-3 XII include Val 37, Ala 50, Lys 52, Val 82, Met 98, Gln 99, Leu 100, Val 101, Gly 104, Leu 151, Ser 164, Trp 306, and Phe 310. Side chains of Trp 306 and Phe 310 appear buried and directed towards GSK-3 XIII suggesting that there is an interaction between them that further stabilizes this region of the AID/CBD. This is in contrast to the apo rat CaMK1A where the equivalent Trp side chain is solvent exposed (Goldberg, Nairn et al. 1996). We therefore believe that the buried orientation of the Trp 306 side chain suggests an allosteric mode of inhibition because stabilization of the AID/CBD and the enforced burial of the Trp 306 side chain will prevent activation by Ca^{2+}/CaM. Our ITC data confirmed the importance of this solvent exposed side chain in mediating the CaMK1D interaction with Ca^{2+}/CaM. In our NMR studies these residues are resolved thus facilitating our analysis of the conformational changes involved which we discuss in the subsequent section.

CaMK1D binds to GSK-3 XIII inhibitor with a high affinity and shows pronounced CSPs around the active site cleft

Although the CaMK1D-GSK-3 XIII co-crystal structure provides insight into the binding mode of GSK-3 XIII and residue contacts within the active site cleft, information about potential long range conformational changes beyond the binding

site are not readily apparent due to the lack of an apo CaMK1D crystal structure for comparison and potential artefacts from enforced lattice contacts in a crystal structure. We therefore examined by NMR the conformational changes of CaMK1D in solution which were induced by GSK-3 XIII binding to gain mechanistic insight into its mode of inhibition and potential allosteric sites that modulate the activity of the CaMK1D.

For this study of possible inhibitor-induced conformational changes we titrated from 0-1.5 equivalents of GSK-3 XIII inhibitor into CaMK1D and monitored CSPs by ^{1}H, ^{15}N TROSY HSQC (figure 140). Spectra collected at each titration point were overlaid, revealing CSPs in slow exchange and therefore indicative of a tight binding interaction (data not shown). Some assigned peaks which exhibit CSPs and intermediate exchange broadening are highlighted in figure 140, which illustrates only the apo CaMK1D and titration endpoint spectra superimposed for simplification. The perturbed amide peaks included those of the activation loop residues Gly 173 and Gly 175, which were found to exhibit slow and intermediate exchange, respectively. Thus the CSPs of Gly 173 and Gly 175 suggests that this region of the activation loop maybe in close proximity to GSK-3 XIII or indirectly experience a conformational change that is conveyed through the active site cleft. Due to the phenomenon of slow exchange the final positions of some peaks in the endpoint spectra had to be predicted. It was not possible to assign the CaMK1D-GSK-3 XIII complex due to time and financial constraints as well as the limited quality of the spectra collected during the titration.

A histogram of the peak intensity reduction induced by GSK-3 XIII (figure 141) indicates a spread of residues from both lobes. Some of these residues occur within or near ATP interacting functional elements such as the P-loop, the catalytic Lys 52, and the hinge. Mapping of the intensity reduction onto the structure (figure 142) reveals that the large reductions occur around the active site cleft but interesting also the RD which is not surprising because this is seen in the crystal structure (figure139).

A few residues in the histogram with large intensity reduction that appear far in sequence from the ATP interacting functional elements include Asn 80, Ile 87, Asn 149, Leu 150, and Leu 214 but mapping of these perturbed residues onto the CaMK1D structure indicates all residues except Leu 214 are in close promixity to the active site cleft (figure142). Leu 214 is buried within the core of the C-terminal lobe but is in close proximity to Pro 147 which is C-terminal of the catalytic loop by one residue. Since the catalytic loop also forms part of the active site cleft think that a long range conformational change induced by GSK-3 XIII is conveyed through to Leu 214 via the catalytic loop and Pro 147.
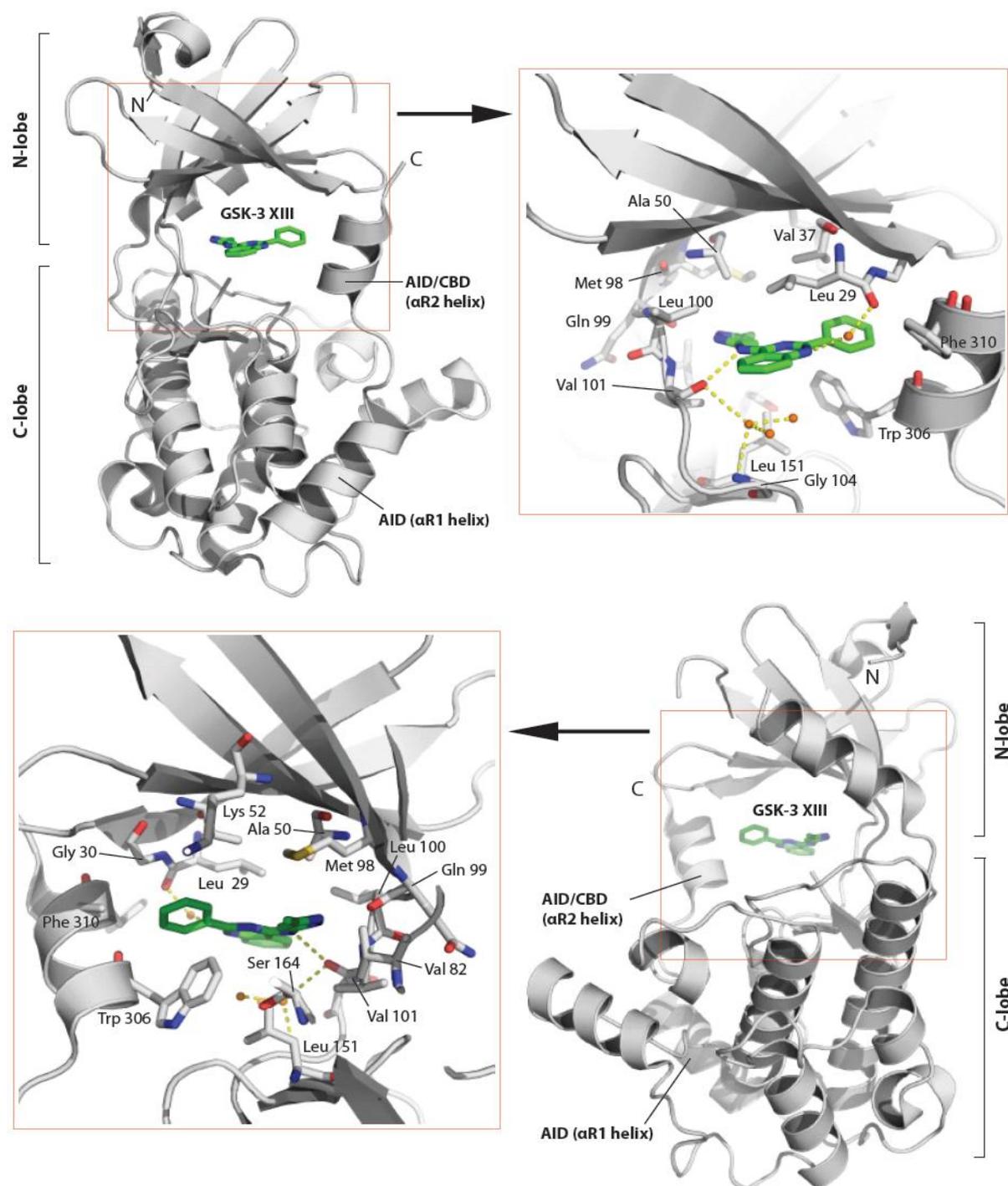
Figure 139. CaMK1D and GSK-3 XIII inhibitor co-crystal interactions. The upper and lower (rotated 180° with respect to upper figure) cartoon figures illustrate the co-crystal structure of CaMK1D with GSK-3 XIII inhibitor bound inside the active site cleft. The expanded active site cleft highlights the CaMK1D residues that are within the vicinity of the GSK-3 XIII inhibitor. Hydrogen bonds (defined as polar contacts between 2.6-3.5 Å in pymol) are shown between CaMK1D and GSK-3 XIII as yellow dashed lines whilst orange spheres denoted water molecules.
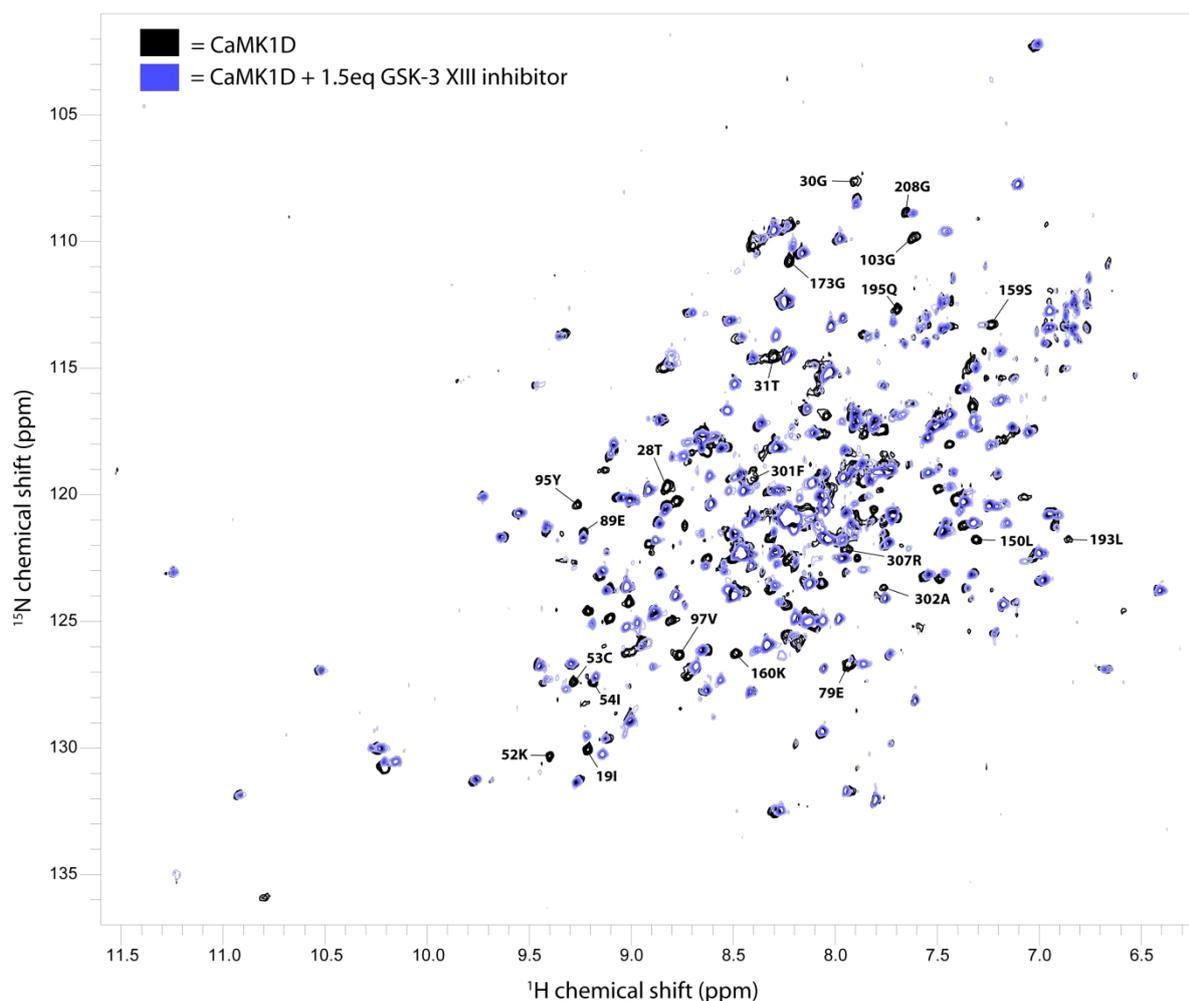
Figure 140. CaMK1D titration with GSK-3 XIII inhibitor in 5 % DMSO. Overlay of the ^{1}H, ^{15}N TROSY HSQC spectra for apo CaMK1D (black) and the titration endpoint of CaMK1D with 1.5 equivalents of GSK-3 XIII inhibitor (blue) reveals peaks that display CSPs and intermediate exchange broadening thus indicating an interaction between CaMK1D with the inhibitor. Titration of CaMK1D with 0.1, 0.25, 0.5, 0.75, and 1.5 equivalents of GSK-3 XIII inhibitor displayed CSPs in slow exchange indicative of a tight binding affinity (spectra not shown). All spectra were acquired at 25 °C on a Varian 800 MHZ spectrometer using 180 µM ^{2}H, ^{15}N CaMK1D, in 50 mM HEPES pH 7.0, 150 mM NaCl, 0.5 mM TCEP, 0.02 % Na Azide, 5 % DMSO. The GSK-3 XIII inhibitor was prepared in an identical buffer.
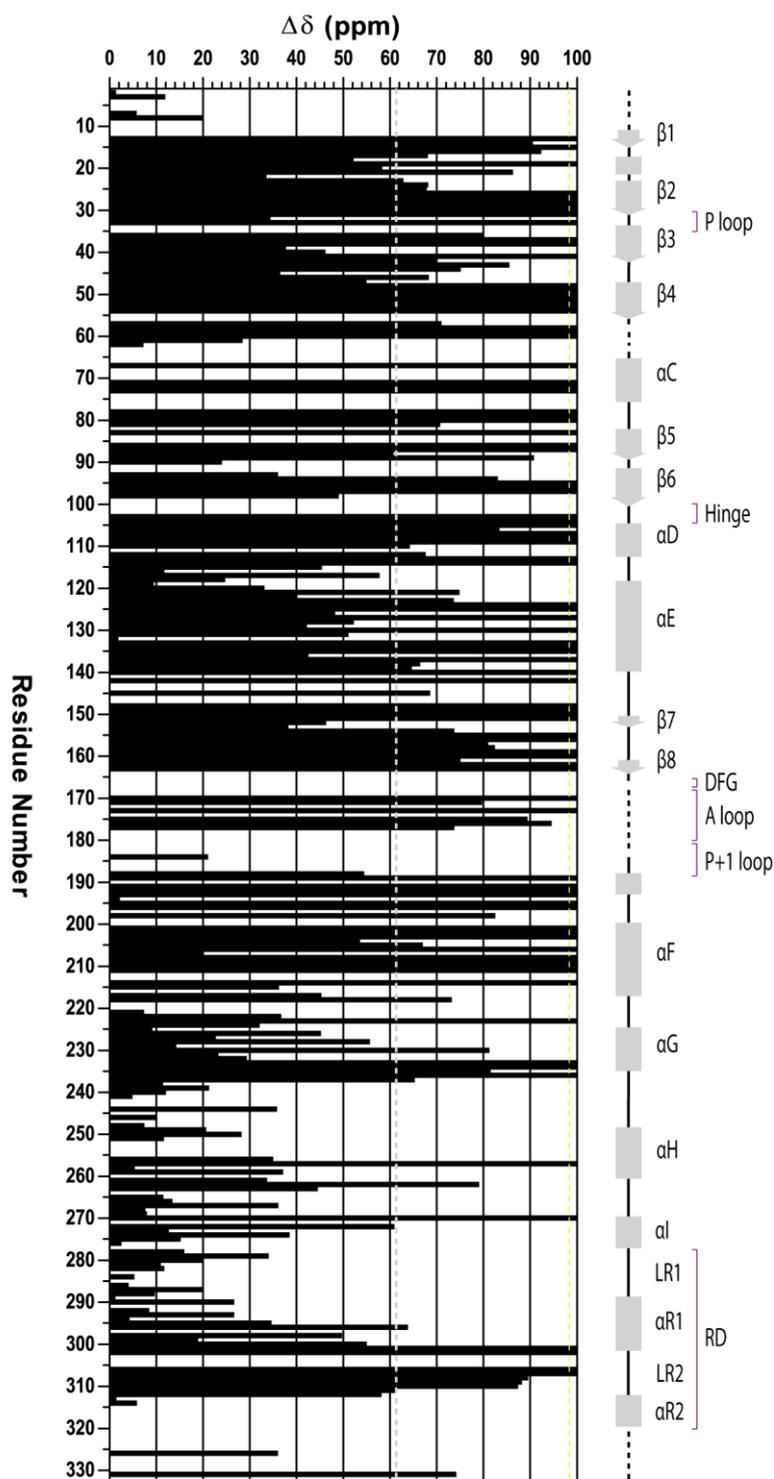
Figure 141. Histogram of CaMK1D amide peak intensity reduction induced by GSK-3 XIII. The average peak intensity reducion is shown indicated by the gray line. Those that are $\geq$1 or 2 S.D are indicated by the yellow and red line respectively.
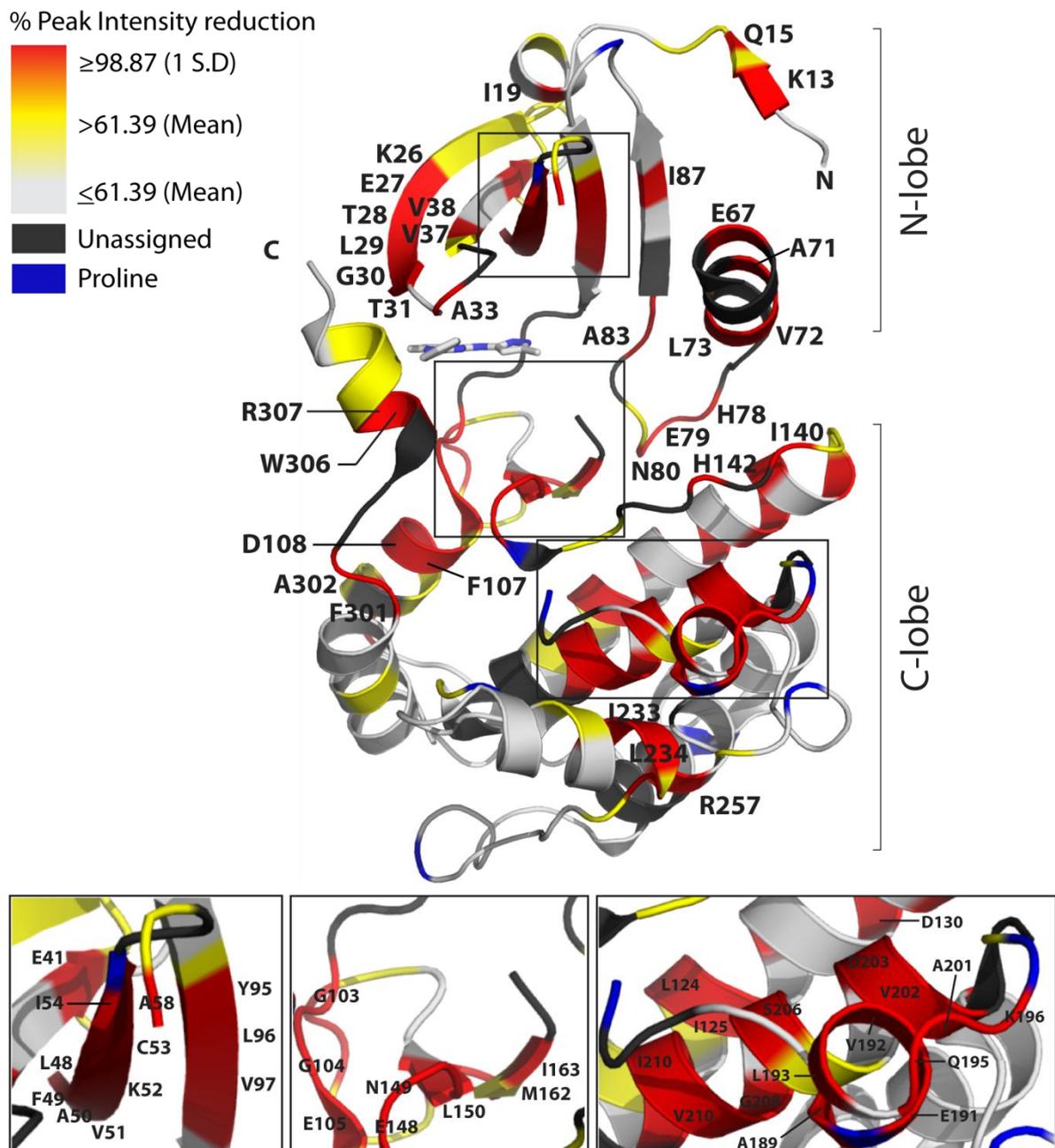
Figure 142 A. CaMK1D CSPs induced by GSK-3 XIII. A cartoon representation of the CaMK1D crystal structure with peak intensity reduction induced by GSK-3 XIII mapped on.
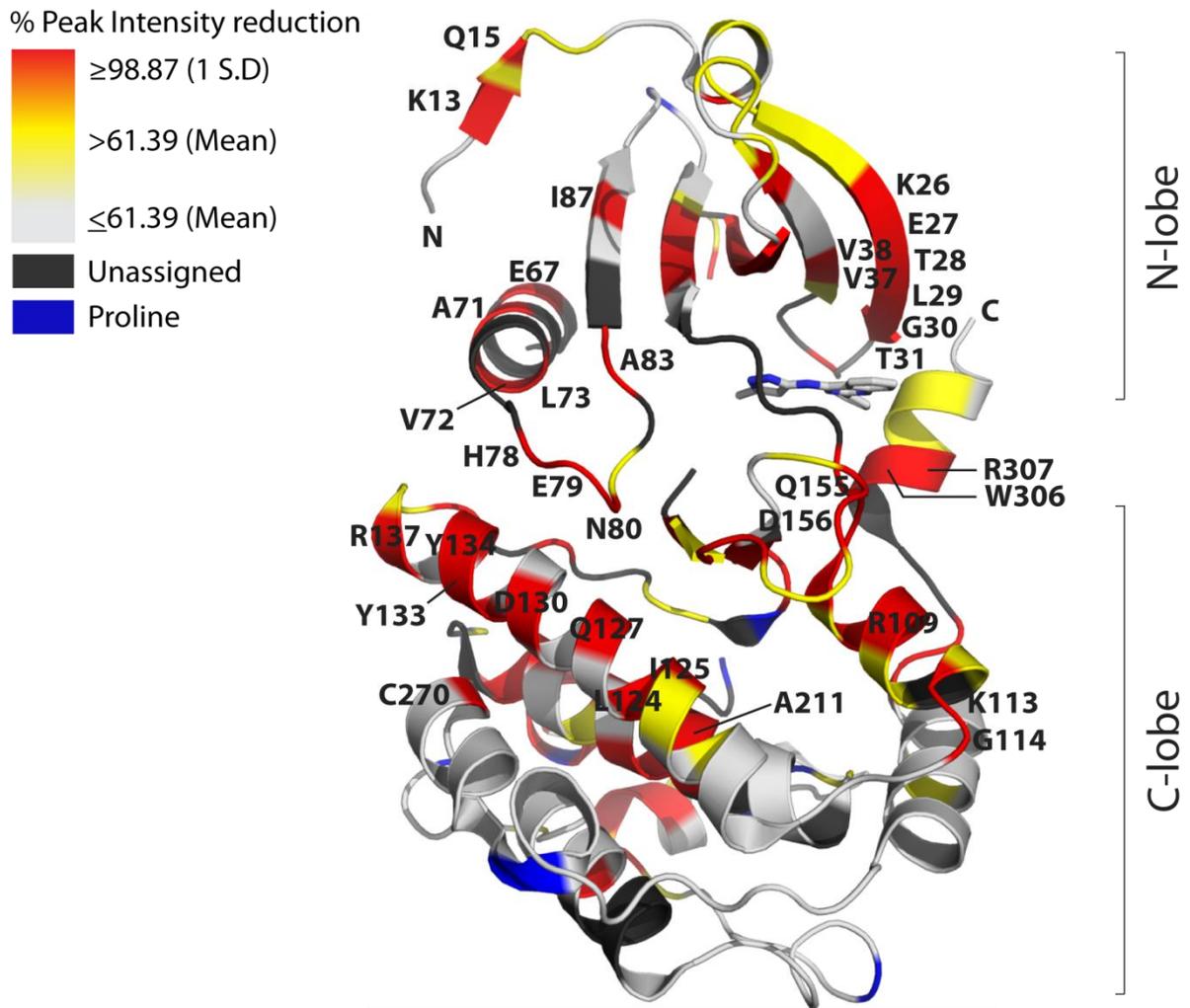
Figure 142 B. CaMK1D CSPs induced by GSK-3 XIII. A 180° rotated view of the previous figure illustrating a cartoon of the CaMK1D crystal structure with peak intensity reduction induced by GSK-3 XIII mapped on.

Chapter IV Discussions and Conclusions

Protein kinases by NMR

Our preliminary analysis of six different kinase proteins by NMR has shown that only two out of the six targets were feasibly tractable to backbone structural studies by NMR without needing extensive optimisation. CaMK1D and PAK5 were the two kinases that produced ^{1}H, ^{15}N HSQC spectra feasible for structural studies. In contrast TOPK, Pim1, CSNK1G3, and CaMK1G were not tractable to backbone structural studies predominantly because the majority of their backbone amide resonances broadened out on an intermediate exchange time scale. We deduce that the main reason for this phenomenon is because of the inherently dynamic nature of kinases which was reflected by their ^{1}H, ^{15}N HSQC spectra and B-factors. This is not surprising because it is well established that kinases are dynamic entities that sample an ensemble of conformations permitted by their conformational plasticity (Vogtherr, Saxena et al. 2006) Kuryian 2002).

Stability, heterogeneity and post translational modifications were additional issues that hindered our study of kinases. Extensive efforts were made to optimise stability by screening different buffer solution conditions because we suspected that an increase in kinase stabilibilty might correlate with improved NMR spectra via changes in the dynamic properties of the kinase. However, our results showed that altering the buffer solution conditions had no significant impact on improving the amide peak count even though an increase in stability was apparent by thermofluor^{TM} in the presence of NaCl as a stabilizer, but these were often small changes typically 2-4 °C. The main notable changes in NMR spectra due to buffer optimisation were increases in peak intensities and pH dependent CSPs. We thus conclude that increasing the Tm of a kinase via changes in buffer conditions does not correlate with an improvement in NMR spectra. Our results emphasise the difficulty in working with kinases particularly by NMR. Hence only a small number of kinases have reportedly been assigned and studied using NMR (Langer, Vogtherr et al. 2004; Vogtherr, Saxena et al. 2005; Gelev, Aktas et al. 2006; Vogtherr, Saxena et al. 2006; Vajpai, Strauss et al. 2008; Masterson, Shi et al. 2009).

The importance of the AID for CaMK1D stability

Both CaMK1D and PAK5 produced ^{1}H, ^{15}N HSQC spectra which were evaluated to be feasible for structural analysis but we decided to pursue CaMK1D as our main target because its NMR spectrum was superior to PAK5 in terms of peak count and intensity. Furthermore during acquisition of ^{1}H, ^{15}N HSQC spectra, CaMK1D displayed fewer tendencies to precipitate at 25 °C compared to PAK5 suggesting that CaMK1D is more stable. There appeared to be a correlation between protein stability and spectral quality. For example the stability of PAK5 and CaMK1D was deduced to be relatively high compared to the other kinases based on thermofluor^{TM} data. CaMK1D in particular exhibited the highest long term stability at 25 °C during acquisition of NMR experiments. In contrast thermofluor^{TM} data suggested that the stability of Pim1, CaMK1G and CSNK1G3 was relatively low with no correlation between Tm and buffer pH. Given that these samples also demonstrated an increase in precipitation 25 °C during acquisition of NMR experiment is consistent with the notion suggested by thermofluor^{TM} that these proteins are relatively unstable. The stability of TOPK was neither relatively low nor high but was considered to have moderate stabiity whilst its ^{1}H, ^{15}N HSQC was superior to Pim1, CaMK1G and CSNK1G3 but inferior to CaMK1D and PAK5.

We examined the B-factors associated with the kinases to infer a correlation with the dynamics suggested by their ^{1}H ^{15}N HSQC spectra. Generally high B-factors were found to correlate with increased dynamics in their spectra often on an intermediate exchange timescale. Another important observation was the difference in stability and dynamics of kinases possessing a regulatory domain and those that were purely catalytic domains. CaMK1D for example contains a relatively rigid AID whereas Pim1 is comprisesd of a catalytic domain only. We believe that the reason CaMK1D produces a high peak count in its spectra with a relatively large subset of consistent and intense peaks is because of its AID which is able to mediate pseudosubstrate contacts and reduce conformational sampling concomitant with a reduction in conformational dynamics.

Conversly Pim1 is constitively active and permitted to sample multiple conformations which we believe correlates to substantial intermediate exchange dynamics. However PAK5 and CaMK1G are execeptions to this because our PAK5 construct for example constitutes the catalytic domain only yet it has an intrinsically stable conformation with stable dynamics. Whereas CaMK1G which we would predict to be similarly well behaved for NMR like CaMK1D is instead contrastingly different exhibiting intermediate exchange dynamics deduced from its NMR specrum. Despite possessing an AID CaMK1G B-factors imply that its conformation is relatively dynamic whilst its stability is low. To summarise these points we believe that based on our results it is effectively very difficult to predict how a kinase and its intrinsic dynamics and stability affects its ^{1}H, ^{15}H HSQC spectrum. Although the presence of a stabilizing AID as exempliefied by CaMK1D suggests that reduced mobility and thus limited sampling of conformations can produce NMR spectra with well dispersed strong amide peak intensites and a high peak count. We have seen that the extent of kinase dynamic behaviour and the NMR time scales on which they occur varies widely with each kinase.

Kinase backbone assignments

Having found a kinase tractable to structural studies by NMR in CaMK1D we have managed to assign 79 % of the backbone amide resonances for apo autoinhibited CaMK1D. The extent of our backbone assignment is comparable to the backbone assignments that have been demonstrated with other kinases. For example PKA in complex with AMPPNP-Mg^{2+} and Abl in complex with Imatinib have been assigned for 80 % and 96 %, of their backbone amides respectively (Vajpai, Strauss et al. 2008; Masterson, Shi et al. 2009). In comparison to other apo kinases, our backbone assignment coverage for apo CaMK1D is also comparable to PKR which had in total 75 % of its backbone assigned and an improvement on the backbone assignments reported for PKA and p38MAPK (Langer, Vogtherr et al. 2004; Vogtherr, Saxena et al. 2005; Gelev, Aktas et al. 2006). It is interesting to note from these few examples that a more complete backbone assignment can be accomplished for kinases in complex with ligand rather than in apo form due to the fact that kinases are dynamic proteins. Therefore a ligand can stabilize and trap a kinase in a particular conformation whereas an apo kinase can sample an ensemble of conformations.

This is also analogous to kinases that are tackled by protein crystallography which often require a ligand in order the kinase to become receptive to conditions that promote crystallisation.

The key regions which were unassigned in CaMK1D correlate predominantly to loops that include the catalytic and activation loops. In the CaMK1D crystal structure the catalytic loop is well defined but the activation loop is disordered. This is often the case if the activation loop is in an unstable conformation usually in the absence of activation loop phosphorylation. Curiously the crystal structure of CaMK1G displays an ordered activation which is in contrast to CaMK1D even though they share high sequence homology. Other unassigned regions correlate to functionally mobile elements such as the αC helix, the hinge, and a portion of the CBD. Apart from the CBD these elements are well defined in the CaMK1D crystal structure. Analysis of the CaMK1D B-factors reveals that the αC helix is relatively dynamic and in our spectrum we presume that this correlates to dynamics on an intermediate exchange timescale.

Assignments could not be acquired for the unassigned regions predominantly because the amide resonances were very weak in intensity or completely broadened out. Some sequential assignments could not be established due to CB resonance broadening. Therefore in order to increase our backbone assignment coverage we attempted to optimise the CaMK1D spectrum by lowering the pH but this had negligible effect in terms of producing new peaks and increasing the intensity of weak resonances. We conclude that these regions must be inherently dynamic on an intermediate exchange timescale. Lowering the pH did however induce pH dependent CSPs and when these were mapped to the CaMK1D structure there was a localise patch of pH sensitive residues near the N-terminal start of the AID. We therefore speculate whether there is a potential pH dependent effect on the disengagement of the AID. For future studies aiming to optimise the peak count we propose the use of a ligand such as AMPPNP to stabilize the active site cleft elements such as the catalytic and activation loops, hinge, and the αC helix. This has been demonstrated by Masteron et al with 80 % of the PKA backbone amides

assigned. A significant improvement on the previous backbone assignment of PKA which was only had 55 % of its total backbone amide resonances assigned (Langer, Vogtherr et al. 2004; Masterson, Shi et al. 2009).

Refinement of CaMK1D crystal structure using NMR

A TALOS based structure prediction was performed using our backbone assignments and compared with the secondary structure defined by the CaMk1D crystal structure. Some differences were revealed in this comparison, notably at the N-terminus β1 strand is absent in our prediction because this secondary structure element is a crystallisation artefact that acts as a dimerisation interface possibly via strand exchange. We had no evidence with size exclusion chromatography, NMR, or SAXS to suggest that the CaMK1D crystal structure dimer occurred in solution. Another noteworthy difference was evident in the CBD which is defined as a helix (αR2 helix) in the crystal structure however in our NMR data this was predicted as random coil. We postulate that this region becomes structured when Ca^{2+}/CaM binds to it. Secondary structural differences were also evident in the length of helices within the C-terminal substrate binding lobe. TALOS predicted these helices to be shorter. These highlighted structural differences were subsequently decided combined to create a hybrid CaMK1D model by using our dihedral angles restraints derived from TALOS to refine the crystal structure. Furthermore we included dihedral angle restraints from assignments that spanned disordered regions in the crystal such as the activation loop, N-termini, and C-termini. As a result of this our final model illustrates a refined crystal structure with missing loops including the activation loop modelled in. The activation model also reflects the dynamics and disordered nature of the loops.

CaMK1D-Ca^{2+}/CaM complex model and insight into mechanism of activation
CaMK proteins are dependent on Ca^{2+}/CaM for activation and it is known that the binding interaction occurs around the CBD. The binding mode of Ca^{2+}/CaM and the CBD has been characterised through studies using CBD peptides which revealed that Ca^{2+}/CaM essentially wraps around the CBD and involves Trp and Met being

buried into two hydrophobic pockets within Ca^{2+}/CaM (Clapperton, Martin et al. 2002). Structural characterisation of full CaMK protein with Ca^{2+}/CaM have not been desmonstrated only until recently with the DAPK- Ca^{2+}/CaM complex (de Diego, Kuper et al. 2010). We were also interested in characterising the complex and attempted co-crystallisation screening of CaMK1D- Ca^{2+}/CaM complex but had no success.

Another technique we used to complement our complex data was SAXS. Our final CaMK1D-Ca^{2+}/CaM complex model illustrates an elongated conformation with Ca^{2+}/CaM positioned at the base of the CaMK1D C-terminal substrate binding lobe. This model is in contrast to the crystal structure of the DAPK- Ca^{2+}/CaM complex (de Diego, Kuper et al. 2010). Based on our model we suggest an activation mechanism that can be described in a three step process. Firstly CaMK1D binds to Ca^{2+}/CaM and forming a tight interaction around the CBD. In this position Ca^{2+}/CaM is docked against CaMK1D as illustrated by the DAPK-Ca^{2+}/CaM complex (de Diego, Kuper et al. 2010).

The second step involves disengagement of the AID and CBD from the substrate binding site and N-terminal lobe. We postulate that this is coupled to a swing out mechanism with a pivot point acting near the beginning of AID which is defined by about seven residues in a loop enabling flexibility of the AID once disengaged. A third and final step to form an active complex with the substrate binding site exposed would involve docking of the Ca^{2+}/CaM-AID-CBD to the base of C-terminal lobe. This final step according to SAXS data represents the active complex as opposed to the structure of DAPK-Ca^{2+} which still has the substrate binding site occluded. Finally we propose that the three step cycle repeats and is dynamic with the Ca^{2+}/CaM-AID-CBD engaging and disengaging from the the substrate binding site until Ca^{2+}/CaM dissociates. Based on this model and current knowledge about other CaMKI regulatory mechanism we have devised a schematic that summaries the conformational and functional states of CaMK1D (figure 143).
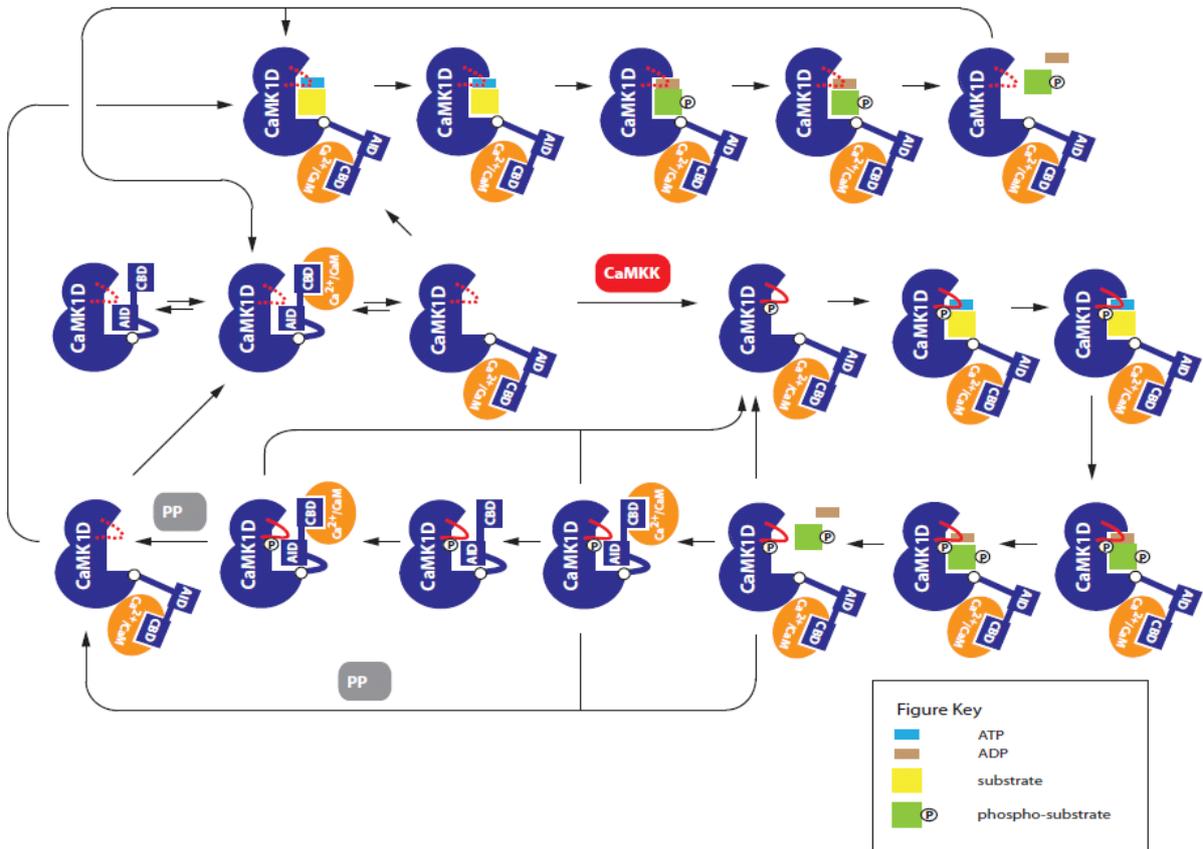
Figure 143. A schematic of CaMK1D regulation. A summary of the different conformational and functional states that CaMK1D undergoes. CaMK1D can exists in main states described as autoinhibited CaMK1D, autoinhibited CaMK1D-Ca2+/CaM, active CaMK1D-Ca2+/CaM, autoinhibited CaMK1D_T180*, autoinhibited CaMK1D_T180*-Ca2+/CaM, maximally active CaMK1D_T180*-Ca2+/CaM,

To help us further elucidate the mechanism of CaMK1D activation we used CSPs and peak intensity changes from our CaMK1D interactions with Ca^{2+}/CaM by NMR to infer conformational changes but the evidence is not clear. Firstly we observe a tight interaction due to evidence of slow exchange coupled with an increase in peaks near the centre of the spectrum. Initially we proposed an unfolding mechanism because of this observation. Furthermore when the CaMK1D titrations with Ca^{2+}/CaM were performed by NMR we noticed in the first increment in a ^{1}H, ^{15}N HSQC TROSY suggest unfolded due to an apparent reduction in resonance dispersions in the ^{1}H dimension in the amide region. However an unfolding mechanism is not consistent with Ca^{2+}/CaM-CBD peptide studies and the DAPK-Ca^{2+}/CaM complex (Clapperton, Martin et al. 2002; de Diego, Kuper et al. 2010). Therefore we suspect that there is an increase in helical content upon and perhaps random coil structure from the activation loop which gives rise to the increased peak overlap in the centre of the CaMK1D-Ca^{2+}/CaM ^{1}H, ^{15}N HSQC spectrum.

Mapping of the CSPs based purely on whether resonances broadened out or exhibited slow exchange were found to localise predominantly on the AID, CBD, a region near the p-loop, our putative Ca^{2+}/CaM docking site base of the C-terminal lobe and a couple of the residues that comprise the first loop of the AID (figure 116). CSPs corresponding to this loop however are not dramatic which we would expect for a swing out mechanism. It is possible that the pivot point is not within the loop but at the N-terminal end of the αR1 helix of the AID. In order to validate this mechanism we would ideally reassign the CaMK1D-Ca^{2+}/CaM because many peaks undergo CSPs on a slow exchange timescale but this could be difficult to accomplish due to the increase in resonance overlap in the centre the spectrum.

In addition to this mechanism of activation regarding the accessibility of the substrate binding site we examined the functional elements involved in catalysis i.e. the catalytic Lys 52, p-loop, αC helix, catalytic loop, activation loop, and hinge. We had assignments that spanned some of these elements. Upon Ca^{2+}/CaM binding the CaMK1D spectrum did not appear to exhibit significant CSPs associated with resonances that correlated to these elements. We conclude that Ca^{2+}/CaM did not

induce any local conformational changes to the active site cleft functional elements. This would suggest that the main role of Ca^{2+}/CaM is to open up the substrate binding site. It is suspected that rearrangement of the active site cleft is induced by activation loop phosphorylation analogous to PKA.

Ca^{2+}/CaM and CBD binding mode

The CaMK1D crystal structure exhibited an interesting structural difference when compared to the rat CaMK1A crystal structure. This key difference was highlight in the CBD and concerned the orientation of Trp 306 in CaMK1D which appeared to adopt a buried orientation as opposed to the solvent exposed Trp 303 in rat CaMK1A (Goldberg, Nairn et al. 1996). Both Trp residues are equivalent and throughout the CaMKI family it is conserved. The implications of this residue have been shown to be critical for Ca^{2+}/CaM binding and its binding mode has been studied in peptide models (Matsushita and Nairn 1998; Clapperton, Martin et al. 2002). Consequently our mutatgenesis studies on this residue show that the residue is not buried because on the basis of our backbone solution structure of CaMK1D_W306S the protein exhibits a fold similar to CaMK1D suggesting that the side chain is not buried. ITC experiments also confirmed the importantance of Trp 306 in mediating a high affinity interaction. In CaMK1D the affinity with Ca^{2+}/CaM was 28.5 nM whereas in CaMK1D_W306S the affinity was reduced significantly something $\geq$ 100 µM. We could not accurately determine the affinity for CaMK1D_W306S because this affinity was at the detection limits of the ITC instrument. A second contact point for Ca^{2+}/CaM is also postulated in CaMK1D Met 319 but our proposed experiment was not possible due to time constraints.

Ligand binding

Our ligand binding studies by NMR and thermofluor have shown that autoinhibited CaMK1D has an accessible active site cleft. Binding to the inhibitor GSK-3 XIII suggested two modes of inhibition firstly via competitive inhibition in the active site cleft and secondly via allosteric modulation of the CBD and Trp 306 side chain into a non productive binding mode with Ca^{2+}/CaM. Furthemore our AMPPNP-Mg^{2+} binding studies have illustrated that autoinhibited CaMK1D has a very weak binding affinity.

This suggests that another activation mechanism is necessary for increased nucleotide affinity in the active complex.

Appendix

Recipes

<u>M9 Minimal media recipe per L</u>

<u>Nutrient mix 14ml, per L</u>

2ml 1M $MgSO_4$

0.5ml metal mix

100ul 1M $CaCl_2$

400ul 3mM $FeCl_2$

10ml 20% (w/v) Glucose (2g), (for ^{13}C labelling use ^{13}C-glucose, for perduteration use deuterated glucose)

1ml 20mg/ml Thiamine

1g $^{15}NH_4Cl$

Sterilise through 0.22uM filter.

<u>M9 10X buffer salts 1L stock</u>

60g Na_2HPO_4

30g KH_2PO_4

5g NaCl

pH 7.2

Sterilise through a 0.22 uM filter, and autoclave if not made up in ^2H_2O, use 100ml per L of minimal media

<u>Metal mix 100ml stock</u>

Dissolve the metals in the following order:

0.115g ZnSO4 7H2O (4mM)

0.0169g of MnSO4 5H2O (1mM)

0.029g of H3BO3 (4.7mM)

0.0175g of CuSO4 (0.7mM)

Sterilise through 0.22uM filter

M9 Minimal Media

100 ml 10X M9 buffer salts pH7.2

14 ml metal mix

Make up the M9 minimal media to a final volume of 1L with mQH_2O or 2H2O for expression of deutrated proteins if inoculated with resuspended cell pellet. Alternatively adjust the volume of mQH_2O or ^2H_2O to account for the volume of pre-culture inoculate. For perdueterated protein nutrient mix and M9 buffer salts in 100 % ^2H_2O

CaMK1D, 1, FL
42.9 kDa, 385 AA

1 23 T180 280-321 279 303-321 385

CaMK1D SGC
37.8 kDa, 333 AA
37.8 kDa, 334 AA

1 23 T180 280-321 279 303-321 333

CaMK1G, 1, FL
53.1 kDa, 476 AA

1 23 T178 279-318 278 302-320 473-476

CaMK1G SGC
34.5 kDa, 302 AA
34.8 kDa, 304 AA

1 23 T178 279-318 278 302-317

CSNK1G3, 1, FL
51.4 kDa, 447AA

1 43 313 447

CSNK1G3 SGC
38.3 kDa, 328 AA
38.5 kDa, 300 AA

35 43 313 328

PAK4, 1, FL
64.1 kDa, 591 AA

1 4-8 10-30 104-108 228-232 277-281, 280-284, 285-289, 288-292 292-323 324 S472 572 591

PAK4 SGC
33.9 kDa, 300 AA
34.3 kDa, 305 AA

292-323 591 324 S472 572 591

PAK5, FL
80.8 kDa, 719 AA

1 4-8 10-30 31 83 104-108 120-133 350-354 401-411 452 S602 719

PAK5 SGC
33.7 kDa, 296 AA
33.7 kDa, 297 AA

425 S602 719

Pim1, 1, FL
45.4 kDa, 404 AA

1 129 381 404

Pim1, 2, FL
35.7 kDa, 313 AA

92 129 381 404

Pim1 SGC
35.6 kDa, 313 AA
35.7 kDa, 314 AA

92 129 381 404

TOPK, FL
36.1 kDa, 322 AA

1 T9 32 T198 234 275 319-322 318

TOPK ΔE2
36.0 kDa, 321 AA
36.2 kDa, 323 AA

1 T9 32 T198 234 275 319-322 318

TOPK 10-322
35.1 kDa, 313 AA
35.4 kDa, 316 AA

10 32 T198 234 275 319-322 318

VRK2, 1, FL
58.1 kDa, 508 AA

1 29 319 492 508

VRK2 SGC,
36.5 kDa, 322 AA
36.7 kDa, 324 AA

14 29 319 335

Legend

- Kinase domain (KD)
- CaMK Autoinhibitory domain (AID)
- Ca^{2+}/CaM binding domain (CBD)
- CAAX motif
- Random coil/unstructured
- Nuclear localization signal (NLS)
- p21 binding domain (PDB) aka Cdc42/Rac1 interactive binding domain (CRIB)
- Proline rich domain
- Guanine nucleotide exchange factor (GEF) interacting domain
- Region containing mitochondrial localization signal (MLS)
- Transmembrane domain
- PAK5 Autoinhibitory domain (PAK5 AID)
- Nuclear export sequence (NES)
- Protein binding region (PBR)
- PDZ binding binding motif

337

Chemical shift list

Backbone assignment Chemical shift list

```
#Atom     Residue
#shift    Seq    Residue Atom   Atom   Shift/   Error/ Ambiguity
#assign   code   Label   Name   Type   ppm      ppm    Code
#-------------------------------------------------------------
#

1         2      Met     C      C      175.779  0.000  1
2         2      Met     CA     C      55.139   0.000  1
3         2      Met     CB     C      32.266   0.000  1
4         3      Ala     C      C      177.349  0.000  1
5         3      Ala     CA     C      52.042   0.000  1
6         3      Ala     CB     C      18.480   0.000  1
7         3      Ala     H      H      8.354    0.000  1
8         3      Ala     N      N      126.137  0.000  1
9         4      Arg     C      C      176.064  0.000  1
10        4      Arg     CA     C      55.406   0.000  1
11        4      Arg     CB     C      30.176   0.000  1
12        4      Arg     H      H      8.232    0.000  1
13        4      Arg     N      N      120.940  0.000  1
14        5      Glu     C      C      175.329  0.000  1
15        5      Glu     CA     C      55.974   0.000  1
16        5      Glu     CB     C      29.820   0.000  1
17        5      Glu     H      H      8.407    0.000  1
18        5      Glu     N      N      122.895  0.000  1
19        6      Asn     C      C      179.348  0.000  1
20        6      Asn     CA     C      54.457   0.000  1
21        6      Asn     CB     C      40.284   0.000  1
22        6      Asn     H      H      8.003    0.000  1
23        6      Asn     N      N      124.987  0.000  1
24        7      Gly     C      C      174.627  0.000  1
25        8      Glu     C      C      176.969  0.000  1
26        8      Glu     CA     C      56.194   0.000  1
27        8      Glu     CB     C      29.489   0.000  1
28        8      Glu     H      H      8.435    0.000  1
29        8      Glu     N      N      121.165  0.000  1
30        9      Ser     C      C      175.059  0.000  1
31        9      Ser     CA     C      58.001   0.000  1
32        9      Ser     CB     C      63.411   0.000  1
33        9      Ser     H      H      8.387    0.000  1
34        9      Ser     N      N      117.251  0.000  1
35        10     Ser     CA     C      58.413   0.000  1
36        11     Ser     C      C      174.821  0.000  1
37        11     Ser     CA     C      58.232   0.000  1
38        11     Ser     CB     C      63.034   0.000  1
39        11     Ser     H      H      8.199    0.000  1
40        11     Ser     N      N      117.563  0.000  1
41        12     Ser     C      C      174.428  0.000  1
42        12     Ser     CA     C      58.488   0.000  1
43        12     Ser     CB     C      63.060   0.000  1
44        12     Ser     H      H      8.120    0.000  1
45        12     Ser     N      N      117.491  0.000  1
46        13     Trp     C      C      175.387  0.000  1
47        13     Trp     CA     C      56.037   0.000  1
48        13     Trp     CB     C      28.643   0.000  1
```

49	13	Trp	H	H	7.476	0.000	1
50	13	Trp	N	N	120.607	0.000	1
51	14	Lys	C	C	175.457	0.000	1
52	14	Lys	CA	C	54.952	0.000	1
53	14	Lys	CB	C	32.153	0.000	1
54	14	Lys	H	H	7.349	0.000	1
55	14	Lys	N	N	121.091	0.000	1
56	15	Lys	C	C	175.958	0.000	1
57	15	Lys	CA	C	55.932	0.000	1
58	15	Lys	CB	C	33.074	0.000	1
59	15	Lys	H	H	8.122	0.000	1
60	15	Lys	N	N	125.057	0.000	1
61	16	Gln	C	C	176.107	0.000	1
62	16	Gln	CA	C	54.709	0.000	1
63	16	Gln	CB	C	29.030	0.000	1
64	16	Gln	H	H	8.516	0.000	1
65	16	Gln	N	N	121.575	0.000	1
66	17	Ala	C	C	175.622	0.000	1
67	17	Ala	CA	C	50.663	0.000	1
68	17	Ala	CB	C	20.012	0.000	1
69	17	Ala	H	H	8.615	0.000	1
70	17	Ala	N	N	126.136	0.000	1
71	18	Glu	C	C	176.558	0.000	1
72	18	Glu	CA	C	58.007	0.000	1
73	18	Glu	CB	C	29.134	0.000	1
74	18	Glu	H	H	8.493	0.000	1
75	18	Glu	N	N	122.678	0.000	1
76	19	Asp	C	C	177.837	0.000	1
77	19	Asp	CA	C	52.745	0.000	1
78	19	Asp	CB	C	41.537	0.000	1
79	19	Asp	H	H	7.750	0.000	1
80	19	Asp	N	N	115.905	0.000	1
81	20	Ile	C	C	176.191	0.000	1
82	20	Ile	CA	C	59.835	0.000	1
83	20	Ile	CB	C	38.303	0.000	1
84	20	Ile	H	H	9.256	0.000	1
85	20	Ile	N	N	129.928	0.000	1
86	21	Lys	C	C	177.579	0.000	1
87	21	Lys	CA	C	56.827	0.000	1
88	21	Lys	CB	C	29.817	0.000	1
89	21	Lys	H	H	8.016	0.000	1
90	21	Lys	N	N	121.551	0.000	1
91	22	Lys	C	C	176.828	0.000	1
92	22	Lys	CA	C	56.715	0.000	1
93	22	Lys	CB	C	31.716	0.000	1
94	22	Lys	H	H	7.404	0.000	1
95	22	Lys	N	N	115.869	0.000	1
96	23	Ile	C	C	176.227	0.000	1
97	23	Ile	CA	C	60.787	0.000	1
98	23	Ile	CB	C	38.127	0.000	1
99	23	Ile	H	H	7.294	0.000	1
100	23	Ile	N	N	117.589	0.000	1
101	24	Phe	C	C	173.681	0.000	1
102	24	Phe	CA	C	56.059	0.000	1
103	24	Phe	CB	C	43.051	0.000	1
104	24	Phe	H	H	7.983	0.000	1
105	24	Phe	N	N	118.392	0.000	1
106	25	Glu	C	C	176.320	0.000	1

107	25	Glu	CA	C	53.996	0.000	1
108	25	Glu	CB	C	30.774	0.000	1
109	25	Glu	H	H	9.067	0.000	1
110	25	Glu	N	N	120.078	0.000	1
111	26	Phe	C	C	176.041	0.000	1
112	26	Phe	CA	C	55.787	0.000	1
113	26	Phe	CB	C	36.617	0.000	1
114	26	Phe	H	H	9.015	0.000	1
115	26	Phe	N	N	128.879	0.000	1
116	27	Lys	C	C	175.781	0.000	1
117	27	Lys	CA	C	55.752	0.000	1
118	27	Lys	CB	C	31.458	0.000	1
119	27	Lys	H	H	9.116	0.000	1
120	27	Lys	N	N	124.959	0.000	1
121	28	Glu	C	C	175.030	0.000	1
122	28	Glu	CA	C	55.722	0.000	1
123	28	Glu	CB	C	30.953	0.000	1
124	28	Glu	H	H	7.767	0.000	1
125	28	Glu	N	N	117.485	0.000	1
126	29	Thr	C	C	174.244	0.000	1
127	29	Thr	CA	C	63.646	0.000	1
128	29	Thr	CB	C	68.674	0.000	1
129	29	Thr	H	H	8.788	0.000	1
130	29	Thr	N	N	119.300	0.000	1
131	30	Leu	C	C	177.504	0.000	1
132	30	Leu	CA	C	54.690	0.000	1
133	30	Leu	CB	C	41.800	0.000	1
134	30	Leu	H	H	9.215	0.000	1
135	30	Leu	N	N	128.565	0.000	1
136	31	Gly	C	C	171.962	0.000	1
137	31	Gly	CA	C	45.106	0.000	1
138	31	Gly	H	H	7.922	0.000	1
139	31	Gly	N	N	107.977	0.000	1
140	32	Thr	C	C	173.973	0.000	1
141	32	Thr	CA	C	60.944	0.000	1
142	32	Thr	CB	C	70.526	0.000	1
143	32	Thr	H	H	8.280	0.000	1
144	32	Thr	N	N	114.339	0.000	1
145	33	Gly	C	C	173.519	0.000	1
146	33	Gly	CA	C	43.813	0.000	1
147	33	Gly	H	H	8.249	0.000	1
148	33	Gly	N	N	112.484	0.000	1
149	34	Ala	CA	C	53.864	0.000	1
150	34	Ala	CB	C	17.976	0.000	1
151	34	Ala	H	H	8.453	0.000	1
152	34	Ala	N	N	123.970	0.000	1
153	36	Ser	C	C	173.217	0.000	1
154	36	Ser	CA	C	57.262	0.000	1
155	36	Ser	CB	C	65.398	0.000	1
156	37	Glu	C	C	174.418	0.000	1
157	37	Glu	CA	C	55.090	0.000	1
158	37	Glu	CB	C	31.827	0.000	1
159	37	Glu	H	H	8.030	0.000	1
160	37	Glu	N	N	119.898	0.000	1
161	38	Val	C	C	175.249	0.000	1
162	38	Val	CA	C	60.568	0.000	1
163	38	Val	CB	C	32.185	0.000	1
164	38	Val	H	H	8.768	0.000	1

165	38	Val	N	N	124.815	0.000	1
166	39	Val	C	C	174.208	0.000	1
167	39	Val	CA	C	58.829	0.000	1
168	39	Val	CB	C	34.185	0.000	1
169	39	Val	H	H	8.990	0.000	1
170	39	Val	N	N	124.114	0.000	1
171	40	Leu	C	C	174.749	0.000	1
172	40	Leu	CA	C	54.737	0.000	1
173	40	Leu	CB	C	41.905	0.000	1
174	40	Leu	H	H	8.648	0.000	1
175	40	Leu	N	N	126.710	0.000	1
176	41	Ala	C	C	174.282	0.000	1
177	41	Ala	CA	C	49.729	0.000	1
178	41	Ala	CB	C	23.833	0.000	1
179	41	Ala	H	H	9.200	0.000	1
180	41	Ala	N	N	129.836	0.000	1
181	42	Glu	C	C	176.328	0.000	1
182	42	Glu	CA	C	53.233	0.000	1
183	42	Glu	CB	C	33.137	0.000	1
184	42	Glu	H	H	8.910	0.000	1
185	42	Glu	N	N	121.906	0.000	1
186	43	Glu	C	C	176.584	0.000	1
187	43	Glu	CA	C	55.924	0.000	1
188	43	Glu	CB	C	29.345	0.000	1
189	43	Glu	H	H	9.766	0.000	1
190	43	Glu	N	N	131.272	0.000	1
191	44	Lys	C	C	178.955	0.000	1
192	44	Lys	CA	C	59.929	0.000	1
193	44	Lys	CB	C	30.658	0.000	1
194	44	Lys	H	H	8.296	0.000	1
195	44	Lys	N	N	132.543	0.000	1
196	45	Ala	C	C	178.436	0.000	1
197	45	Ala	CA	C	53.928	0.000	1
198	45	Ala	CB	C	18.981	0.000	1
199	45	Ala	H	H	8.842	0.000	1
200	45	Ala	N	N	117.053	0.000	1
201	46	Thr	C	C	176.212	0.000	1
202	46	Thr	CA	C	60.893	0.000	1
203	46	Thr	CB	C	71.135	0.000	1
204	46	Thr	H	H	7.006	0.000	1
205	46	Thr	N	N	102.381	0.000	1
206	47	Gly	C	C	173.750	0.000	1
207	47	Gly	CA	C	44.965	0.000	1
208	47	Gly	H	H	8.705	0.000	1
209	47	Gly	N	N	112.957	0.000	1
210	48	Lys	C	C	174.789	0.000	1
211	48	Lys	CA	C	56.742	0.000	1
212	48	Lys	CB	C	33.069	0.000	1
213	48	Lys	H	H	8.038	0.000	1
214	48	Lys	N	N	121.845	0.000	1
215	49	Leu	C	C	176.150	0.000	1
216	49	Leu	CA	C	53.140	0.000	1
217	49	Leu	CB	C	43.199	0.000	1
218	49	Leu	H	H	7.688	0.000	1
219	49	Leu	N	N	120.914	0.000	1
220	50	Phe	C	C	173.401	0.000	1
221	50	Phe	CA	C	57.154	0.000	1
222	50	Phe	CB	C	44.217	0.000	1

223	50	Phe	H	H	9.091	0.000	1
224	50	Phe	N	N	118.364	0.000	1
225	51	Ala	C	C	175.863	0.000	1
226	51	Ala	CA	C	50.138	0.000	1
227	51	Ala	CB	C	19.374	0.000	1
228	51	Ala	H	H	8.312	0.000	1
229	51	Ala	N	N	123.700	0.000	1
230	52	Val	C	C	174.900	0.000	1
231	52	Val	CA	C	60.235	0.000	1
232	52	Val	CB	C	32.518	0.000	1
233	52	Val	H	H	9.154	0.000	1
234	52	Val	N	N	123.133	0.000	1
235	53	Lys	C	C	174.826	0.000	1
236	53	Lys	CA	C	54.956	0.000	1
237	53	Lys	CB	C	32.980	0.000	1
238	53	Lys	H	H	9.375	0.000	1
239	53	Lys	N	N	130.181	0.000	1
240	54	Cys	C	C	174.733	0.000	1
241	54	Cys	CA	C	56.790	0.000	1
242	54	Cys	CB	C	28.050	0.000	1
243	54	Cys	H	H	9.267	0.000	1
244	54	Cys	N	N	127.319	0.000	1
245	55	Ile	C	C	173.512	0.000	1
246	55	Ile	CA	C	57.640	0.000	1
247	55	Ile	CB	C	40.691	0.000	1
248	55	Ile	H	H	9.172	0.000	1
249	55	Ile	N	N	127.443	0.000	1
250	57	Lys	C	C	179.533	0.000	1
251	57	Lys	CA	C	59.371	0.000	1
252	57	Lys	CB	C	31.430	0.000	1
253	58	Lys	C	C	177.433	0.000	1
254	58	Lys	CA	C	58.231	0.000	1
255	58	Lys	CB	C	31.170	0.000	1
256	58	Lys	H	H	8.460	0.000	1
257	58	Lys	N	N	119.283	0.000	1
258	59	Ala	C	C	177.529	0.000	1
259	59	Ala	CA	C	52.520	0.000	1
260	59	Ala	CB	C	18.170	0.000	1
261	59	Ala	H	H	7.367	0.000	1
262	59	Ala	N	N	120.285	0.000	1
263	60	Leu	C	C	176.487	0.000	1
264	60	Leu	CA	C	53.644	0.000	1
265	60	Leu	CB	C	41.043	0.000	1
266	60	Leu	H	H	7.332	0.000	1
267	60	Leu	N	N	116.792	0.000	1
268	61	Lys	C	C	177.665	0.000	1
269	61	Lys	CA	C	57.129	0.000	1
270	61	Lys	CB	C	31.378	0.000	1
271	61	Lys	H	H	7.122	0.000	1
272	61	Lys	N	N	120.233	0.000	1
273	62	Gly	C	C	174.516	0.000	1
274	62	Gly	CA	C	45.005	0.000	1
275	62	Gly	H	H	8.786	0.000	1
276	62	Gly	N	N	114.605	0.000	1
277	63	Lys	CA	C	56.879	0.000	1
278	63	Lys	CB	C	32.759	0.000	1
279	63	Lys	H	H	7.743	0.000	1
280	63	Lys	N	N	126.300	0.000	1

281	67	Ile	CA	C	61.152	0.000	1
282	67	Ile	CB	C	37.729	0.000	1
283	68	Glu	CA	C	59.063	0.000	1
284	68	Glu	CB	C	28.935	0.000	1
285	68	Glu	H	H	7.814	0.000	1
286	68	Glu	N	N	120.631	0.000	1
287	71	Ile	CA	C	61.177	0.000	1
288	72	Ala	C	C	179.396	0.000	1
289	72	Ala	CA	C	54.424	0.000	1
290	72	Ala	CB	C	16.646	0.000	1
291	72	Ala	H	H	7.945	0.000	1
292	72	Ala	N	N	122.971	0.000	1
293	73	Val	C	C	177.515	0.000	1
294	73	Val	CA	C	59.143	0.000	1
295	73	Val	CB	C	31.221	0.000	1
296	73	Val	H	H	6.866	0.000	1
297	73	Val	N	N	115.034	0.000	1
298	74	Leu	C	C	177.404	0.000	1
299	74	Leu	CA	C	57.315	0.000	1
300	74	Leu	CB	C	42.915	0.000	1
301	74	Leu	H	H	7.912	0.000	1
302	74	Leu	N	N	119.064	0.000	1
303	78	Lys	C	C	172.981	0.000	1
304	79	His	C	C	175.213	0.000	1
305	79	His	CA	C	57.293	0.000	1
306	79	His	CB	C	33.729	0.000	1
307	79	His	H	H	8.988	0.000	1
308	79	His	N	N	126.128	0.000	1
309	80	Glu	C	C	176.612	0.000	1
310	80	Glu	CA	C	58.396	0.000	1
311	80	Glu	CB	C	29.403	0.000	1
312	80	Glu	H	H	7.911	0.000	1
313	80	Glu	N	N	126.583	0.000	1
314	81	Asn	C	C	173.722	0.000	1
315	81	Asn	CA	C	53.852	0.000	1
316	81	Asn	CB	C	39.861	0.000	1
317	81	Asn	H	H	11.519	0.000	1
318	81	Asn	N	N	119.225	0.000	1
319	82	Ile	C	C	175.414	0.000	1
320	82	Ile	CA	C	61.089	0.000	1
321	82	Ile	CB	C	40.500	0.000	1
322	82	Ile	H	H	7.709	0.000	1
323	82	Ile	N	N	118.802	0.000	1
324	83	Val	CA	C	61.793	0.000	1
325	83	Val	CB	C	31.827	0.000	1
326	84	Ala	CA	C	52.577	0.000	1
327	84	Ala	CB	C	18.304	0.000	1
328	84	Ala	H	H	8.185	0.000	1
329	84	Ala	N	N	125.802	0.000	1
330	86	Glu	C	C	176.760	0.000	1
331	86	Glu	CA	C	57.025	0.000	1
332	86	Glu	CB	C	29.678	0.000	1
333	87	Asp	C	C	173.982	0.000	1
334	87	Asp	CA	C	53.239	0.000	1
335	87	Asp	CB	C	45.639	0.000	1
336	87	Asp	H	H	8.015	0.000	1
337	87	Asp	N	N	116.869	0.000	1
338	88	Ile	C	C	174.198	0.000	1

339	88	Ile	CA	C	60.390	0.000	1
340	88	Ile	CB	C	40.637	0.000	1
341	88	Ile	H	H	8.603	0.000	1
342	88	Ile	N	N	118.345	0.000	1
343	89	Tyr	C	C	174.601	0.000	1
344	89	Tyr	CA	C	55.774	0.000	1
345	89	Tyr	CB	C	43.116	0.000	1
346	89	Tyr	H	H	9.104	0.000	1
347	89	Tyr	N	N	123.689	0.000	1
348	90	Glu	C	C	174.217	0.000	1
349	90	Glu	CA	C	54.737	0.000	1
350	90	Glu	CB	C	32.554	0.000	1
351	90	Glu	H	H	9.217	0.000	1
352	90	Glu	N	N	121.630	0.000	1
353	91	Ser	C	C	173.767	0.000	1
354	91	Ser	CA	C	56.171	0.000	1
355	91	Ser	CB	C	64.901	0.000	1
356	91	Ser	H	H	9.468	0.000	1
357	91	Ser	N	N	127.001	0.000	1
358	92	Pro	C	C	177.431	0.000	1
359	93	Asn	C	C	175.503	0.000	1
360	93	Asn	CA	C	52.700	0.000	1
361	93	Asn	CB	C	40.695	0.000	1
362	93	Asn	H	H	8.201	0.000	1
363	93	Asn	N	N	108.322	0.000	1
364	94	His	C	C	172.268	0.000	1
365	94	His	CA	C	57.243	0.000	1
366	94	His	CB	C	34.299	0.000	1
367	94	His	H	H	8.079	0.000	1
368	94	His	N	N	121.586	0.000	1
369	95	Leu	C	C	174.316	0.000	1
370	95	Leu	CA	C	53.417	0.000	1
371	95	Leu	CB	C	45.029	0.000	1
372	95	Leu	H	H	8.869	0.000	1
373	95	Leu	N	N	120.928	0.000	1
374	96	Tyr	C	C	174.787	0.000	1
375	96	Tyr	CA	C	56.024	0.000	1
376	96	Tyr	CB	C	39.002	0.000	1
377	96	Tyr	H	H	9.277	0.000	1
378	96	Tyr	N	N	120.338	0.000	1
379	97	Leu	C	C	174.719	0.000	1
380	97	Leu	CA	C	53.788	0.000	1
381	97	Leu	CB	C	40.718	0.000	1
382	97	Leu	H	H	9.192	0.000	1
383	97	Leu	N	N	124.497	0.000	1
384	98	Val	C	C	173.931	0.000	1
385	98	Val	CA	C	62.554	0.000	1
386	98	Val	CB	C	30.064	0.000	1
387	98	Val	H	H	8.767	0.000	1
388	98	Val	N	N	126.159	0.000	1
389	99	Met	C	C	173.767	0.000	1
390	99	Met	CA	C	53.217	0.000	1
391	99	Met	CB	C	33.769	0.000	1
392	99	Met	H	H	9.446	0.000	1
393	99	Met	N	N	127.409	0.000	1
394	103	Ser	C	C	175.019	0.000	1
395	103	Ser	CA	C	57.956	0.000	1
396	103	Ser	CB	C	64.269	0.000	1

397	104	Gly	C	C	173.835	0.000	1
398	104	Gly	CA	C	45.776	0.000	1
399	104	Gly	H	H	7.608	0.000	1
400	104	Gly	N	N	109.868	0.000	1
401	105	Gly	C	C	172.396	0.000	1
402	105	Gly	CA	C	43.363	0.000	1
403	105	Gly	H	H	8.189	0.000	1
404	105	Gly	N	N	109.460	0.000	1
405	106	Glu	C	C	177.924	0.000	1
406	106	Glu	CA	C	56.213	0.000	1
407	106	Glu	CB	C	30.070	0.000	1
408	106	Glu	H	H	8.561	0.000	1
409	106	Glu	N	N	117.915	0.000	1
410	107	Leu	C	C	178.455	0.000	1
411	107	Leu	CA	C	58.335	0.000	1
412	107	Leu	CB	C	41.428	0.000	1
413	107	Leu	H	H	8.070	0.000	1
414	107	Leu	N	N	123.450	0.000	1
415	108	Phe	C	C	178.128	0.000	1
416	108	Phe	CA	C	58.005	0.000	1
417	108	Phe	CB	C	36.321	0.000	1
418	108	Phe	H	H	8.798	0.000	1
419	108	Phe	N	N	114.970	0.000	1
420	109	Asp	C	C	178.048	0.000	1
421	109	Asp	CA	C	57.147	0.000	1
422	109	Asp	CB	C	39.975	0.000	1
423	109	Asp	H	H	7.169	0.000	1
424	109	Asp	N	N	117.784	0.000	1
425	110	Arg	C	C	179.444	0.000	1
426	110	Arg	CA	C	56.734	0.000	1
427	110	Arg	CB	C	28.339	0.000	1
428	110	Arg	H	H	7.432	0.000	1
429	110	Arg	N	N	118.027	0.000	1
430	111	Ile	C	C	178.533	0.000	1
431	111	Ile	CA	C	64.852	0.000	1
432	111	Ile	CB	C	36.809	0.000	1
433	111	Ile	H	H	7.744	0.000	1
434	111	Ile	N	N	118.789	0.000	1
435	112	Val	CA	C	66.352	0.000	1
436	112	Val	CB	C	30.813	0.000	1
437	113	Glu	CA	C	56.710	0.000	1
438	113	Glu	CB	C	29.561	0.000	1
439	113	Glu	H	H	7.503	0.000	1
440	113	Glu	N	N	117.157	0.000	1
441	114	Lys	C	C	177.056	0.000	1
442	114	Lys	CA	C	58.247	0.000	1
443	114	Lys	CB	C	33.057	0.000	1
444	114	Lys	H	H	7.387	0.000	1
445	114	Lys	N	N	120.316	0.000	1
446	115	Gly	C	C	172.530	0.000	1
447	115	Gly	CA	C	45.133	0.000	1
448	115	Gly	H	H	6.844	0.000	1
449	115	Gly	N	N	104.228	0.000	1
450	116	Phe	C	C	173.693	0.000	1
451	116	Phe	CA	C	56.835	0.000	1
452	116	Phe	CB	C	40.750	0.000	1
453	116	Phe	H	H	6.977	0.000	1
454	116	Phe	N	N	120.808	0.000	1

455	117	Tyr	C	C	172.305	0.000	1
456	117	Tyr	CA	C	57.269	0.000	1
457	117	Tyr	CB	C	41.784	0.000	1
458	117	Tyr	H	H	8.055	0.000	1
459	117	Tyr	N	N	129.204	0.000	1
460	118	Thr	C	C	174.685	0.000	1
461	118	Thr	CA	C	58.544	0.000	1
462	118	Thr	CB	C	73.374	0.000	1
463	118	Thr	H	H	8.533	0.000	1
464	118	Thr	N	N	118.063	0.000	1
465	119	Glu	C	C	178.429	0.000	1
466	119	Glu	CA	C	59.437	0.000	1
467	119	Glu	CB	C	30.037	0.000	1
468	119	Glu	H	H	9.711	0.000	1
469	119	Glu	N	N	119.972	0.000	1
470	120	Lys	C	C	179.175	0.000	1
471	120	Lys	CA	C	58.832	0.000	1
472	120	Lys	CB	C	30.523	0.000	1
473	120	Lys	H	H	7.962	0.000	1
474	120	Lys	N	N	119.326	0.000	1
475	121	Asp	C	C	178.751	0.000	1
476	121	Asp	CA	C	57.385	0.000	1
477	121	Asp	CB	C	40.753	0.000	1
478	121	Asp	H	H	7.471	0.000	1
479	121	Asp	N	N	121.375	0.000	1
480	122	Ala	C	C	178.454	0.000	1
481	122	Ala	CA	C	55.305	0.000	1
482	122	Ala	CB	C	17.891	0.000	1
483	122	Ala	H	H	7.363	0.000	1
484	122	Ala	N	N	123.728	0.000	1
485	123	Ser	C	C	176.160	0.000	1
486	123	Ser	CA	C	60.179	0.000	1
487	123	Ser	CB	C	60.859	0.000	1
488	123	Ser	H	H	8.471	0.000	1
489	123	Ser	N	N	113.858	0.000	1
490	124	Thr	C	C	177.286	0.000	1
491	124	Thr	CA	C	66.096	0.000	1
492	124	Thr	CB	C	68.483	0.000	1
493	124	Thr	H	H	7.825	0.000	1
494	124	Thr	N	N	116.940	0.000	1
495	125	Leu	C	C	178.635	0.000	1
496	125	Leu	CA	C	58.075	0.000	1
497	125	Leu	CB	C	41.617	0.000	1
498	125	Leu	H	H	7.457	0.000	1
499	125	Leu	N	N	123.300	0.000	1
500	126	Ile	C	C	177.594	0.000	1
501	126	Ile	CA	C	61.351	0.000	1
502	126	Ile	CB	C	34.016	0.000	1
503	126	Ile	H	H	8.394	0.000	1
504	126	Ile	N	N	119.052	0.000	1
505	127	Arg	C	C	178.597	0.000	1
506	127	Arg	CA	C	60.180	0.000	1
507	127	Arg	CB	C	28.985	0.000	1
508	127	Arg	H	H	9.112	0.000	1
509	127	Arg	N	N	122.190	0.000	1
510	128	Gln	C	C	177.578	0.000	1
511	128	Gln	CA	C	58.832	0.000	1
512	128	Gln	CB	C	30.301	0.000	1

513	128	Gln	H	H	7.026	0.000	1
514	128	Gln	N	N	117.457	0.000	1
515	129	Val	C	C	177.629	0.000	1
516	129	Val	CA	C	66.569	0.000	1
517	129	Val	CB	C	30.675	0.000	1
518	129	Val	H	H	7.490	0.000	1
519	129	Val	N	N	118.829	0.000	1
520	130	Leu	C	C	178.445	0.000	1
521	130	Leu	CA	C	57.608	0.000	1
522	130	Leu	CB	C	41.994	0.000	1
523	130	Leu	H	H	9.101	0.000	1
524	130	Leu	N	N	118.384	0.000	1
525	131	Asp	C	C	178.583	0.000	1
526	131	Asp	CA	C	57.467	0.000	1
527	131	Asp	CB	C	41.373	0.000	1
528	131	Asp	H	H	7.859	0.000	1
529	131	Asp	N	N	120.032	0.000	1
530	132	Ala	C	C	178.634	0.000	1
531	132	Ala	CA	C	54.081	0.000	1
532	132	Ala	CB	C	18.740	0.000	1
533	132	Ala	H	H	7.754	0.000	1
534	132	Ala	N	N	121.891	0.000	1
535	133	Val	C	C	176.413	0.000	1
536	133	Val	CA	C	65.917	0.000	1
537	133	Val	CB	C	29.402	0.000	1
538	133	Val	H	H	8.658	0.000	1
539	133	Val	N	N	117.882	0.000	1
540	134	Tyr	C	C	177.573	0.000	1
541	134	Tyr	CA	C	58.999	0.000	1
542	134	Tyr	CB	C	37.202	0.000	1
543	134	Tyr	H	H	8.895	0.000	1
544	134	Tyr	N	N	122.253	0.000	1
545	135	Tyr	C	C	177.276	0.000	1
546	135	Tyr	CA	C	61.365	0.000	1
547	135	Tyr	CB	C	37.078	0.000	1
548	135	Tyr	H	H	7.484	0.000	1
549	135	Tyr	N	N	119.700	0.000	1
550	136	Leu	C	C	179.489	0.000	1
551	136	Leu	CA	C	57.777	0.000	1
552	136	Leu	CB	C	39.142	0.000	1
553	136	Leu	H	H	7.785	0.000	1
554	136	Leu	N	N	119.954	0.000	1
555	137	His	C	C	181.194	0.000	1
556	137	His	CA	C	57.810	0.000	1
557	137	His	CB	C	30.768	0.000	1
558	137	His	H	H	9.407	0.000	1
559	137	His	N	N	121.229	0.000	1
560	138	Arg	C	C	178.444	0.000	1
561	138	Arg	CA	C	58.874	0.000	1
562	138	Arg	CB	C	28.935	0.000	1
563	138	Arg	H	H	8.611	0.000	1
564	138	Arg	N	N	122.507	0.000	1
565	139	Met	C	C	176.078	0.000	1
566	139	Met	CA	C	54.574	0.000	1
567	139	Met	CB	C	31.880	0.000	1
568	139	Met	H	H	7.494	0.000	1
569	139	Met	N	N	117.318	0.000	1
570	140	Gly	C	C	173.779	0.000	1

571	140	Gly	CA	C	45.365	0.000	1
572	140	Gly	H	H	7.888	0.000	1
573	140	Gly	N	N	108.330	0.000	1
574	141	Ile	C	C	173.316	0.000	1
575	141	Ile	CA	C	59.830	0.000	1
576	141	Ile	CB	C	38.579	0.000	1
577	141	Ile	H	H	8.218	0.000	1
578	141	Ile	N	N	124.395	0.000	1
579	142	Val	C	C	177.955	0.000	1
580	142	Val	CA	C	61.636	0.000	1
581	142	Val	CB	C	30.718	0.000	1
582	143	His	C	C	179.558	0.000	1
583	143	His	CA	C	59.492	0.000	1
584	143	His	CB	C	31.272	0.000	1
585	143	His	H	H	8.178	0.000	1
586	143	His	N	N	125.529	0.000	1
587	145	Asp	C	C	175.437	0.000	1
588	145	Asp	CA	C	52.859	0.000	1
589	145	Asp	CB	C	42.609	0.000	1
590	146	Leu	C	C	175.209	0.000	1
591	146	Leu	CA	C	55.833	0.000	1
592	146	Leu	CB	C	40.770	0.000	1
593	146	Leu	H	H	8.476	0.000	1
594	146	Leu	N	N	123.702	0.000	1
595	148	Pro	C	C	176.963	0.000	1
596	149	Glu	C	C	176.262	0.000	1
597	149	Glu	CA	C	58.635	0.000	1
598	149	Glu	CB	C	28.193	0.000	1
599	149	Glu	H	H	9.586	0.000	1
600	149	Glu	N	N	114.802	0.000	1
601	150	Asn	C	C	173.285	0.000	1
602	150	Asn	CA	C	51.914	0.000	1
603	150	Asn	CB	C	38.440	0.000	1
604	150	Asn	H	H	7.338	0.000	1
605	150	Asn	N	N	114.986	0.000	1
606	151	Leu	C	C	173.585	0.000	1
607	151	Leu	CA	C	53.144	0.000	1
608	151	Leu	CB	C	40.649	0.000	1
609	151	Leu	H	H	7.289	0.000	1
610	151	Leu	N	N	121.731	0.000	1
611	152	Leu	C	C	177.043	0.000	1
612	152	Leu	CA	C	51.915	0.000	1
613	152	Leu	CB	C	45.572	0.000	1
614	152	Leu	H	H	8.286	0.000	1
615	152	Leu	N	N	121.399	0.000	1
616	153	Tyr	C	C	176.174	0.000	1
617	153	Tyr	CA	C	58.998	0.000	1
618	153	Tyr	CB	C	38.171	0.000	1
619	153	Tyr	H	H	8.822	0.000	1
620	153	Tyr	N	N	120.307	0.000	1
621	154	Tyr	C	C	175.634	0.000	1
622	154	Tyr	CA	C	60.611	0.000	1
623	154	Tyr	CB	C	39.038	0.000	1
624	154	Tyr	H	H	7.954	0.000	1
625	154	Tyr	N	N	121.624	0.000	1
626	155	Ser	C	C	173.921	0.000	1
627	155	Ser	CA	C	56.033	0.000	1
628	155	Ser	CB	C	66.171	0.000	1

629	155	Ser	H	H	9.348	0.000	1
630	155	Ser	N	N	113.812	0.000	1
631	156	Gln	C	C	175.302	0.000	1
632	156	Gln	CA	C	56.321	0.000	1
633	156	Gln	CB	C	27.802	0.000	1
634	156	Gln	H	H	8.621	0.000	1
635	156	Gln	N	N	118.015	0.000	1
636	157	Asp	C	C	177.625	0.000	1
637	157	Asp	CA	C	54.510	0.000	1
638	157	Asp	CB	C	41.188	0.000	1
639	157	Asp	H	H	7.862	0.000	1
640	157	Asp	N	N	119.548	0.000	1
641	158	Glu	C	C	177.863	0.000	1
642	158	Glu	CA	C	58.693	0.000	1
643	158	Glu	CB	C	28.872	0.000	1
644	158	Glu	H	H	8.947	0.000	1
645	158	Glu	N	N	125.785	0.000	1
646	159	Glu	C	C	175.745	0.000	1
647	159	Glu	CA	C	55.190	0.000	1
648	159	Glu	CB	C	27.817	0.000	1
649	159	Glu	H	H	8.625	0.000	1
650	159	Glu	N	N	117.492	0.000	1
651	160	Ser	C	C	175.366	0.000	1
652	160	Ser	CA	C	58.151	0.000	1
653	160	Ser	CB	C	64.403	0.000	1
654	160	Ser	H	H	7.236	0.000	1
655	160	Ser	N	N	113.430	0.000	1
656	161	Lys	C	C	175.787	0.000	1
657	161	Lys	CA	C	55.818	0.000	1
658	161	Lys	CB	C	32.820	0.000	1
659	161	Lys	H	H	8.485	0.000	1
660	161	Lys	N	N	126.337	0.000	1
661	162	Ile	C	C	172.713	0.000	1
662	162	Ile	CA	C	56.289	0.000	1
663	162	Ile	CB	C	40.379	0.000	1
664	162	Ile	H	H	8.271	0.000	1
665	162	Ile	N	N	122.157	0.000	1
666	163	Met	C	C	175.759	0.000	1
667	163	Met	CA	C	52.316	0.000	1
668	163	Met	CB	C	34.877	0.000	1
669	163	Met	H	H	8.765	0.000	1
670	163	Met	N	N	120.235	0.000	1
671	164	Ile	C	C	176.519	0.000	1
672	164	Ile	CA	C	61.638	0.000	1
673	164	Ile	CB	C	37.431	0.000	1
674	164	Ile	H	H	8.713	0.000	1
675	164	Ile	N	N	121.189	0.000	1
676	170	Ser	C	C	179.850	0.000	1
677	170	Ser	CA	C	61.571	0.000	1
678	171	Lys	C	C	176.796	0.000	1
679	171	Lys	CA	C	57.744	0.000	1
680	171	Lys	CB	C	29.442	0.000	1
681	171	Lys	H	H	8.052	0.000	1
682	171	Lys	N	N	115.940	0.000	1
683	172	Met	C	C	177.342	0.000	1
684	172	Met	CA	C	55.889	0.000	1
685	172	Met	CB	C	32.443	0.000	1
686	172	Met	H	H	7.491	0.000	1

687	172	Met	N	N	117.066	0.000	1
688	173	Glu	C	C	174.534	0.000	1
689	173	Glu	CA	C	57.757	0.000	1
690	173	Glu	CB	C	28.147	0.000	1
691	174	Gly	C	C	173.746	0.000	1
692	174	Gly	CA	C	44.691	0.000	1
693	174	Gly	H	H	8.212	0.000	1
694	174	Gly	N	N	110.802	0.000	1
695	175	Lys	C	C	177.081	0.000	1
696	175	Lys	CA	C	55.858	0.000	1
697	175	Lys	CB	C	31.977	0.000	1
698	176	Gly	C	C	173.769	0.000	1
699	176	Gly	CA	C	44.942	0.000	1
700	176	Gly	H	H	8.407	0.000	1
701	176	Gly	N	N	110.252	0.000	1
702	177	Asp	C	C	176.356	0.000	1
703	177	Asp	CA	C	53.825	0.000	1
704	177	Asp	CB	C	40.734	0.000	1
705	177	Asp	H	H	8.137	0.000	1
706	177	Asp	N	N	120.890	0.000	1
707	178	Val	CA	C	62.032	0.000	1
708	178	Val	CB	C	31.934	0.000	1
709	178	Val	H	H	8.054	0.000	1
710	178	Val	N	N	120.212	0.000	1
711	183	Cys	C	C	177.119	0.000	1
712	184	Gly	C	C	174.289	0.000	1
713	184	Gly	CA	C	44.916	0.000	1
714	184	Gly	H	H	8.592	0.000	1
715	184	Gly	N	N	109.781	0.000	1
716	185	Thr	C	C	174.732	0.000	1
717	185	Thr	CA	C	57.751	0.000	1
718	185	Thr	CB	C	65.427	0.000	1
719	185	Thr	H	H	8.792	0.000	1
720	185	Thr	N	N	118.044	0.000	1
721	188	Tyr	C	C	175.311	0.000	1
722	188	Tyr	CA	C	56.872	0.000	1
723	188	Tyr	CB	C	38.574	0.000	1
724	189	Val	C	C	173.428	0.000	1
725	189	Val	CA	C	61.249	0.000	1
726	189	Val	CB	C	32.116	0.000	1
727	189	Val	H	H	7.436	0.000	1
728	189	Val	N	N	121.202	0.000	1
729	190	Ala	C	C	175.899	0.000	1
730	190	Ala	CA	C	49.095	0.000	1
731	190	Ala	CB	C	17.539	0.000	1
732	190	Ala	H	H	7.597	0.000	1
733	190	Ala	N	N	125.163	0.000	1
734	191	Pro	C	C	178.193	0.000	1
735	191	Pro	CA	C	64.524	0.000	1
736	191	Pro	CB	C	30.560	0.000	1
737	192	Glu	C	C	177.131	0.000	1
738	192	Glu	CA	C	58.496	0.000	1
739	192	Glu	CB	C	25.544	0.000	1
740	192	Glu	H	H	9.808	0.000	1
741	192	Glu	N	N	115.450	0.000	1
742	193	Val	C	C	180.729	0.000	1
743	193	Val	CA	C	64.434	0.000	1
744	193	Val	CB	C	30.233	0.000	1

745	193	Val	H	H	6.906	0.000	1
746	193	Val	N	N	120.984	0.000	1
747	194	Leu	C	C	178.388	0.000	1
748	194	Leu	CA	C	56.758	0.000	1
749	194	Leu	CB	C	39.166	0.000	1
750	194	Leu	H	H	6.843	0.000	1
751	194	Leu	N	N	121.679	0.000	1
752	195	Ala	C	C	175.713	0.000	1
753	195	Ala	CA	C	51.543	0.000	1
754	195	Ala	CB	C	17.945	0.000	1
755	195	Ala	H	H	7.745	0.000	1
756	195	Ala	N	N	119.203	0.000	1
757	196	Gln	C	C	175.297	0.000	1
758	196	Gln	CA	C	56.362	0.000	1
759	196	Gln	CB	C	24.729	0.000	1
760	196	Gln	H	H	7.699	0.000	1
761	196	Gln	N	N	112.920	0.000	1
762	197	Lys	C	C	173.996	0.000	1
763	197	Lys	CA	C	53.227	0.000	1
764	197	Lys	CB	C	30.689	0.000	1
765	197	Lys	H	H	7.859	0.000	1
766	197	Lys	N	N	119.474	0.000	1
767	198	Pro	C	C	176.455	0.000	1
768	198	Pro	CA	C	62.574	0.000	1
769	198	Pro	CB	C	30.945	0.000	1
770	199	Tyr	C	C	174.228	0.000	1
771	199	Tyr	CA	C	56.803	0.000	1
772	199	Tyr	CB	C	39.190	0.000	1
773	199	Tyr	H	H	7.880	0.000	1
774	199	Tyr	N	N	119.098	0.000	1
775	201	Lys	C	C	175.744	0.000	1
776	201	Lys	CA	C	58.835	0.000	1
777	201	Lys	CB	C	30.474	0.000	1
778	202	Ala	C	C	182.667	0.000	1
779	202	Ala	CA	C	54.293	0.000	1
780	202	Ala	CB	C	18.427	0.000	1
781	202	Ala	H	H	7.905	0.000	1
782	202	Ala	N	N	120.871	0.000	1
783	203	Val	C	C	178.558	0.000	1
784	203	Val	CA	C	65.139	0.000	1
785	203	Val	CB	C	33.249	0.000	1
786	203	Val	H	H	8.370	0.000	1
787	203	Val	N	N	118.324	0.000	1
788	204	Asp	C	C	178.123	0.000	1
789	204	Asp	CA	C	56.485	0.000	1
790	204	Asp	CB	C	40.652	0.000	1
791	204	Asp	H	H	7.772	0.000	1
792	204	Asp	N	N	121.553	0.000	1
793	205	Cys	C	C	175.757	0.000	1
794	205	Cys	CA	C	63.295	0.000	1
795	205	Cys	CB	C	26.545	0.000	1
796	205	Cys	H	H	7.209	0.000	1
797	205	Cys	N	N	119.245	0.000	1
798	206	Trp	C	C	177.429	0.000	1
799	206	Trp	CA	C	60.538	0.000	1
800	206	Trp	CB	C	27.528	0.000	1
801	206	Trp	H	H	7.359	0.000	1
802	206	Trp	N	N	119.734	0.000	1

803	207	Ser	C	C	175.971	0.000	1
804	207	Ser	CA	C	62.962	0.000	1
805	207	Ser	CB	C	66.611	0.000	1
806	207	Ser	H	H	7.857	0.000	1
807	207	Ser	N	N	113.580	0.000	1
808	208	Ile	C	C	176.808	0.000	1
809	208	Ile	CA	C	65.913	0.000	1
810	208	Ile	CB	C	36.749	0.000	1
811	208	Ile	H	H	7.951	0.000	1
812	208	Ile	N	N	120.456	0.000	1
813	209	Gly	C	C	173.185	0.000	1
814	209	Gly	CA	C	48.149	0.000	1
815	209	Gly	H	H	7.648	0.000	1
816	209	Gly	N	N	108.906	0.000	1
817	210	Val	C	C	178.236	0.000	1
818	210	Val	CA	C	65.920	0.000	1
819	210	Val	CB	C	30.711	0.000	1
820	210	Val	H	H	7.635	0.000	1
821	210	Val	N	N	121.960	0.000	1
822	211	Ile	C	C	177.042	0.000	1
823	211	Ile	CA	C	65.794	0.000	1
824	211	Ile	CB	C	37.123	0.000	1
825	211	Ile	H	H	8.098	0.000	1
826	211	Ile	N	N	118.814	0.000	1
827	212	Ala	C	C	176.959	0.000	1
828	212	Ala	CA	C	55.693	0.000	1
829	212	Ala	CB	C	17.088	0.000	1
830	212	Ala	H	H	8.631	0.000	1
831	212	Ala	N	N	121.680	0.000	1
832	214	Ile	C	C	179.222	0.000	1
833	214	Ile	CA	C	64.857	0.000	1
834	214	Ile	CB	C	40.066	0.000	1
835	215	Leu	C	C	177.874	0.000	1
836	215	Leu	CA	C	56.553	0.000	1
837	215	Leu	CB	C	40.373	0.000	1
838	215	Leu	H	H	8.727	0.000	1
839	215	Leu	N	N	118.699	0.000	1
840	216	Leu	C	C	175.259	0.000	1
841	216	Leu	CA	C	53.797	0.000	1
842	216	Leu	CB	C	41.651	0.000	1
843	216	Leu	H	H	7.100	0.000	1
844	216	Leu	N	N	115.379	0.000	1
845	217	Cys	C	C	174.223	0.000	1
846	218	Gly	C	C	170.480	0.000	1
847	218	Gly	CA	C	45.960	0.000	1
848	218	Gly	H	H	9.433	0.000	1
849	218	Gly	N	N	115.601	0.000	1
850	219	Tyr	C	C	170.653	0.000	1
851	219	Tyr	CA	C	53.452	0.000	1
852	219	Tyr	CB	C	37.968	0.000	1
853	219	Tyr	H	H	7.796	0.000	1
854	219	Tyr	N	N	116.754	0.000	1
855	221	Pro	C	C	175.372	0.000	1
856	221	Pro	CA	C	63.986	0.000	1
857	222	Phe	C	C	174.961	0.000	1
858	222	Phe	CA	C	54.713	0.000	1
859	222	Phe	CB	C	40.199	0.000	1
860	222	Phe	H	H	6.937	0.000	1

861	222	Phe	N	N	114.085	0.000	1
862	223	Tyr	C	C	173.129	0.000	1
863	223	Tyr	CA	C	60.245	0.000	1
864	223	Tyr	CB	C	39.909	0.000	1
865	223	Tyr	H	H	8.857	0.000	1
866	223	Tyr	N	N	123.049	0.000	1
867	224	Asp	C	C	173.822	0.000	1
868	224	Asp	CA	C	54.233	0.000	1
869	224	Asp	CB	C	44.693	0.000	1
870	224	Asp	H	H	6.656	0.000	1
871	224	Asp	N	N	126.639	0.000	1
872	225	Glu	C	C	175.682	0.000	1
873	225	Glu	CA	C	57.639	0.000	1
874	225	Glu	CB	C	28.884	0.000	1
875	225	Glu	H	H	8.485	0.000	1
876	225	Glu	N	N	124.081	0.000	1
877	226	Asn	C	C	175.104	0.000	1
878	226	Asn	CA	C	51.894	0.000	1
879	226	Asn	CB	C	39.157	0.000	1
880	226	Asn	H	H	9.566	0.000	1
881	226	Asn	N	N	120.783	0.000	1
882	227	Asp	C	C	177.346	0.000	1
883	227	Asp	CA	C	57.247	0.000	1
884	227	Asp	CB	C	39.800	0.000	1
885	227	Asp	H	H	8.675	0.000	1
886	227	Asp	N	N	126.762	0.000	1
887	228	Ser	C	C	176.961	0.000	1
888	228	Ser	CA	C	61.373	0.000	1
889	228	Ser	H	H	8.053	0.000	1
890	228	Ser	N	N	115.330	0.000	1
891	229	Lys	C	C	178.856	0.000	1
892	229	Lys	CA	C	57.771	0.000	1
893	229	Lys	CB	C	30.739	0.000	1
894	229	Lys	H	H	7.555	0.000	1
895	229	Lys	N	N	123.326	0.000	1
896	230	Leu	C	C	177.709	0.000	1
897	230	Leu	CA	C	57.688	0.000	1
898	230	Leu	CB	C	40.312	0.000	1
899	230	Leu	H	H	7.969	0.000	1
900	230	Leu	N	N	122.527	0.000	1
901	231	Phe	C	C	177.039	0.000	1
902	231	Phe	CA	C	60.062	0.000	1
903	231	Phe	CB	C	37.658	0.000	1
904	231	Phe	H	H	8.029	0.000	1
905	231	Phe	N	N	117.573	0.000	1
906	232	Glu	C	C	178.649	0.000	1
907	232	Glu	CA	C	58.901	0.000	1
908	232	Glu	CB	C	28.949	0.000	1
909	232	Glu	H	H	7.430	0.000	1
910	232	Glu	N	N	116.634	0.000	1
911	233	Gln	C	C	178.791	0.000	1
912	233	Gln	CA	C	58.705	0.000	1
913	233	Gln	CB	C	28.281	0.000	1
914	233	Gln	H	H	7.898	0.000	1
915	233	Gln	N	N	118.991	0.000	1
916	234	Ile	C	C	177.835	0.000	1
917	234	Ile	CA	C	65.001	0.000	1
918	234	Ile	CB	C	36.758	0.000	1

919	234	Ile	H	H	7.725	0.000	1
920	234	Ile	N	N	121.068	0.000	1
921	235	Leu	C	C	179.350	0.000	1
922	235	Leu	CA	C	57.422	0.000	1
923	235	Leu	CB	C	41.206	0.000	1
924	235	Leu	H	H	7.750	0.000	1
925	235	Leu	N	N	117.348	0.000	1
926	236	Lys	C	C	176.383	0.000	1
927	236	Lys	CA	C	56.750	0.000	1
928	236	Lys	CB	C	32.292	0.000	1
929	236	Lys	H	H	7.314	0.000	1
930	236	Lys	N	N	114.898	0.000	1
931	237	Ala	C	C	175.089	0.000	1
932	237	Ala	CA	C	53.020	0.000	1
933	237	Ala	CB	C	17.305	0.000	1
934	237	Ala	H	H	8.186	0.000	1
935	237	Ala	N	N	122.729	0.000	1
936	238	Glu	C	C	174.300	0.000	1
937	238	Glu	CA	C	55.478	0.000	1
938	238	Glu	CB	C	29.422	0.000	1
939	238	Glu	H	H	8.274	0.000	1
940	238	Glu	N	N	119.830	0.000	1
941	239	Tyr	C	C	172.674	0.000	1
942	239	Tyr	CA	C	55.042	0.000	1
943	239	Tyr	CB	C	38.776	0.000	1
944	239	Tyr	H	H	7.298	0.000	1
945	239	Tyr	N	N	120.378	0.000	1
946	240	Glu	C	C	174.168	0.000	1
947	240	Glu	CA	C	53.212	0.000	1
948	240	Glu	CB	C	33.152	0.000	1
949	240	Glu	H	H	8.311	0.000	1
950	240	Glu	N	N	118.288	0.000	1
951	241	Phe	C	C	175.003	0.000	1
952	241	Phe	CA	C	54.280	0.000	1
953	241	Phe	CB	C	35.239	0.000	1
954	241	Phe	H	H	8.733	0.000	1
955	241	Phe	N	N	123.750	0.000	1
956	242	Asp	C	C	176.562	0.000	1
957	242	Asp	CA	C	55.240	0.000	1
958	242	Asp	CB	C	40.627	0.000	1
959	242	Asp	H	H	7.542	0.000	1
960	242	Asp	N	N	119.104	0.000	1
961	243	Ser	C	C	174.338	0.000	1
962	243	Ser	CA	C	57.115	0.000	1
963	243	Ser	CB	C	62.706	0.000	1
964	243	Ser	H	H	8.917	0.000	1
965	243	Ser	N	N	119.678	0.000	1
966	244	Pro	C	C	177.630	0.000	1
967	244	Pro	CA	C	62.818	0.000	1
968	244	Pro	CB	C	33.450	0.000	1
969	245	Tyr	C	C	177.588	0.000	1
970	245	Tyr	CA	C	59.447	0.000	1
971	245	Tyr	CB	C	37.488	0.000	1
972	245	Tyr	H	H	6.561	0.000	1
973	245	Tyr	N	N	124.546	0.000	1
974	246	Trp	C	C	177.520	0.000	1
975	246	Trp	CA	C	57.422	0.000	1
976	246	Trp	CB	C	28.662	0.000	1

977	246	Trp	H	H	8.372	0.000	1
978	246	Trp	N	N	114.680	0.000	1
979	247	Asp	C	C	177.350	0.000	1
980	247	Asp	CA	C	57.320	0.000	1
981	247	Asp	CB	C	38.671	0.000	1
982	247	Asp	H	H	7.757	0.000	1
983	247	Asp	N	N	124.020	0.000	1
984	248	Asp	C	C	175.526	0.000	1
985	248	Asp	CA	C	53.658	0.000	1
986	248	Asp	CB	C	40.607	0.000	1
987	248	Asp	H	H	8.568	0.000	1
988	248	Asp	N	N	116.625	0.000	1
989	249	Ile	C	C	176.487	0.000	1
990	249	Ile	CA	C	57.277	0.000	1
991	249	Ile	CB	C	35.404	0.000	1
992	249	Ile	H	H	7.007	0.000	1
993	249	Ile	N	N	122.412	0.000	1
994	250	Ser	C	C	174.813	0.000	1
995	250	Ser	CA	C	59.357	0.000	1
996	250	Ser	CB	C	63.102	0.000	1
997	250	Ser	H	H	10.543	0.000	1
998	250	Ser	N	N	126.978	0.000	1
999	251	Asp	C	C	179.538	0.000	1
1000	251	Asp	CA	C	56.630	0.000	1
1001	251	Asp	CB	C	39.310	0.000	1
1002	251	Asp	H	H	8.605	0.000	1
1003	251	Asp	N	N	120.309	0.000	1
1004	255	Asp	C	C	177.531	0.000	1
1005	255	Asp	CA	C	57.374	0.000	1
1006	255	Asp	CB	C	42.955	0.000	1
1007	256	Phe	C	C	176.029	0.000	1
1008	256	Phe	CA	C	61.653	0.000	1
1009	256	Phe	CB	C	37.713	0.000	1
1010	256	Phe	H	H	8.070	0.000	1
1011	256	Phe	N	N	121.599	0.000	1
1012	257	Ile	C	C	178.294	0.000	1
1013	257	Ile	CA	C	65.400	0.000	1
1014	257	Ile	CB	C	37.147	0.000	1
1015	257	Ile	H	H	7.706	0.000	1
1016	257	Ile	N	N	116.817	0.000	1
1017	258	Arg	C	C	178.006	0.000	1
1018	258	Arg	CA	C	58.380	0.000	1
1019	258	Arg	CB	C	29.354	0.000	1
1020	258	Arg	H	H	8.514	0.000	1
1021	258	Arg	N	N	119.338	0.000	1
1022	259	Asn	C	C	175.211	0.000	1
1023	259	Asn	CA	C	55.073	0.000	1
1024	259	Asn	CB	C	37.890	0.000	1
1025	259	Asn	H	H	6.962	0.000	1
1026	259	Asn	N	N	113.435	0.000	1
1027	260	Leu	C	C	177.936	0.000	1
1028	260	Leu	CA	C	56.248	0.000	1
1029	260	Leu	CB	C	41.398	0.000	1
1030	260	Leu	H	H	6.912	0.000	1
1031	260	Leu	N	N	121.249	0.000	1
1032	261	Met	C	C	175.499	0.000	1
1033	261	Met	CA	C	54.891	0.000	1
1034	261	Met	CB	C	28.794	0.000	1

1035	262	Glu	C	C	176.417	0.000	1
1036	262	Glu	CA	C	52.391	0.000	1
1037	262	Glu	CB	C	30.493	0.000	1
1038	262	Glu	H	H	7.036	0.000	1
1039	262	Glu	N	N	122.527	0.000	1
1040	263	Lys	C	C	174.955	0.000	1
1041	263	Lys	CA	C	59.798	0.000	1
1042	263	Lys	CB	C	31.975	0.000	1
1043	263	Lys	H	H	9.257	0.000	1
1044	263	Lys	N	N	131.328	0.000	1
1045	264	Asp	C	C	176.199	0.000	1
1046	264	Asp	CA	C	50.745	0.000	1
1047	264	Asp	CB	C	40.823	0.000	1
1048	264	Asp	H	H	9.619	0.000	1
1049	264	Asp	N	N	121.545	0.000	1
1050	265	Pro	C	C	177.091	0.000	1
1051	265	Pro	CA	C	64.197	0.000	1
1052	265	Pro	CB	C	31.045	0.000	1
1053	266	Asn	C	C	176.051	0.000	1
1054	266	Asn	CA	C	54.638	0.000	1
1055	266	Asn	CB	C	38.074	0.000	1
1056	266	Asn	H	H	8.481	0.000	1
1057	266	Asn	N	N	115.620	0.000	1
1058	267	Lys	C	C	176.121	0.000	1
1059	267	Lys	CA	C	55.928	0.000	1
1060	267	Lys	CB	C	33.557	0.000	1
1061	267	Lys	H	H	7.409	0.000	1
1062	267	Lys	N	N	119.134	0.000	1
1063	268	Arg	C	C	175.702	0.000	1
1064	268	Arg	CA	C	55.970	0.000	1
1065	268	Arg	CB	C	30.737	0.000	1
1066	268	Arg	H	H	7.338	0.000	1
1067	268	Arg	N	N	123.110	0.000	1
1068	269	Tyr	C	C	177.782	0.000	1
1069	269	Tyr	CA	C	55.414	0.000	1
1070	269	Tyr	CB	C	36.677	0.000	1
1071	269	Tyr	H	H	8.415	0.000	1
1072	269	Tyr	N	N	127.747	0.000	1
1073	270	Thr	C	C	176.827	0.000	1
1074	270	Thr	CA	C	60.177	0.000	1
1075	270	Thr	CB	C	71.107	0.000	1
1076	270	Thr	H	H	8.161	0.000	1
1077	270	Thr	N	N	110.487	0.000	1
1078	271	Cys	C	C	176.258	0.000	1
1079	271	Cys	CA	C	64.461	0.000	1
1080	271	Cys	CB	C	26.014	0.000	1
1081	271	Cys	H	H	9.142	0.000	1
1082	271	Cys	N	N	119.024	0.000	1
1083	272	Glu	C	C	178.432	0.000	1
1084	272	Glu	CA	C	58.785	0.000	1
1085	272	Glu	CB	C	28.477	0.000	1
1086	272	Glu	H	H	8.267	0.000	1
1087	272	Glu	N	N	118.160	0.000	1
1088	273	Gln	C	C	179.448	0.000	1
1089	273	Gln	CA	C	57.740	0.000	1
1090	273	Gln	CB	C	28.247	0.000	1
1091	273	Gln	H	H	7.131	0.000	1
1092	273	Gln	N	N	117.472	0.000	1

1093	274	Ala	C	C	178.375	0.000	1
1094	274	Ala	CA	C	54.889	0.000	1
1095	274	Ala	CB	C	18.762	0.000	1
1096	274	Ala	H	H	8.471	0.000	1
1097	274	Ala	N	N	122.386	0.000	1
1098	275	Ala	C	C	177.505	0.000	1
1099	275	Ala	CA	C	53.332	0.000	1
1100	275	Ala	CB	C	17.477	0.000	1
1101	275	Ala	H	H	8.141	0.000	1
1102	275	Ala	N	N	116.622	0.000	1
1103	276	Arg	C	C	176.080	0.000	1
1104	276	Arg	CA	C	54.544	0.000	1
1105	276	Arg	CB	C	30.235	0.000	1
1106	276	Arg	H	H	6.936	0.000	1
1107	276	Arg	N	N	112.798	0.000	1
1108	277	His	C	C	175.376	0.000	1
1109	277	His	CA	C	55.438	0.000	1
1110	277	His	CB	C	32.603	0.000	1
1111	277	His	H	H	7.724	0.000	1
1112	277	His	N	N	126.359	0.000	1
1113	278	Pro	C	C	176.938	0.000	1
1114	278	Pro	CA	C	64.955	0.000	1
1115	279	Trp	C	C	178.489	0.000	1
1116	279	Trp	CA	C	61.932	0.000	1
1117	279	Trp	CB	C	27.699	0.000	1
1118	279	Trp	H	H	11.225	0.000	1
1119	279	Trp	N	N	123.062	0.000	1
1120	280	Ile	C	C	177.273	0.000	1
1121	280	Ile	CA	C	58.526	0.000	1
1122	280	Ile	CB	C	35.813	0.000	1
1123	280	Ile	H	H	7.628	0.000	1
1124	280	Ile	N	N	116.368	0.000	1
1125	281	Ala	C	C	178.052	0.000	1
1126	281	Ala	CA	C	52.527	0.000	1
1127	281	Ala	CB	C	19.457	0.000	1
1128	281	Ala	H	H	8.411	0.000	1
1129	281	Ala	N	N	120.826	0.000	1
1130	282	Gly	C	C	175.413	0.000	1
1131	282	Gly	CA	C	44.607	0.000	1
1132	282	Gly	H	H	7.976	0.000	1
1133	282	Gly	N	N	109.977	0.000	1
1134	283	Asp	C	C	176.585	0.000	1
1135	283	Asp	CA	C	53.233	0.000	1
1136	283	Asp	CB	C	40.802	0.000	1
1137	283	Asp	H	H	8.469	0.000	1
1138	283	Asp	N	N	122.272	0.000	1
1139	284	Thr	C	C	175.608	0.000	1
1140	284	Thr	CA	C	61.851	0.000	1
1141	284	Thr	CB	C	68.641	0.000	1
1142	284	Thr	H	H	7.093	0.000	1
1143	284	Thr	N	N	107.887	0.000	1
1144	285	Ala	C	C	178.170	0.000	1
1145	285	Ala	CA	C	52.486	0.000	1
1146	285	Ala	CB	C	16.904	0.000	1
1147	285	Ala	H	H	8.136	0.000	1
1148	285	Ala	N	N	123.486	0.000	1
1149	286	Leu	C	C	173.854	0.000	1
1150	286	Leu	CA	C	54.173	0.000	1

1151	286	Leu	CB	C	41.443	0.000	1
1152	286	Leu	H	H	6.993	0.000	1
1153	286	Leu	N	N	123.494	0.000	1
1154	287	Asn	C	C	174.895	0.000	1
1155	287	Asn	CA	C	50.991	0.000	1
1156	287	Asn	CB	C	39.323	0.000	1
1157	287	Asn	H	H	8.017	0.000	1
1158	287	Asn	N	N	115.080	0.000	1
1159	288	Lys	C	C	176.812	0.000	1
1160	288	Lys	CA	C	54.899	0.000	1
1161	288	Lys	CB	C	32.987	0.000	1
1162	288	Lys	H	H	7.310	0.000	1
1163	288	Lys	N	N	121.129	0.000	1
1164	289	Asn	C	C	176.454	0.000	1
1165	289	Asn	CA	C	53.196	0.000	1
1166	289	Asn	CB	C	36.714	0.000	1
1167	289	Asn	H	H	9.025	0.000	1
1168	289	Asn	N	N	123.539	0.000	1
1169	290	Ile	C	C	175.264	0.000	1
1170	290	Ile	CA	C	61.102	0.000	1
1171	290	Ile	CB	C	37.298	0.000	1
1172	290	Ile	H	H	9.304	0.000	1
1173	290	Ile	N	N	126.705	0.000	1
1174	291	His	C	C	177.522	0.000	1
1175	291	His	CA	C	61.707	0.000	1
1176	291	His	CB	C	32.058	0.000	1
1177	291	His	H	H	8.064	0.000	1
1178	291	His	N	N	120.296	0.000	1
1179	292	Glu	C	C	178.439	0.000	1
1180	292	Glu	CA	C	60.308	0.000	1
1181	292	Glu	CB	C	28.412	0.000	1
1182	292	Glu	H	H	9.010	0.000	1
1183	292	Glu	N	N	120.299	0.000	1
1184	293	Ser	C	C	177.041	0.000	1
1185	293	Ser	CA	C	60.309	0.000	1
1186	293	Ser	CB	C	62.941	0.000	1
1187	293	Ser	H	H	8.272	0.000	1
1188	293	Ser	N	N	112.288	0.000	1
1189	294	Val	C	C	178.552	0.000	1
1190	294	Val	CA	C	66.375	0.000	1
1191	294	Val	CB	C	31.377	0.000	1
1192	294	Val	H	H	7.531	0.000	1
1193	294	Val	N	N	117.726	0.000	1
1194	295	Ser	C	C	175.932	0.000	1
1195	295	Ser	CA	C	61.563	0.000	1
1196	295	Ser	CB	C	62.336	0.000	1
1197	295	Ser	H	H	8.501	0.000	1
1198	295	Ser	N	N	112.978	0.000	1
1199	296	Ala	C	C	181.242	0.000	1
1200	296	Ala	CA	C	54.745	0.000	1
1201	296	Ala	CB	C	17.531	0.000	1
1202	296	Ala	H	H	7.159	0.000	1
1203	296	Ala	N	N	124.171	0.000	1
1204	297	Gln	C	C	179.273	0.000	1
1205	297	Gln	CA	C	57.508	0.000	1
1206	297	Gln	CB	C	26.357	0.000	1
1207	297	Gln	H	H	7.884	0.000	1
1208	297	Gln	N	N	117.173	0.000	1

1209	298	Ile	C	C	178.866	0.000	1
1210	298	Ile	CA	C	65.006	0.000	1
1211	298	Ile	CB	C	36.774	0.000	1
1212	298	Ile	H	H	8.885	0.000	1
1213	298	Ile	N	N	124.588	0.000	1
1214	299	Arg	C	C	178.171	0.000	1
1215	299	Arg	CA	C	58.785	0.000	1
1216	299	Arg	CB	C	29.466	0.000	1
1217	299	Arg	H	H	7.905	0.000	1
1218	299	Arg	N	N	116.951	0.000	1
1219	300	Lys	C	C	177.864	0.000	1
1220	300	Lys	CA	C	57.560	0.000	1
1221	300	Lys	CB	C	33.181	0.000	1
1222	300	Lys	H	H	7.199	0.000	1
1223	300	Lys	N	N	116.395	0.000	1
1224	301	Asn	C	C	174.694	0.000	1
1225	301	Asn	CA	C	55.248	0.000	1
1226	301	Asn	CB	C	40.697	0.000	1
1227	301	Asn	H	H	8.094	0.000	1
1228	301	Asn	N	N	114.786	0.000	1
1229	302	Phe	C	C	175.039	0.000	1
1230	302	Phe	CA	C	56.525	0.000	1
1231	302	Phe	CB	C	38.628	0.000	1
1232	302	Phe	H	H	8.395	0.000	1
1233	302	Phe	N	N	119.653	0.000	1
1234	303	Ala	C	C	177.895	0.000	1
1235	303	Ala	CA	C	52.787	0.000	1
1236	303	Ala	CB	C	18.590	0.000	1
1237	303	Ala	H	H	7.782	0.000	1
1238	303	Ala	N	N	123.728	0.000	1
1239	306	Lys	C	C	176.713	0.000	1
1240	306	Lys	CA	C	56.597	0.000	1
1241	306	Lys	CB	C	31.483	0.000	1
1242	307	Trp	C	C	176.490	0.000	1
1243	307	Trp	CA	C	56.696	0.000	1
1244	307	Trp	CB	C	28.873	0.000	1
1245	307	Trp	H	H	7.900	0.000	1
1246	307	Trp	N	N	121.222	0.000	1
1247	308	Arg	C	C	176.176	0.000	1
1248	308	Arg	CA	C	55.994	0.000	1
1249	308	Arg	CB	C	29.814	0.000	1
1250	308	Arg	H	H	7.953	0.000	1
1251	308	Arg	N	N	122.213	0.000	1
1252	309	Gln	C	C	175.732	0.000	1
1253	309	Gln	CA	C	55.584	0.000	1
1254	309	Gln	CB	C	28.716	0.000	1
1255	309	Gln	H	H	8.143	0.000	1
1256	309	Gln	N	N	121.209	0.000	1
1257	310	Ala	C	C	177.394	0.000	1
1258	310	Ala	CA	C	52.110	0.000	1
1259	310	Ala	CB	C	18.558	0.000	1
1260	310	Ala	H	H	8.153	0.000	1
1261	310	Ala	N	N	124.791	0.000	1
1262	311	Phe	C	C	175.329	0.000	1
1263	311	Phe	CA	C	57.401	0.000	1
1264	311	Phe	CB	C	38.890	0.000	1
1265	311	Phe	H	H	8.070	0.000	1
1266	311	Phe	N	N	119.399	0.000	1

1267	312	Asn	C	C	174.513	0.000	1
1268	312	Asn	CA	C	52.582	0.000	1
1269	312	Asn	CB	C	38.437	0.000	1
1270	312	Asn	H	H	8.186	0.000	1
1271	312	Asn	N	N	120.978	0.000	1
1272	313	Ala	C	C	177.747	0.000	1
1273	313	Ala	CA	C	52.358	0.000	1
1274	313	Ala	CB	C	18.657	0.000	1
1275	313	Ala	H	H	8.117	0.000	1
1276	313	Ala	N	N	125.059	0.000	1
1277	314	Thr	C	C	174.137	0.000	1
1278	314	Thr	CA	C	61.527	0.000	1
1279	314	Thr	CB	C	69.160	0.000	1
1280	314	Thr	H	H	7.971	0.000	1
1281	314	Thr	N	N	113.203	0.000	1
1282	315	Ala	C	C	177.310	0.000	1
1283	315	Ala	CA	C	51.934	0.000	1
1284	315	Ala	CB	C	18.670	0.000	1
1285	315	Ala	H	H	8.062	0.000	1
1286	315	Ala	N	N	126.949	0.000	1
1287	316	Val	C	C	175.362	0.000	1
1288	316	Val	CA	C	61.982	0.000	1
1289	316	Val	CB	C	31.826	0.000	1
1290	316	Val	H	H	8.053	0.000	1
1291	316	Val	N	N	120.902	0.000	1
1292	317	Val	C	C	180.921	0.000	1
1293	317	Val	CA	C	63.072	0.000	1
1294	317	Val	CB	C	32.601	0.000	1
1295	317	Val	H	H	7.635	0.000	1
1296	317	Val	N	N	128.406	0.000	1
1297	326	Gly	C	C	174.244	0.000	1
1298	326	Gly	CA	C	44.814	0.000	1
1299	327	Ser	C	C	174.785	0.000	1
1300	327	Ser	CA	C	57.912	0.000	1
1301	327	Ser	CB	C	63.274	0.000	1
1302	327	Ser	H	H	8.151	0.000	1
1303	327	Ser	N	N	116.144	0.000	1
1304	328	Ser	C	C	174.611	0.000	1
1305	328	Ser	CA	C	58.213	0.000	1
1306	328	Ser	CB	C	63.514	0.000	1
1307	329	Leu	C	C	177.259	0.000	1
1308	329	Leu	CA	C	54.938	0.000	1
1309	329	Leu	CB	C	41.218	0.000	1
1310	329	Leu	H	H	8.195	0.000	1
1311	329	Leu	N	N	124.036	0.000	1
1312	330	Asp	C	C	176.432	0.000	1
1313	330	Asp	CA	C	54.023	0.000	1
1314	330	Asp	CB	C	40.611	0.000	1
1315	330	Asp	H	H	8.230	0.000	1
1316	330	Asp	N	N	121.436	0.000	1
1317	331	Ser	C	C	174.864	0.000	1
1318	331	Ser	CA	C	58.190	0.000	1
1319	331	Ser	CB	C	63.398	0.000	1
1320	331	Ser	H	H	8.185	0.000	1
1321	331	Ser	N	N	117.014	0.000	1
1322	332	Ser	C	C	174.350	0.000	1
1323	332	Ser	CA	C	58.630	0.000	1
1324	332	Ser	CB	C	63.234	0.000	1

```
1325    332    Ser    H     H    8.317   0.000   1
1326    332    Ser    N     N    118.136 0.000   1
1327    333    Asn    C     C    173.764 0.000   1
1328    333    Asn    CA    C    52.824  0.000   1
1329    333    Asn    CB    C    38.717  0.000   1
1330    333    Asn    H     H    8.242   0.000   1
1331    333    Asn    N     N    121.197 0.000   1
1332    334    Ala    C     C    182.491 0.000   1
1333    334    Ala    CA    C    53.472  0.000   1
1334    334    Ala    CB    C    19.314  0.000   1
1335    334    Ala    H     H    7.745   0.000   1
1336    334    Ala    N     N    129.850 0.000   1

stop_
```

Abe, Y., S. Matsumoto, et al. (2000). "Cloning and expression of a novel MAPKK-like protein kinase, lymphokine-activated killer T-cell-originated protein kinase, specifically expressed in the testis and activated lymphoid cells." J Biol Chem 275(28): 21525-31.

Abe, Y., T. Takeuchi, et al. (2007). "A mitotic kinase TOPK enhances Cdk1/cyclin B1-dependent phosphorylation of PRC1 and promotes cytokinesis." J Mol Biol 370(2): 231-45.

Abo, A., J. Qu, et al. (1998). "PAK4, a novel effector for Cdc42Hs, is implicated in the reorganization of the actin cytoskeleton and in the formation of filopodia." EMBO J 17(22): 6527-40.

Adams, J. A. (2001). "Kinetic and catalytic mechanisms of protein kinases." Chem Rev 101(8): 2271-90.

Adams, J. A. (2003). "Activation loop phosphorylation and catalysis in protein kinases: is there functional evidence for the autoinhibitor model?" Biochemistry 42(3): 601-7.

Akasaka, H., T. Akasaka, et al. (2000). "Molecular anatomy of BCL6 translocations revealed by long-distance polymerase chain reaction-based assays." Cancer Res 60(9): 2335-41.

Amson, R., F. Sigaux, et al. (1989). "The human protooncogene product p33pim is expressed during fetal hematopoiesis and in diverse leukemias." Proc Natl Acad Sci U S A 86(22): 8857-61.

Anderson, K. A., R. L. Means, et al. (1998). "Components of a calmodulin-dependent protein kinase cascade. Molecular cloning, functional characterization and cellular localization of Ca2+/calmodulin-dependent protein kinase kinase beta." J Biol Chem 273(48): 31880-9.

Bachmann, M. and T. Moroy (2005). "The serine/threonine kinase Pim-1." Int J Biochem Cell Biol 37(4): 726-30.

Baselga, J. (2006). "Targeting tyrosine kinases in cancer: the second wave." Science 312(5777): 1175-8.

Beitz, E. (2000). "TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon." Bioinformatics 16(2): 135-9.

Bergamaschi, A., Y. H. Kim, et al. (2008). "CAMK1D amplification implicated in epithelial-mesenchymal transition in basal-like breast cancer." Mol Oncol 2(4): 327-39.

Berridge, M. J., M. D. Bootman, et al. (1998). "Calcium--a life and death signal." Nature 395(6703): 645-8.

Brunger, A. T. (2007). "Version 1.2 of the Crystallography and NMR system." Nat Protoc 2(11): 2728-33.

Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr 54(Pt 5): 905-21.

Bullock, A. N., J. Debreczeni, et al. (2005). "Structure and substrate specificity of the Pim-1 kinase." J Biol Chem 280(50): 41675-82.

Burnett, G. and E. P. Kennedy (1954). "The enzymatic phosphorylation of proteins." J Biol Chem 211(2): 969-80.

Carmel, G., B. Leichus, et al. (1994). "Expression, purification, crystallization, and preliminary x-ray analysis of casein kinase-1 from Schizosaccharomyces pombe." J Biol Chem 269(10): 7304-9.

Chin, D. and A. R. Means (2000). "Calmodulin: a prototypical calcium sensor." Trends Cell Biol 10(8): 322-8.

Ching, Y. P., V. Y. Leong, et al. (2003). "Identification of an autoinhibitory domain of p21-activated protein kinase 5." J Biol Chem 278(36): 33621-4.

Clapham, D. E. (2007). "Calcium signaling." Cell 131(6): 1047-58.

Clapperton, J. A., S. R. Martin, et al. (2002). "Structure of the complex of calmodulin with the target sequence of calmodulin-dependent protein kinase I: studies of the kinase activation mechanism." Biochemistry 41(50): 14669-79.

Cohen, P. (2001). "The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture." Eur J Biochem 268(19): 5001-10.

Cohen, P. (2002). "The origins of protein phosphorylation." Nat Cell Biol 4(5): E127-30.

Cohen, P. (2002). "Protein kinases--the major drug targets of the twenty-first century?" Nat Rev Drug Discov 1(4): 309-15.

Colomer, J. and A. R. Means (2007). "Physiological roles of the Ca2+/CaM-dependent protein kinase cascade in health and disease." Subcell Biochem 45: 169-214.

Cote, S., C. Simard, et al. (2002). "Regulation of growth-related genes by interleukin-6 in murine myeloma cells." Cytokine 20(3): 113-20.

Cotteret, S. and J. Chernoff (2006). "Nucleocytoplasmic shuttling of Pak5 regulates its antiapoptotic properties." Mol Cell Biol 26(8): 3215-30.

Cotteret, S., Z. M. Jaffer, et al. (2003). "p21-Activated kinase 5 (Pak5) localizes to mitochondria and inhibits apoptosis by phosphorylating BAD." Mol Cell Biol 23(16): 5526-39.

Cuypers, H. T., G. Selten, et al. (1984). "Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region." Cell 37(1): 141-50.

Dan, C., N. Nath, et al. (2002). "PAK5, a new brain-specific kinase, promotes neurite outgrowth in N1E-115 cells." Mol Cell Biol 22(2): 567-77.

de Diego, I., J. Kuper, et al. (2010). "Molecular basis of the death-associated protein kinase-calcium/calmodulin regulator complex." Sci Signal 3(106): ra6.

Delaglio, F., S. Grzesiek, et al. (1995). "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." J Biomol NMR 6(3): 277-93.

Dhanasekaran, S. M., T. R. Barrette, et al. (2001). "Delineation of prognostic biomarkers in prostate cancer." Nature 412(6849): 822-6.

Dolmetsch, R. E., K. Xu, et al. (1998). "Calcium oscillations increase the efficiency and specificity of gene expression." Nature 392(6679): 933-6.

Domen, J., N. M. van der Lugt, et al. (1993). "Analysis of Pim-1 function in mutant mice." Leukemia 7 Suppl 2: S108-12.

Elias, L., A. P. Li, et al. (1981). "Cyclic adenosine 3':5'-monophosphate-dependent and -independent protein kinase in acute myeloblastic leukemia." Cancer Res 41(6): 2182-8.

Enslen, H., P. Sun, et al. (1994). "Characterization of Ca2+/calmodulin-dependent protein kinase IV. Role in transcriptional regulation." J Biol Chem 269(22): 15520-7.

Eswaran, J., W. H. Lee, et al. (2007). "Crystal Structures of the p21-activated kinases PAK4, PAK5, and PAK6 reveal catalytic domain plasticity of active group II PAKs." Structure 15(2): 201-13.

Fabbro, D., S. Ruetz, et al. (2002). "Protein kinases as targets for anticancer agents: from inhibitors to useful drugs." Pharmacol Ther 93(2-3): 79-98.

Fedorov, O., B. Marsden, et al. (2007). "A systematic interaction map of validated kinase inhibitors with Ser/Thr kinases." Proc Natl Acad Sci U S A 104(51): 20523-8.

Fernandez, C. and G. Wider (2003). "TROSY in NMR studies of the structure and function of large biological macromolecules." Curr Opin Struct Biol 13(5): 570-80.

Fischer, E. H., D. J. Graves, et al. (1959). "Structure of the site phosphorylated in the phosphorylase b to a reaction." J Biol Chem 234(7): 1698-704.

Fischer, E. H. and E. G. Krebs (1955). "Conversion of phosphorylase b to phosphorylase a in muscle extracts." J Biol Chem 216(1): 121-32.

Fish, K. J., A. Cegielska, et al. (1995). "Isolation and characterization of human casein kinase I epsilon (CKI), a novel member of the CKI gene family." J Biol Chem 270(25): 14875-83.

Flotow, H., P. R. Graves, et al. (1990). "Phosphate groups as substrate determinants for casein kinase I action." J Biol Chem 265(24): 14264-9.

Foster, M. P., C. A. McElroy, et al. (2007). "Solution NMR of large molecules and assemblies." Biochemistry 46(2): 331-40.

Fujibuchi, T., Y. Abe, et al. (2005). "Expression and phosphorylation of TOPK during spermatogenesis." Dev Growth Differ 47(9): 637-44.

Fukukawa, C., K. Ueda, et al. (2010). "Critical roles of LGN/GPSM2 phosphorylation by PBK/TOPK in cell division of breast cancer cells." Genes Chromosomes Cancer 49(10): 861-72.

Gardner, K. H. and L. E. Kay (1998). "The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins." Annu Rev Biophys Biomol Struct 27: 357-406.

Gaudet, S., D. Branton, et al. (2000). "Characterization of PDZ-binding kinase, a mitotic kinase." Proc Natl Acad Sci U S A 97(10): 5167-72.

GCUA "Graphical Codon Usage Analyser."

Gelev, V., H. Aktas, et al. (2006). "Mapping of the auto-inhibitory interactions of protein kinase R by nuclear magnetic resonance." J Mol Biol 364(3): 352-63.

Goldberg, J., A. C. Nairn, et al. (1996). "Structural basis for the autoinhibition of calcium/calmodulin-dependent protein kinase I." Cell 84(6): 875-87.

Graves, P. R., D. W. Haas, et al. (1993). "Molecular cloning, expression, and characterization of a 49-kilodalton casein kinase I isoform from rat testis." J Biol Chem 268(9): 6394-401.

Graves, P. R. and P. J. Roach (1995). "Role of COOH-terminal phosphorylation in the regulation of casein kinase I delta." J Biol Chem 270(37): 21689-94.

Gunther, U. L., C. Ludwig, et al. (2000). "NMRLAB-Advanced NMR data processing in matlab." J Magn Reson 145(2): 201-8.

Hanks, S. K. and T. Hunter (1995). "Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification." FASEB J 9(8): 576-96.

Hanks, S. K., A. M. Quinn, et al. (1988). "The protein kinase family: conserved features and deduced phylogeny of the catalytic domains." Science 241(4861): 42-52.

Haribabu, B., S. S. Hook, et al. (1995). "Human calcium-calmodulin dependent protein kinase I: cDNA cloning, domain structure and activation by phosphorylation at threonine-177 by calcium-calmodulin dependent protein kinase I kinase." EMBO J 14(15): 3679-86.

Herrero-Martin, D., D. Osuna, et al. (2009). "Stable interference of EWS-FLI1 in an Ewing sarcoma cell line impairs IGF-1/IGF-1R signalling and reveals TOPK as a new target." Br J Cancer 101(1): 80-90.

Ho, Y., S. Mason, et al. (1997). "Role of the casein kinase I isoform, Hrr25, and the cell cycle-regulatory transcription factor, SBF, in the transcriptional response to DNA damage in Saccharomyces cerevisiae." Proc Natl Acad Sci U S A 94(2): 581-6.

Hochuli, M., T. Szyperski, et al. (2000). "Deuterium isotope effects on the central carbon metabolism of Escherichia coli cells grown on a D2O-containing minimal medium." J Biomol NMR 17(1): 33-42.

Hook, S. S. and A. R. Means (2001). "Ca(2+)/CaM-dependent kinases: from activation to function." Annu Rev Pharmacol Toxicol 41: 471-505.

Howe, C. J., M. M. LaHair, et al. (2002). "Participation of the calcium/calmodulin-dependent kinases in hydrogen peroxide-induced Ikappa B phosphorylation in human T lymphocytes." J Biol Chem 277(34): 30469-76.

Hu, F., R. B. Gartenhaus, et al. (2010). "PBK/TOPK interacts with the DBD domain of tumor suppressor p53 and modulates expression of transcriptional targets including p21." Oncogene 29(40): 5464-74.

Hu, S. H., M. W. Parker, et al. (1994). "Insights into autoregulation from the crystal structure of twitchin kinase." Nature 369(6481): 581-4.

Huse, M. and J. Kuriyan (2002). "The conformational plasticity of protein kinases." Cell 109(3): 275-82.

Invitrogen (2007). "TOPK/PBK certificate of analysis."

Ishikawa, Y., H. Tokumitsu, et al. (2003). "Identification and characterization of novel components of a Ca2+/calmodulin-dependent protein kinase cascade in HeLa cells." FEBS Lett 550(1-3): 57-63.

Jacobs, M. D., J. Black, et al. (2005). "Pim-1 ligand-bound structures reveal the mechanism of serine/threonine kinase inhibition by LY294002." J Biol Chem 280(14): 13728-34.

Jaffer, Z. M. and J. Chernoff (2002). "p21-activated kinases: three more join the Pak." Int J Biochem Cell Biol 34(7): 713-7.

Jensen, K. F., C. A. Ohmstede, et al. (1991). "Nuclear and axonal localization of Ca2+/calmodulin-dependent protein kinase type Gr in rat cerebellar cortex." Proc Natl Acad Sci U S A 88(7): 2850-3.

Kahl, C. R. and A. R. Means (2003). "Regulation of cell cycle progression by calcium/calmodulin-dependent pathways." Endocr Rev 24(6): 719-36.

Kemp, B. E., R. B. Pearson, et al. (1987). "The calmodulin binding domain of chicken smooth muscle myosin light chain kinase contains a pseudosubstrate sequence." J Biol Chem 262(6): 2542-8.

kinbase.

Kitani, T., S. Okuno, et al. (1997). "Molecular cloning of Ca2+/calmodulin-dependent protein kinase kinase beta." J Biochem 122(1): 243-50.

Knighton, D. R., J. H. Zheng, et al. (1991). "Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science 253(5018): 407-14.

Knighton, D. R., J. H. Zheng, et al. (1991). "Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase." Science 253(5018): 414-20.

Knippschild, U., A. Gocht, et al. (2005). "The casein kinase 1 family: participation in multiple cellular processes in eukaryotes." Cell Signal 17(6): 675-89.

Kornev, A. P., S. S. Taylor, et al. (2008). "A helix scaffold for the assembly of active protein kinases." Proc Natl Acad Sci U S A 105(38): 14377-82.

Krebs, E. G. and E. H. Fischer (1956). "The phosphorylase b to a converting enzyme of rabbit skeletal muscle." Biochim Biophys Acta 20(1): 150-7.

Kwon, H. R., K. W. Lee, et al. (2010). "Requirement of T-lymphokine-activated killer cell-originated protein kinase for TRAIL resistance of human HeLa cervical cancer cells." Biochem Biophys Res Commun 391(1): 830-4.

Ladbury, J. E. (2007). "Measurement of the formation of complexes in tyrosine kinase-mediated signal transduction." Acta Crystallogr D Biol Crystallogr 63(Pt 1): 26-31.

Ladbury, J. E. and B. Z. Chowdhry (1996). "Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions." Chem Biol 3(10): 791-801.

Langan, T. A. (1969). "Phosphorylation of liver histone following the administration of glucagon and insulin." Proc Natl Acad Sci U S A 64(4): 1276-83.

Langer, T., M. Vogtherr, et al. (2004). "NMR backbone assignment of a protein kinase catalytic domain by a combination of several approaches: application to the catalytic subunit of cAMP-dependent protein kinase." Chembiochem 5(11): 1508-16.

Leavitt, S. and E. Freire (2001). "Direct measurement of protein binding energetics by isothermal titration calorimetry." Curr Opin Struct Biol 11(5): 560-6.

Linn, T. C., F. H. Pettit, et al. (1969). "Alpha-keto acid dehydrogenase complexes. X. Regulation of the activity of the pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation." Proc Natl Acad Sci U S A 62(1): 234-41.

Longenecker, K. L., P. J. Roach, et al. (1998). "Crystallographic studies of casein kinase I delta toward a structural understanding of auto-inhibition." Acta Crystallogr D Biol Crystallogr 54(Pt 3): 473-5.

Manning, G., D. B. Whyte, et al. (2002). "The protein kinase complement of the human genome." Science 298(5600): 1912-34.

Marsden, B. D. and S. Knapp (2008). "Doing more than just the structure-structural genomics in kinase drug discovery." Curr Opin Chem Biol 12(1): 40-5.

Masterson, L. R., L. Shi, et al. (2009). "Backbone NMR resonance assignment of the catalytic subunit of cAMP-dependent protein kinase A in complex with AMP-PNP." Biomol NMR Assign 3(1): 115-7.

Matsumoto, S., Y. Abe, et al. (2004). "Characterization of a MAPKK-like protein kinase TOPK." Biochem Biophys Res Commun 325(3): 997-1004.

Matsushita, M. and A. C. Nairn (1998). "Characterization of the mechanism of regulation of Ca2+/calmodulin-dependent protein kinase I by calmodulin and by Ca2+/calmodulin-dependent protein kinase kinase." J Biol Chem 273(34): 21473-81.

Matsushita, M. and A. C. Nairn (1999). "Inhibition of the Ca2+/calmodulin-dependent protein kinase I cascade by cAMP-dependent protein kinase." J Biol Chem 274(15): 10086-93.

Matthews, R. P., C. R. Guthrie, et al. (1994). "Calcium/calmodulin-dependent protein kinase types II and IV differentially regulate CREB-dependent gene expression." Mol Cell Biol 14(9): 6107-16.

Means, A. R. (2000). "Regulatory cascades involving calmodulin-dependent protein kinases." Mol Endocrinol 14(1): 4-13.

Means, A. R. (2008). "The Year in Basic Science: calmodulin kinase cascades." Mol Endocrinol 22(12): 2759-65.

Meggio, F., J. W. Perich, et al. (1992). "The comparative efficiencies of the Ser(P)-, Thr(P)- and Tyr(P)-residues as specificity determinants for casein kinase-1." Biochem Biophys Res Commun 182(3): 1460-5.

Meggio, F., J. W. Perich, et al. (1991). "A synthetic beta-casein phosphopeptide and analogues as model substrates for casein kinase-1, a ubiquitous, phosphate directed protein kinase." FEBS Lett 283(2): 303-6.

Melnikova, I. and J. Golden (2004). "Targeting protein kinases." Nat Rev Drug Discov 3(12): 993-4.

Mikkers, H., M. Nawijn, et al. (2004). "Mice deficient for all PIM kinases display reduced body size and impaired responses to hematopoietic growth factors." Mol Cell Biol 24(13): 6104-15.

Mukherjee, K., M. Sharma, et al. (2008). "CASK Functions as a Mg2+-independent neurexin kinase." Cell 133(2): 328-39.

Nakamura, Y., S. Okuno, et al. (1996). "Distribution of Ca2+/calmodulin-dependent protein kinase kinase alpha in the rat central nervous system: an immunohistochemical study." Neurosci Lett 204(1-2): 61-4.

Nandi, A., M. Tidwell, et al. (2004). "Protein expression of PDZ-binding kinase is up-regulated in hematologic malignancies and strongly down-regulated during terminal differentiation of HL-60 leukemic cells." Blood Cells Mol Dis 32(1): 240-5.

Oh, S. M., F. Zhu, et al. (2007). "T-lymphokine-activated killer cell-originated protein kinase functions as a positive regulator of c-Jun-NH2-kinase 1 signaling and H-Ras-induced cell transformation." Cancer Res 67(11): 5186-94.

P. V. Konarev, V. V. V., A. V. Sokolova, M. H. J. Koch and D. I. Svergun (2003). "PRIMUS: a Windows PC-based system for small-angle scattering data analysis." Journal of Applied Crystallography 36: 1277-1282.

Padma, R. and L. Nagarajan (1991). "The human PIM-1 gene product is a protein serine kinase." Cancer Res 51(9): 2486-9.

Pandey, A., I. Dan, et al. (2002). "Cloning and characterization of PAK5, a novel member of mammalian p21-activated kinase-II subfamily that is predominantly expressed in brain." Oncogene 21(24): 3939-48.

Park, J. H., M. L. Lin, et al. (2006). "PDZ-binding kinase/T-LAK cell-originated protein kinase, a putative cancer/testis antigen with an oncogenic activity in breast cancer." Cancer Res 66(18): 9186-95.

Park, J. H., T. Nishidate, et al. (2009). "Critical roles of T-LAK cell-originated protein kinase in cytokinesis." Cancer Sci 101(2): 403-11.

Pasqualucci, L., P. Neumeister, et al. (2001). "Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas." Nature 412(6844): 341-6.

Paul, M. K. and A. K. Mukhopadhyay (2004). "Tyrosine kinase - Role and significance in Cancer." Int J Med Sci 1(2): 101-115.

Pearson, R. B., R. E. Wettenhall, et al. (1988). "Autoregulation of enzymes by pseudosubstrate prototopes: myosin light chain kinase." Science 241(4868): 970-3.

Peersen, O. B., T. S. Madsen, et al. (1997). "Intermolecular tuning of calmodulin by target peptides and proteins: differential effects on Ca2+ binding and implications for kinase activation." Protein Sci 6(4): 794-807.

Pervushin, K. (2000). "Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology." Q Rev Biophys 33(2): 161-97.

Pervushin, K., R. Riek, et al. (1997). "Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution." Proc Natl Acad Sci U S A 94(23): 12366-71.

Picciotto, M. R., J. A. Cohn, et al. (1992). "Phosphorylation of the cystic fibrosis transmembrane conductance regulator." J Biol Chem 267(18): 12742-52.

Picciotto, M. R., A. J. Czernik, et al. (1993). "Calcium/calmodulin-dependent protein kinase I. cDNA cloning and identification of autophosphorylation site." J Biol Chem 268(35): 26512-21.

Pierce, M. M., C. S. Raman, et al. (1999). "Isothermal titration calorimetry of protein-protein interactions." Methods 19(2): 213-21.

Qian, K. C., L. Wang, et al. (2005). "Structural basis of constitutive activity and a unique nucleotide binding mode of human Pim-1 kinase." J Biol Chem 280(7): 6130-7.

Qin, H., B. Raught, et al. (2003). "Phosphorylation screening identifies translational initiation factor 4GII as an intracellular target of Ca(2+)/calmodulin-dependent protein kinase I." J Biol Chem 278(49): 48570-9.

Riek, R., K. Pervushin, et al. (2000). "TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution." Trends Biochem Sci 25(10): 462-8.

Rina, S., A. A. Jusuf, et al. (2001). "Distribution of Ca(2+)/calmodulin-dependent protein kinase I beta 2 in the central nervous system of the rat." Brain Res 911(1): 1-11.

Ringe, D. and G. A. Petsko (1986). "Study of protein dynamics by X-ray diffraction." Methods Enzymol 131: 389-433.

Rodriguez-Mora, O. G., M. M. Lahair, et al. (2006). "Inhibition of the CaM-kinases augments cell death in response to oxygen radicals and oxygen radical inducing cancer therapies in MCF-7 human breast cancer cells." Cancer Biol Ther 5(8): 1022-30.

Rodriguez-Mora, O. G., M. M. LaHair, et al. (2005). "Calcium/calmodulin-dependent kinase I and calcium/calmodulin-dependent kinase kinase participate in the control of cell cycle progression in MCF-7 human breast cancer cells." Cancer Res 65(12): 5408-16.

Sakagami, H., A. Kamata, et al. (2005). "Prominent expression and activity-dependent nuclear translocation of Ca2+/calmodulin-dependent protein kinase Idelta in hippocampal neurons." Eur J Neurosci 22(11): 2697-707.

Sakagami, H., S. Saito, et al. (1998). "Localization of the mRNAs for two isoforms of Ca2+/calmodulin-dependent protein kinase kinases in the adult rat brain." Brain Res Mol Brain Res 54(2): 311-5.

Sattler, M. and S. W. Fesik (1996). "Use of deuterium labeling in NMR: overcoming a sizeable problem." Structure 4(11): 1245-9.

Schrodinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3r1.

Shen, Y., F. Delaglio, et al. (2009). "TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts." J Biomol NMR 44(4): 213-23.

Simons-Evelyn, M., K. Bailey-Dell, et al. (2001). "PBK/TOPK is a novel mitotic kinase which is upregulated in Burkitt's lymphoma and other highly proliferative malignant cells." Blood Cells Mol Dis 27(5): 825-9.

Soderling, T. R. (1999). "The Ca-calmodulin-dependent protein kinase cascade." Trends Biochem Sci 24(6): 232-6.

Soderling, T. R. and J. T. Stull (2001). "Structure and regulation of calcium/calmodulin-dependent protein kinases." Chem Rev 101(8): 2341-52.

Subramaniam, S. (1998). "The Biology Workbench--a seamless database and analysis environment for the biologist." Proteins 32(1): 1-2.

Sun, P., L. Lou, et al. (1996). "Regulation of activating transcription factor-1 and the cAMP response element-binding protein by Ca2+/calmodulin-dependent protein kinases type I, II, and IV." J Biol Chem 271(6): 3066-73.

Svergun, D. F. a. D. I. (2009). "DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering." Journal of Applied Crystallography 42: 342-346.

Svergun, D. I. (1992). "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." Journal of Applied Crystallography 25: 495-503.

Svergun, M. B. K. a. D. I. (2001). "Automated matching of high- and low-resolution structural models." Journal of Applied Crystallography 34: 33-41.

Svergun, V. V. V. a. D. I. (2003). "Uniqueness of ab initio shape determination in small-angle scattering." Journal of Applied Crystallography 36: 860-864.

Swindells, M. B. and M. Ikura (1996). "Pre-formation of the semi-open conformation by the apo-calmodulin C-terminal domain and implications binding IQ-motifs." Nat Struct Biol 3(6): 501-4.

Swulius, M. T. and M. N. Waxham (2008). "Ca(2+)/calmodulin-dependent protein kinases." Cell Mol Life Sci 65(17): 2637-57.

Takemoto-Kimura, S., N. Ageta-Ishihara, et al. (2007). "Regulation of dendritogenesis via a lipid-raft-associated Ca2+/calmodulin-dependent protein kinase CLICK-III/CaMKIgamma." Neuron 54(5): 755-70.

Takemoto-Kimura, S., H. Terai, et al. (2003). "Molecular cloning and characterization of CLICK-III/CaMKIgamma, a novel membrane-anchored neuronal Ca2+/calmodulin-dependent protein kinase (CaMK)." J Biol Chem 278(20): 18597-605.

Taylor, S. S. and A. P. Kornev "Protein kinases: evolution of dynamic regulatory proteins." Trends Biochem Sci 36(2): 65-77.

Taylor, S. S. and A. P. Kornev (2010). "Protein kinases: evolution of dynamic regulatory proteins." Trends Biochem Sci 36(2): 65-77.

Thompson, J. D., D. G. Higgins, et al. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res 22(22): 4673-80.

Tokumitsu, H., H. Enslen, et al. (1995). "Characterization of a Ca2+/calmodulin-dependent protein kinase cascade. Molecular cloning and expression of calcium/calmodulin-dependent protein kinase kinase." J Biol Chem 270(33): 19320-4.

Tokumitsu, H., M. Iwabu, et al. (2001). "Differential regulatory mechanism of Ca2+/calmodulin-dependent protein kinase kinase isoforms." Biochemistry 40(46): 13925-32.

Tokumitsu, H., N. Takahashi, et al. (1999). "Substrate recognition by Ca2+/Calmodulin-dependent protein kinase kinase. Role of the arg-pro-rich insert domain." J Biol Chem 274(22): 15803-10.

Ueda, T., H. Sakagami, et al. (1999). "Distribution and intracellular localization of a mouse homologue of Ca2+/calmodulin-dependent protein kinase Ibeta2 in the nervous system." J Neurochem 73(5): 2119-29.

uniprot.

Vajpai, N., A. Strauss, et al. (2008). "Solution conformations and dynamics of ABL kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib." J Biol Chem 283(26): 18292-302.

Vajpai, N., A. Strauss, et al. (2008). "Backbone NMR resonance assignment of the Abelson kinase domain in complex with imatinib." Biomol NMR Assign 2(1): 41-2.

van Lohuizen, M., S. Verbeek, et al. (1989). "Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors." Cell 56(4): 673-82.

Velazquez-Campoy, A., H. Ohtaka, et al. (2004). "Isothermal titration calorimetry." Curr Protoc Cell Biol Chapter 17: Unit 17 8.

Verbeek, S., M. van Lohuizen, et al. (1991). "Mice bearing the E mu-myc and E mu-pim-1 transgenes develop pre-B-cell leukemia prenatally." Mol Cell Biol 11(2): 1176-9.

Verploegen, S., J. W. Lammers, et al. (2000). "Identification and characterization of CKLiK, a novel granulocyte Ca(++)/calmodulin-dependent kinase." Blood 96(9): 3215-23.

Verploegen, S., L. Ulfman, et al. (2005). "Characterization of the role of CaMKI-like kinase (CKLiK) in human granulocyte function." Blood 106(3): 1076-83.

Vlahovic, G. and J. Crawford (2003). "Activation of tyrosine kinases in cancer." Oncologist 8(6): 531-8.

Vogtherr, M., K. Saxena, et al. (2005). "NMR backbone assignment of the mitogen-activated protein (MAP) kinase p38." J Biomol NMR 32(2): 175.

Vogtherr, M., K. Saxena, et al. (2006). "NMR characterization of kinase p38 dynamics in free and ligand-bound forms." Angew Chem Int Ed Engl 45(6): 993-7.

Vranken, W. F., W. Boucher, et al. (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." Proteins 59(4): 687-96.

Wang, Z., N. Bhattacharya, et al. (2002). "Phosphorylation of the cell cycle inhibitor p21Cip1/WAF1 by Pim-1 kinase." Biochim Biophys Acta 1593(1): 45-55.

Wang, Z., N. Bhattacharya, et al. (2001). "Pim-1: a serine/threonine kinase with a role in cell survival, proliferation, differentiation and tumorigenesis." J Vet Sci 2(3): 167-79.

Wayman, G. A., H. Tokumitsu, et al. (1997). "Inhibitory cross-talk by cAMP kinase on the calmodulin-dependent protein kinase cascade." J Biol Chem 272(26): 16073-6.

Wells, C. M. and G. E. Jones (2010). "The emerging importance of group II PAKs." Biochem J 425(3): 465-73.

Willis, B. T. M. (1975). Thermal vibrations in crystallography, Cambridge University press.

Wiseman, T., S. Williston, et al. (1989). "Rapid measurement of binding constants and heats of binding using a new titration calorimeter." Anal Biochem 179(1): 131-7.

Wu, X., H. S. Carr, et al. (2008). "p21 activated kinase 5 activates Raf-1 and targets it to mitochondria." J Cell Biochem 105(1): 167-75.

Wuthrich, K. (2003). "NMR studies of structure and function of biological macromolecules (Nobel Lecture)." J Biomol NMR 27(1): 13-39.

Xie, Y., K. Xu, et al. (2006). "The 44 kDa Pim-1 kinase directly interacts with tyrosine kinase Etk/BMX and protects human prostate cancer cells from apoptosis induced by chemotherapeutic drugs." Oncogene 25(1): 70-8.

Yokokura, H., M. R. Picciotto, et al. (1995). "The regulatory region of calcium/calmodulin-dependent protein kinase I contains closely associated autoinhibitory and calmodulin-binding domains." J Biol Chem 270(40): 23851-9.

Zhai, L., P. R. Graves, et al. (1995). "Casein kinase I gamma subfamily. Molecular cloning, expression, and characterization of three mammalian isoforms and complementation of defects in the Saccharomyces cerevisiae YCK genes." J Biol Chem 270(21): 12717-24.

Zhu, F., T. A. Zykova, et al. (2007). "Bidirectional signals transduced by TOPK-ERK interaction increase tumorigenesis of HCT116 colorectal cancer cells." Gastroenterology 133(1): 219-31.

Zhu, N., L. M. Ramirez, et al. (2002). "CD40 signaling in B cells regulates the expression of the Pim-1 kinase via the NF-kappa B pathway." J Immunol 168(2): 744-54.

Zykova, T. A., F. Zhu, et al. (2006). "Lymphokine-activated killer T-cell-originated protein kinase phosphorylation of histone H2AX prevents arsenite-induced apoptosis in RPMI7951 melanoma cells." Clin Cancer Res 12(23): 6884-93.

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository