STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF SIGNALLING PROTEIN COMPLEXES

NG CHERLYN
BSc (Hons), NUS

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES
AT THE NATIONAL UNIVERSITY OF SINGAPORE

2009

For Dad and Mum

Acknowledgements

My heartfelt gratitude first extends to A/P J. Sivaraman, a boss who wrote himself into being a teacher, counsellor and role-model. Your expertise and passion for research is truly admirable, and your dedication to us, students, has inspired me to follow suit should I run a lab one day.

Also to A/P Sheu Fwu Shan, a mentor of eight years who introduced me to research, first-handedly taught the techniques of protein expression, and saw me through the undergraduate years till today, thank you.

I thank A/P Graeme Guy, Dr. Rebecca A. Jackson, Dr. Jan P. Buschorf and all at IMCB for the clones, peptides and immunoprecipitation assays for the Cbl-TKB study; more importantly, your guidance, patience and opportunity. I hope to join in your comradeship.

Dr Anand Saxena from Brookhaven National Laboratories National Synchrotron Light Source, for assistance in data collection and and Mrs Mala Saxena too. I have tasted your hospitality (literally) and experienced a sincere friendship that trenscends age, language, race and religion.

Finally to friends brought together by the constraints of lab space:

Sunita - a wonderful sister and confidente, Jeremy and Tzer Fong, Dr. Zhou Xingding and Dr. Li Mo, I think we were gelled by joys and frustrations that kept us going. A special tribute to Lissa who does a thankless job, I hope to learn your goodness.

I thank NUS for having given me the opportunity to pursue my Ph.D. with a research scholarship .

Table of Contents

Summary

The ERK/MAPK pathway is a ubiquitous serine/threonine kinase cascade that directs growth, differentiation and plasticity in various tissues. Transmembrane receptor proteins acts as a bridge between extracellular signals and the ERK/MAPK pathway, so regulation of these receptors is of crucial importance towards maintaining a healthy cell. Many regulatory mechanisms exist, all of which make use of protein interaction domains to bind to their substrates. This thesis draws examples from two regulatory proteins, c-Cbl and calmodulin.

c-Cbl is an E2 ubiquitin ligase and a major regulator of tyrosine kinases at the membrane. Through x-ray crystallography of five Cbl-TKB: phosphopeptide complexes, our work demonstrate the mechanism by which the Cbl-TKB domain binds to its substrates through a conserved, specificity determining intrapeptidyl hydrogen bond. The ability of Cbl to bind to its substrates in a reverse orientation given the TKB atypical binding motif found in the Met family of proteins was also uncovered. This finding implicates that there may be a group of yet undiscovered Cbl substrates.

Calmodulin is a calcium binding protein that modify its substrates' activity by conferring calcium sensitivity when it binds with its substrates. One of the domains responsible for this interaction is the IQ motif. All voltage gated sodium channels possess this motif but bind differentially to calmodulin. Through biophysical and computational analyses, we characterised the way calmodulin binds to two high affinity sodium channel isoforms Nav1.4 and Nav1.6. Together with mutation of two residues predicted to be involved in Nav1.4 association but not Nav1.6, we explained differences in binding.

List of Tables

List of Figures

List of Abbreviations

4H	Four-helix bundle
Å	angstrom (10^{-10} m)
AMPAR	α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate receptor
APS	Adapter with a Plekstrin homology and Src homology-2 domains
ATP	Adenosine triphosphate
apoCaM	Calmodulin without calcium ions bound
Ca^{2+}/CaM	Calcium loaded calmodulin
CaM	Calmodulin
Cbl	Casitas b-lineage lymphoma
Cbl-TKB	Tyrosine kinase binding domain of Cbl
CCD	Charged coupled device
cDNA	Complementary Deoxyribonucleic Acid
CNS	crystallography and NMR system
DTT	Dithiothreitol
E. coli	*Escherichia Coli*
EDTA	ethylenediamine tetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ERK	Extracellular signal regulated kinase
FGFR	Fibroblast growth factor receptor
ΔG	Gibbs free energy change
GDP	Guanosine diphosphate
ΔH	Enthalpy change

H-bond	Hydrogen bond
IP	Immunoprecipitation
IPTG	Isopropylthogalactoside
IQ motif	Protein interaction motif of consensus sequence [I,L,V]QXXXRXXXX[R,K]
ITC	isothermal titration calorimetry
K_a	Association constant
K_d	Dissociation constant
kDa	Kilodaltons
LB	Luria-Bertrani
LTP	Long term potentiation
MAPK	Mitogen activated protein kinase
MAPKK	Mitogen activated protein kinase kinase (also MAP2K, MKK or MEK)
MAPKKK	Mitogen activated protein kinase kinase kinase (also MAP3K, MEKK)
Met	Hepatocyte growth factor receptor
MPD	2-Methyl-2,4-pentanediol
N	Number of binding sites
Nav	Voltage gated sodium channel isoform
Nav1.4IQ	IQ motif of voltage gated sodium channel isoform 1.4
Nav1.6IQ	IQ motif of voltage gated sodium channel isoform 1.6
ΔNav1.6	Cytoplasmic domain of voltage gated sodium channel isoform 1.6
NCS	Non crystallographic restraints
Nm	Nanometer
NMDAR	N-methyl-D-aspartate receptor
NRTK	Non-receptor tyrosine kinase

PCR	polymerase chain reaction
PEG	polyethylene glycol
PTB	Phosphotyrosine binding
PTK	Protein tyrosine kinase
PTM	Post-translational modification
pTyr / pY	Phosphotyrosine
RING	Really interesting new gene
rmsd	Root mean square deviation
Ron	Macrophage stimulating 1 receptor
RTK	Receptor tyrosine kinase
ΔS	Entropy change
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SH2	Src homology 2
SH3	Src homology 3
SOS	Son of sevenless
Spry	Sprouty protein homologues
Syk	Spleen tyrosine kinase
TCL	Total cell lysate
TKB	Tyrosine kinase binding
VEGFR	Vascular endothelial growth factor receptor
VGCC	Voltage gated calcium channel
VGSC	Voltage gated sodium channel
WT	Wild-type
ZAP-70	ξ-chain associated protein kinase 70kDa

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
Leu (L)	Leucine
Lys (K)	Lysine
Met (M)	Methionine
Phe (F)	Phenylalanine
Pro (P)	Proline
Ser (S)	Serine
Thr (T)	Threonine
Trp (W)	Tryptophan
Tyr (Y)	Tyrosine
Val (V)	Valine

Publications

Ng C, Jackson RA, Buschdorf JP, Sun Q, Guy G, Sivaraman J (2008) Structural basis for a novel intrapeptidyl H-bond and reverse binding of c-Cbl-TKB domain substrates. EMBO J 27:804-816.

Chapter I

Introduction

1.1 General Introduction

In order to survive and adapt, every organism must be able to accurately respond to spatial and temporal cues. In unicellular organisms, the response to these cues can be rapidly transduced within the cell to achieve the desired change. However in multicellular organisms where billions of cells are organised into specific tissues and organs, transport of cues required to effect a change within tissues becomes more complex. Cells respond to these cues through numerous molecules such as proteins, peptides, lipids and inorganic molecules linked via intricate intracellular networks that efficiently transport these external signals across the plasma membrane into the cytosol and nucleus of the desired cell, where the effects occur. This process of transporting an external signal within the cell to induce a desired outcome is termed signal transduction or cell signalling. Certain conserved signalling pathways are reiteratively employed in biological systems to propagate, regulate, integrate and evoke tissue-specific responses. At the heart of these, protein-protein interactions form the basis through which signals are selectively modulated to determine a cell's fate.

1.2 Signal sensing

Cell to cell communication is essential to ensure that responses to external signals are appropriately coordinated. Cells signal to each other in several ways, depending on the distance between the signalling cell and the target cell. Not all molecules can be trafficked across the lipid bilayer: cells therefore relay messages via a cascade of intracellular events that is tightly regulated at key steps where pathways converge. All signal transduction pathways that are initiated by a non-permeable extracellular stimulus are received by a cell surface receptor, often

a transmembrane protein which causes a conformational change in the receptor upon sensing the stimulus. If the cells are touching, signalling may be through pores in the membrane, such as gap junctions or plasmodesmata, or due to a membrane bound ligand being identified by a receptor in the membrane of a neighbouring cell (Fig. 1.1b and c). If the cells are further apart, they may communicate via the release of molecules (Fig. 1.1a) in the form of cytokines, chemical agonists, growth factors and ions which are detected by the target cell via endocrine, paracrine, autocrine or synaptic signalling.

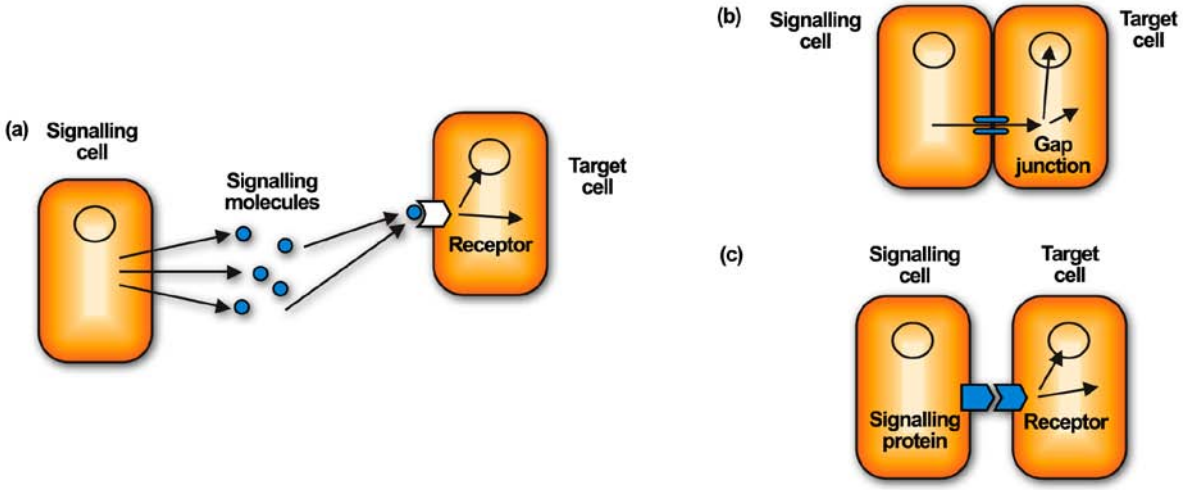


Fig. 1.1. Different ways cells signal to each other (Adapted from Hancock, 2005). (a) Communication to a distantly located cell through the release of signalling molecules into the extracellular space which travel to the target cell. (b) Crosstalk between neighbouring cells via gap junctions through which molecules can pass. (c) Signalling through a membrane bound ligand that is detected by a receptor on the neighbouring cell.

Synaptic: A fast and efficient method of signalling over the length of an axon via changes in the electrical potential across the membrane of a cell.

Endocrine: Cells release signalling molecules, such as hormones, that travel vast distances through the bloodstream to evoke a response in a different tissue.

Paracrine: Cells release signalling molecules that diffuse and are detected by adjacent cells. The signal is often quickly terminated by endocytosis, degradation or immobilisation.

Autocrine: Similar to paracrine signalling, where the signalling molecules act upon the cell that released them. This type of signalling is often found in differentiating cells as reinforcement towards a committed cell fate.

1.3 Types of receptors

Extracellular molecular signals are usually found at low concentrations (~10^{-8}mol/L). Detection of a signal is usually accomplished by specific receptors on the cell surface that have high binding affinity in the concentration range of the ligand. Binding of the ligand will then stimulate the required intracellular response via a signal transduction pathway. Despite the vast array of extracellular molecules that need to be detected by a single cell, receptors fall into four main classes: ion channel linked, G-protein linked, those containing intrinsic enzymatic activity (e.g., receptor tyrosine kinases - RTKs) and receptors without enzymatic activity but associate with cytosolic enzymes (e.g. non-receptor tyrosine kinase - NRTKs) (Fig. 1.2). Accordingly, they each initiate a series of distinct intracellular enzymatic activities in sequential order, to be known as signal transduction pathways that eventually result in changes to gene expression levels. In this thesis, receptors with tyrosine kinase activity and ion channel receptors will be described.

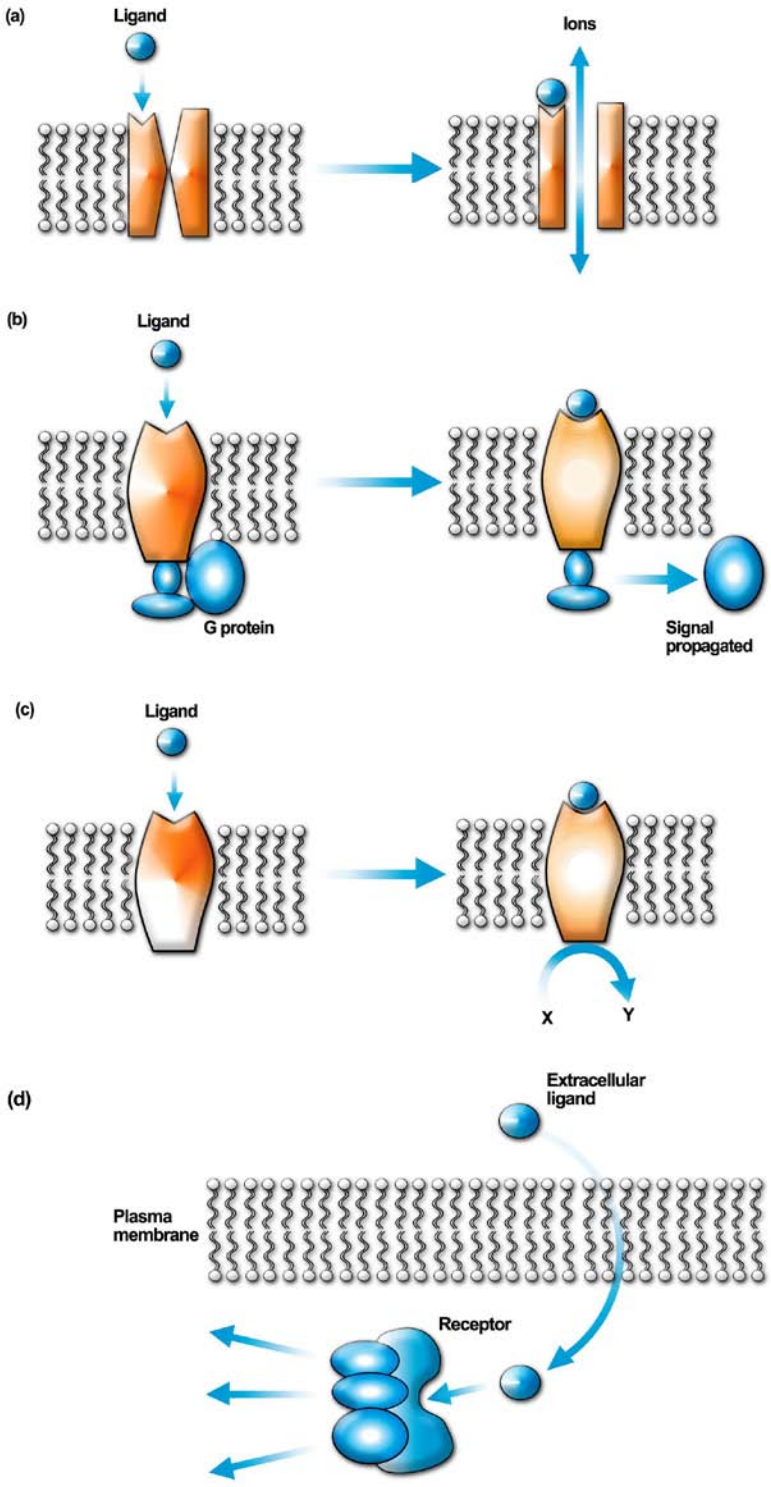


Fig. 1.2. Different receptor classes in an organism (Adapted from Hancock, 2005). (a) Ion channel receptors that open to allow the free passage of ions across the membrane upon stimulation. (b) Receptors linked to G-proteins. Dissociation of the G-protein from the receptor upon stimulation initiates a cascade of signalling events. (c) Receptors with enzymatic activities are activated through binding of their cognate ligands. (d) Intracellular receptors are not membrane bound but may be linked to receptors without any enzymatic activity. They affect signals similar those from receptors with innate enzymatic activity and are mainly involved in immune responses.

1.3.1 Example 1: Receptors with tyrosine kinase activity

Of the 32000 coding genes in the human genome, approximately 20% of these encode for proteins involved in signal transduction - more than 520 are protein kinases and 130 are phosphatases. Signal transduction pathways make extensive use of phosphorylated proteins (phosphoproteins) in which serine, threonine and tyrosine (pTyr) are the residues most commonly phosphorylated. Protein tyrosine kinases (PTKs) transfer the γ phosphate of ATP to the hydroxyl group of a tyrosine in a protein substrate. The 90 known PTK genes are distributed into two pools, 58 encoding receptor tyrosine kinases (RTKs) and 30 encoding non-receptor tyrosine kinases (NRTKs).

Receptor tyrosine kinases

RTKs are transmembrane glycoproteins activated by the binding of their cognate ligands that transduce extracellular signals to the cytoplasm to stimulate changes within the cell. This is accomplished by phosphorylating tyrosine residues on themselves (autophosphorylation) and on downstream adaptor proteins upon ligand binding. The RTK family includes the receptors for insulin and for many growth factors, such as epidermal growth factor (EGFR), fibroblast growth factor (FGFR), platelet-derived growth factor (PDGFR), vascular endothelial growth factor (VEGFR), nerve growth factor (NGFR) and hepatocyte growth factor (Met).

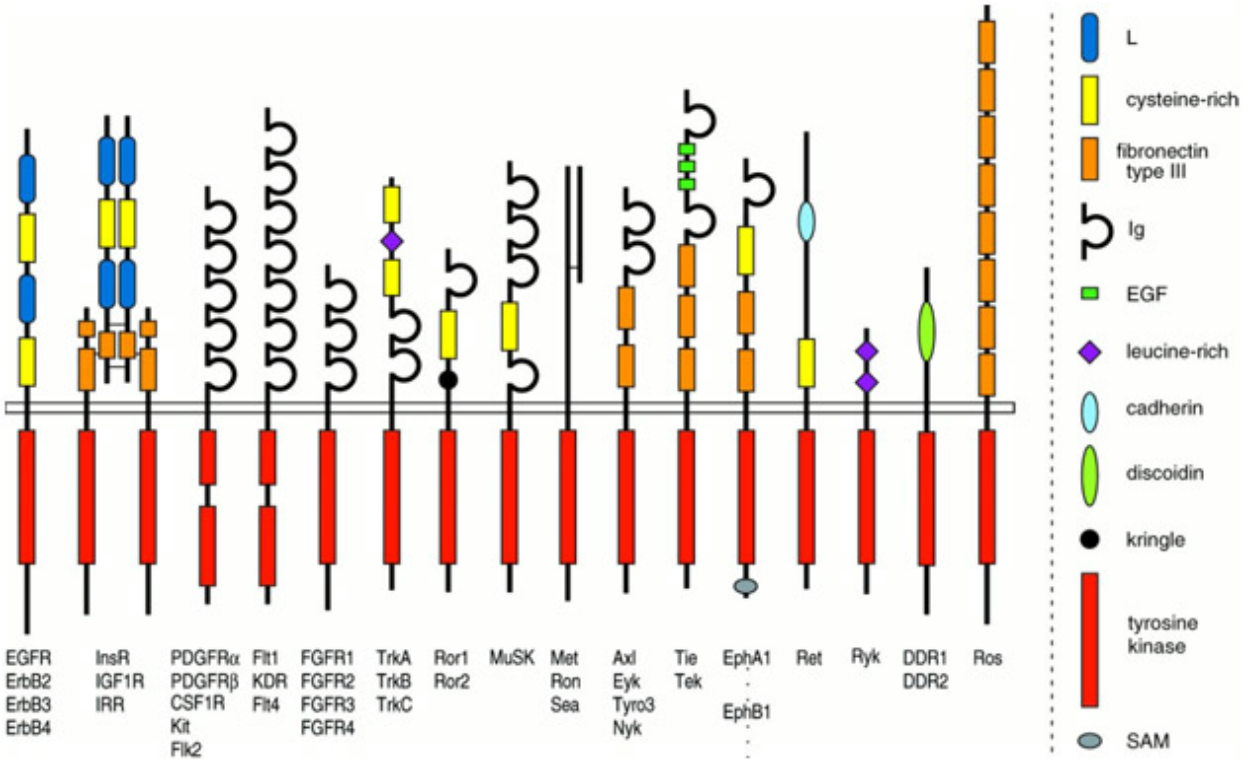


Fig. 1.3. Domain architecture of different RTK families (Hubbard and Till, 2000). RTK family members of the same structural organization are indicated under each module. The portion above the black horizontal lines represents extracellular domains involved in ligand binding, and tyrosine kinase domains are exclusively found in the intracellular region. The Met (Met, Ron, Sea) and insulin receptor (InsR, IGF1R, IRR) family members comprise of multiple subunits within a single receptor, hence the discontinuous extracellular domains.

With the exception of Met, the insulin receptor and their respective family members, RTKs exist as a single polypeptide chain that dimerize upon stimulation. The extracellular portion of RTKs contain a varied array of globular domains that are used for ligand binding. The domain organization in the cytoplasmic portion of RTKs consists of a juxtamembrane region after the transmembrane helix followed by the tyrosine kinase catalytic domain and a carboxy-terminal region (Fig. 1.3). All RTKs possess between one to three critical tyrosine residues in the kinase activation loop (Hanks *et al.*, 1991). With the exception of EGFR, phosphorylation of these residues seems to be essential for the activation of its catalytic activity. The

juxtamembrane and C-terminal regions are varied in length, and may also contain tyrosine residues that are autophosphorylated upon ligand binding.

Non-receptor tyrosine kinases

In addition to the RTKs, there exists a large family of non-receptor tyrosine kinases (NRTKs), including Src, Spleen tyrosine kinase (Syk) / ξ-chain associated protein kinase 70kDa (ZAP-70) and Abelson murine leukemia tyrosine kinase (Abl) among others. These proteins exist in the cytosol as soluble components, or may be receptor associated. The Src subfamily of proteins is the largest of all NRTKs, with nine members and implications in many human carcinomas. NRTKs are integral components of signal cascades from receptors with no intrinsic tyrosine kinase activity by binding to and triggering responses that are similar to RTKs (Gomperts *et al.*, 2003). The receptors they associate with typically mediate immune and inflammatory responses in leucocytes and lymphocytes. In addition, certain NRTKs like c-Src modify specific signals by acting on proteins that are part of the pathway, or proteins that regulate the pathway (Luttrell *et al.*, 1996).

As with RTKs, NRTKs are activated when the activation loop tyrosines are phosphorylated. Phosphorylation can occur in-trans or by a different NRTK. In addition to a tyrosine kinase domain, NRTKs possess domains that mediate protein-protein, protein-lipid, and protein-DNA interactions (Fig. 1.4). The most common protein-protein interaction domains in NRTKs are the Src homology 2 (SH2) and 3 (SH3) domains.

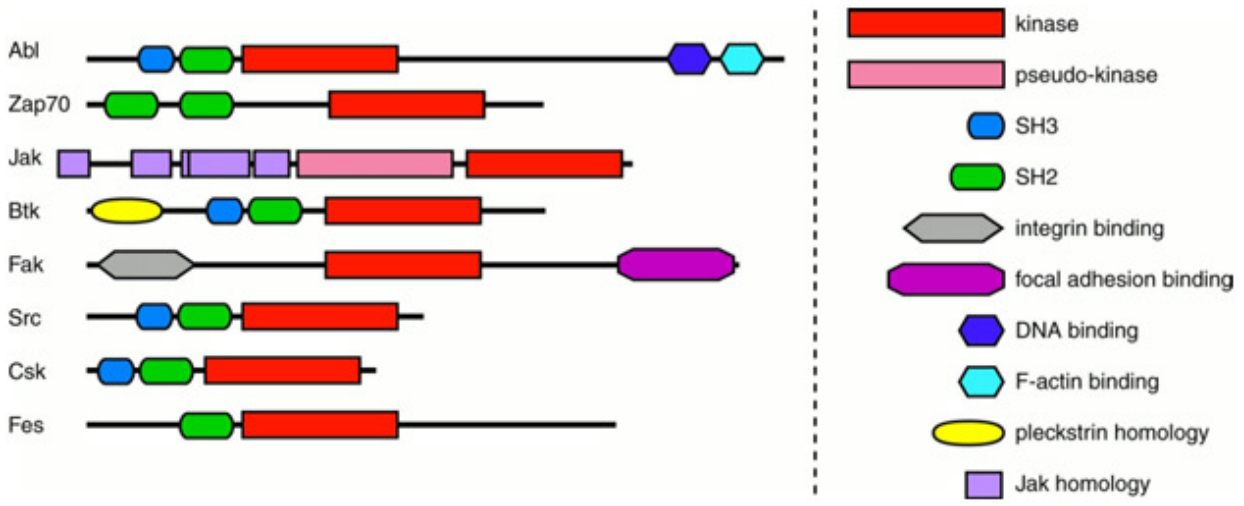


Fig. 1.4. Domain architecture of different NRTK families (Hubbard and Till, 2000). Each NRTK family has the same structural organization and is indicated to the left of each module. Besides the conserved tyrosine kinase domain, other domains mainly for protein-protein interactions are also found in each protein.

1.3.2 Example 2: Ion channel receptors

The human brain consists of approximately 20 billion neurons, with an average of 7500 synaspes per neuron (Pakkenberg *et al.*, 2003; Pakkenberg *et al.*, 1997). At every synaptic interface, chemical signals are transmitted from the axon of the effector cell to the dendrites of the receiving cell. These extracellular chemical signals released and sensed by neurons are collectively known as neurotransmitters, and serve to convert and relay electrical signals in the form of chemical messengers across the space between two communicating neurons.

There are two types of neurotransmitter receptors – those that are able to form ion channel pores (ionotropic), and those that are G-protein coupled receptors (metabotropic). When captured by receptors on the receiving dendrite, neurotransmitter ligands initiate depolarisation of the cell either through the opening of an innate channel in ionotropic receptors, or through signal transduction mechanisms via G-proteins linking to ion channels.

Within each neuron, transmembrane ion pumps and receptor channels maintain a resting potential of -70mV by transporting sodium ions out of the cell and potassium ions into the cell against a concentration gradient. Neurotransmitter activation initiates the opening of sodium and calcium receptors through conformational changes. If this event causes a depolarization beyond a threshold of -60mV, positive feedback causes further depolarization to +40mV through the opening of more voltage-gated sodium channels (VGSC), followed by the efflux of potassium ions when its channel opens later. After this electrical impulse called the action potential has passed, the ion channels rapidly close and sodium and potassium pumps return the cell membrane to its resting potential (Fig. 1.5). During this time, ion channels undergo an inactivation period during a refractory stage when they cannot be stimulated or require a larger stimulus to be activated. This process is known as the action potential, and is essential in all nervous signal transduction.

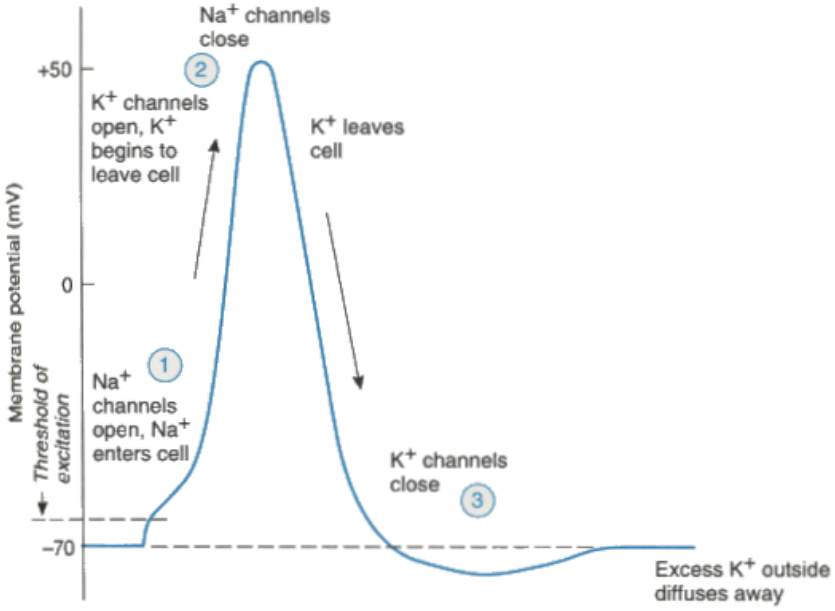


Fig. 1.5. The action potential during nervous transmission (Adapted from mindcreators.com; Cofer, 2002). The axes represent change in membrane potential against time in millisecond. The cell is normally kept at -70mV in resting state with high concentration of K+ and low Na+ inside the cell against a concentration gradient. An initial

10mV depolarization is the threshold required to initiate an action potential in which the sodium channels first open to allow an influx of Na+ ions (1). Potassium channels later open (2) allowing the efflux of K+ to hyperpolarize the cell. At the end of the action potential, K+ channels close (3) and ion pumps revert the Na+/K+ concentrations back to its resting state in preparation for another wave of excitation.

Ion channel receptors are classified according to ligand specificities. All are assembled from subunits of homologous polypeptides that arise from alternative splicing and arranged in a ring structure with a water filled channel running perpendicular to the plasma membrane. The cytoplasmic extension of the channel may contain stretches of amino acids for interacting with other proteins.

1.4 Mitogen-activated protein kinase pathways

Among the pathways used to transduce a signal is the highly conserved mitogen-activated protein kinase (MAPK) pathway. This pathway was initially found to be activated by mitogens, and a critical regulator of cell division and differentiation. Subsequent studies have shown it to be also inducible by physical and chemical stressors such as UV irradiation, heat and osmotic shock. The MAPK pathway is the prototypic signalling cascade found in all eukaryotic organisms and in its most basic form, consists of the sequential phosphorylation and activation of at least three protein kinases, MAPKKK, MAPKK and MAPK.

Mitogen-activated protein (MAP) kinase kinase kinases (MAPKKK), also known as MAP3K or MEKK, integrate multiple inputs which, in turn, phosphorylate MAP kinase kinases (denoted MAPKK, MAP2K, MKK or MEK). MAPKK enzymes are dual-specificity enzymes with limited substrate specificity, belonging to the MAP/ERK (MEK or MKK) family of kinases. Although activated by multiple MAPKKKs, MAPKKs only phosphorylates serine/threonine

(Ser/Thr) and Tyr residues on one or a few MAP kinase (MAPK) proteins (Kosako *et al.*, 1992). Phosphorylated MAPKs are known to control almost all cellular processes ranging from gene expression to cell death (Chang and Karin, 2001) by migrating to the nucleus to recognise transcription factors with the motif (S/T)P, PX(S/T) or (S/T)G (Lewis *et al.*, 1998). In mammals, this pathway can be sub-divided into at least four distinct groups – ERK/MAPK, JNK/SAPK, p38 and ERK5 pathways – according to the specific MAPK activated (Fig. 1.6).

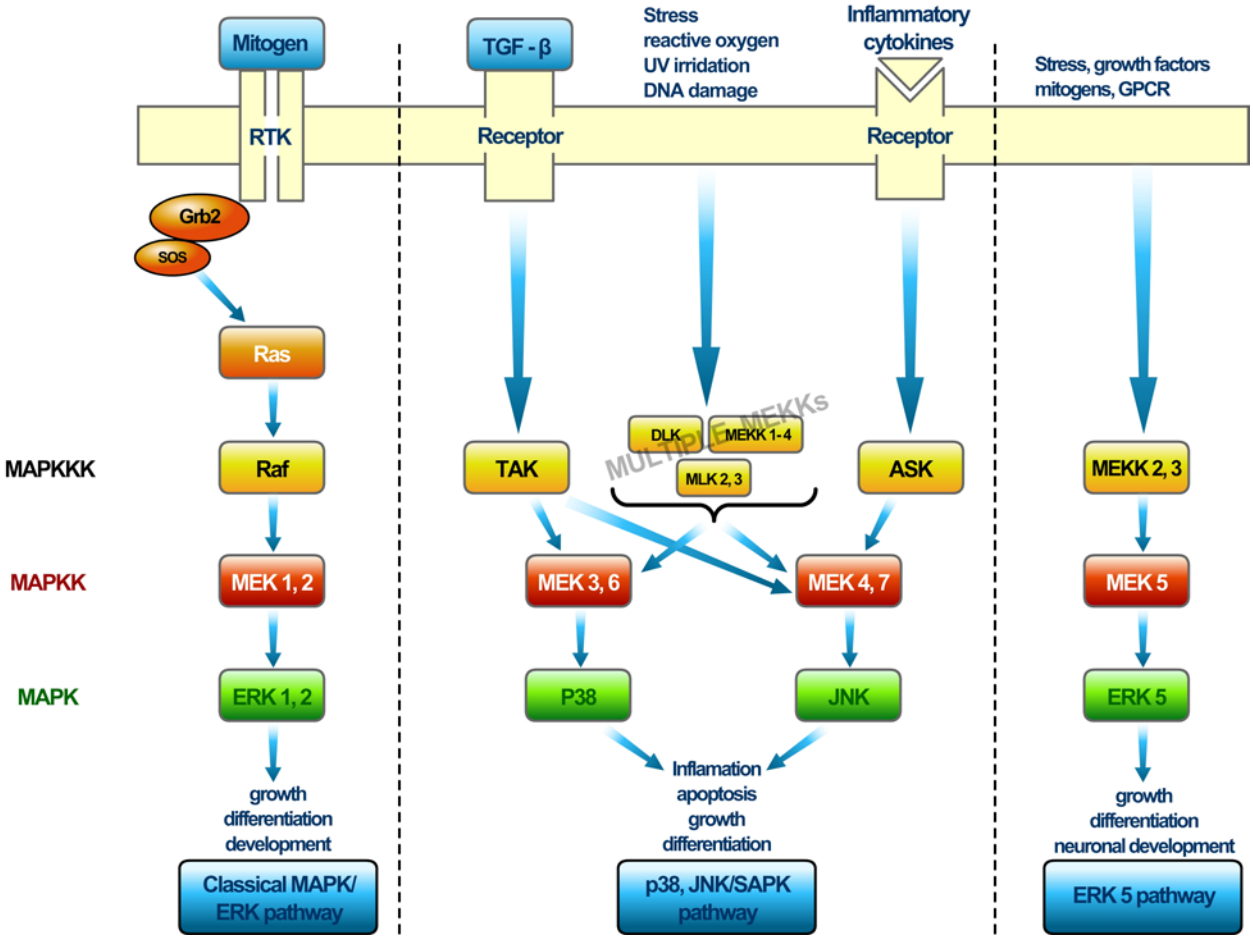


Fig. 1.6. The different classes MAPK signalling pathways in humans (ERK/MAPK, p38 and JNK/SAPK and ERK5). MAPKKK (MAP kinase kinase kinases) are represented as yellow boxes, MAPKK (MAP kinase kinases) as red boxes and MAPK (MAP kinases) as green boxes.

1.4.1 Classical ERK/MAPK pathway

The MAPK/ERK pathway (Fig. 1.5) is induced by ligands received by specific RTKs; and, upon binding to their respective growth factors, oligomerize and autophosphorylate tyrosine residues in trans. These phosphorylated tyrosines signal the recruitment of adaptor proteins such as Grb2. Guanine nucleotide exchange factor Son of Sevenless (SOS) translocates to the membrane, associates with Grb2 and relays the signal by catalysing the activation of Ras by exchanging GDP to GTP. Ras-GTP activates Raf-1, a ubiquitous MAPKKK, by binding to and relieving its inhibition by the N-terminal regulatory domain (Yip-Schnider *et al.,* 2000). Raf-1 regulation is complex, involving several kinases and phosphatases (Avruch *et al.,* 2001; Jelinek *et al.,* 1996; Warne *et al.,* 1993; Moodie *et al.,* 1993, Pearson *et al.,* 2001; von Kriegsheim *et al.,* 2006; Jaumot and Hancock, 2001). Once activated, Raf-1 dually phosphorylates either serine or threonine in the activation loop of MAPKKs MEK1 and MEK2.

The extracellular signal-regulated kinases (ERK) are the most common and best characterised serine/threonine protein kinaes that are activated downstream of the ERK/MAPK cascades. Occuring in two closely related isoforms ERK1/p44 and ERK2/p42, they are possibly unique substrates to MEK1/2. Although ubiquitously expressed, their relative expression levels vary between tissues and they are predominantly associated with positive cellular responses such as proliferation, differentiation, migration and tissue growth. One of the most explored functions of MAPK signalling modules is the regulation of gene expression in response to extracellular stimuli. While the primary site of MAPK action, such as ERKs, is inside the nucleus to phosphorylate transcription factors prebound to DNA, much remains in the cytoplasm to regulate gene expression through post transcriptional means such as mRNA stabilisation (Chen *et al.,*

2000; Winzen *et al.*, 1999; Lasa *et al.*, 2000) and translational control (Kotlyarov *et al.*, 1999; Pyronnet *et al.*, 1999).

1.4.2 An adaptation of the ERK/MAPK pathway: Synaptic plasticity

There are two forms of memory: short-term working memory, and long-term memory. Memory serves a basic purpose to increase an organism's survivability by recognizing food, danger, prospective mates and environmental cues. Organisms have thus developed a complicated yet reliable system of constant addition, expansion, alteration and eradication of memories to better serve their interests.

Memories are formed in the hippocampus and amygdala of the brain where the binding of glutamate to N-methyl-D-aspartate receptor (NMDAR) causes both the influx of calcium through opening of the ion channels and activation of calcium-dependent kinases that lasts between 60 to 90 minutes (Sweatt, 2001). At the same time, calcium channel, α-amino-3-hydroxy-5-methylisoxazole-4-proprionate receptor (AMPAR) increases in number at the synapse, is phosphorylated and generates a greater influx of calcium. This phenomenon reinforces the synaptic connections between the effector and effected neurons and is known as long term potentiation (LTP), more specifically the early phase of LTP. LTP is the long lasting communication improvement between two neurons that results from stimulating them simultaneously, while synaptic plasticity is the ability of two neurons to establish this connection. The LTP theory is a signalling model for memory formation, and the early phase of LTP indicates the start of memory formation.

1.4.3 ERK/MAPK pathway in late LTP

Maintaining LTP in activated neurons is essential for the formation of long-term memory. Following NMDAR activation, a cascade of kinase activities ensue that ultimately changes gene expression levels in the late phase of LTP. If this is not sustained, short-term memory formed during early LTP will be erased. During late LTP, the classical ERK/MAPK pathway of Ras/Raf, MEK and ERK integrates signals from neurotransmitter receptors, Src and Ca^{2+} (English and Sweatt, 1997; Gottschalk *et al.*, 1999; Boxall and Lancaster, 1998; Thomas and Brugge, 1997) (Fig. 1.7). ERK2 is essential in memory consolidation by phosphorylating calcium and cAMP response element binding transcription factor, enabling transcription of its genes (Cestari *et al.*, 2006). Activated ERK also phosphorylates the voltage-gated potassium channel (Schrader *et al.*, 2006); tyrosine phosphorylation of NMDAR or associated proteins also lead to increased channel activity and influx of calcium ions, which in turn strengthen AMPA mediated synaptic transmission (Purcell and Carew, 2003).

Increased post synaptic receptor activity as a result of heightened cell sensitivity towards signals during learning of is important for enhancing communication between neurons (Bliss and Lomo, 1973; Anderson *et al*., 1980). This is attributed in part to the increased sensitivity of the voltage-gated sodium channel (VGSC) to membrane depolarization through a progressive decrease in action potential threshold. Together with an increased input resistance due to a hyperpolarized shift in slow inactivation curve of VGSC, the excitability of postsynaptic neurons is enhanced (Xu *et al*., 2005). This change in activation kinetics of the VGSC was found to be Ca^{2+}/CaM dependent, and a result of protein synthesis (Xu *et al.*, 2005).

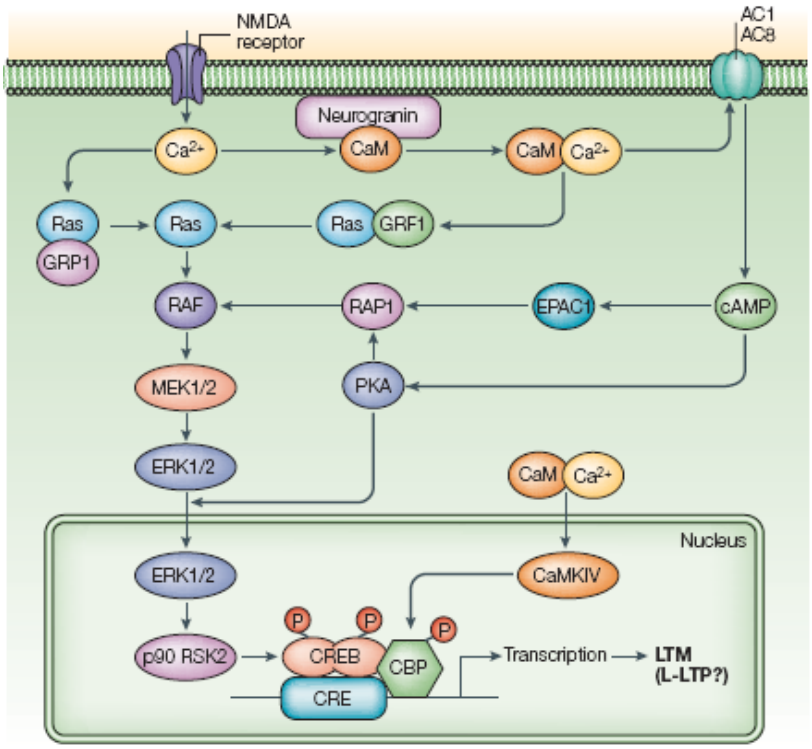


Fig. 1.7. The role of CaM in a simplified ERK/MAPK pathway during neuronal LTP for the formation of memories (Xia and Storm, 2005). CaM acts as a Ca^{2+} sensor to modulate the function of various proteins in different concentrations of Ca^{2+}, caused by neurotransmitter ligands that initiate the opening of NMDA and AMPA receptors to allow an influx of Ca^{2+}.

1.5 Regulating signal transduction pathways at the receptor level

To create varying cellular responses, cells manipulate the same signalling pathways by altering the amplitude, duration and location of its activation. Consequently, several mechanisms exist to ensure that the appropriate signal thresholds are achieved and maintained for the correct length of time or aptly attenuated for strict regulation of receptor activity. This need for tight regulation of catalytic activity is underscored by the numerous diseased states that result from deregulated pathways. Controlled signalling within the cell is predominantly achieved by inbuilt on/off switches and negative feedback mechanisms. Diverse classes of molecules have evolved for this function, ranging from enzymes to inorganic factors. Those

mechanisms that do not degrade a pool of proteins are deemed as transient - their targets can be re-activated once inhibition is lifted. On the other hand, the most common mechanism to permanently attenuate a signal is through protein internalization and degradation.

1.5.1 The role of regulatory ligands and proteins in modulating receptor activity

Just as the essential components of signal transduction pathways have evolved methods to enhance and efficiently propagate a signal, many other proteins exist to control amplitude, duration and location of pathway factors. Known as regulatory proteins, these proteins are key determinants of specific cellular outcomes, and although critical to the fate of an organism, they are not essential for the initiation and propagation of signal.

Regulatory mechanisms often engage competitive inhibitory oligomerization with ligands or other proteins to carry out their function. For example, the Tie2 receptor whose signal is initiated by angiopoietins binding has a natural inhibitory ligand that prevents further activation of the receptor (Maisonpierre *et al.*, 1997). Certain tissues purposefully express receptor variants deficient in tyrosine kinase activity for dominant negative inhibition through the generation of inactive heterodimers. In other cases, PTKs require association with other proteins or homodimerization to become functional. For example, engagement of Syk with the tyrosine-phosphorylated ζ-chain of the T-cell receptor relieves inhibiton on the kinase domain and stimulates its catalytic activity (Shiue *et al.*, 1995).

An activated receptor usually induces positive and negative pathways simultaneously that are functionally connected by numerous feedback mechanisms. For example, the EGF receptor in *Drosophila* mediates the expression of Spitz and Argos, positive and negative auto-regulators

of the EGF induced MAPK pathway respectively (Wasserman and Matthew, 1998) to define body axes and patterning in developing oocytes. Src, when in its activated state, phosphorylates a conserved tyrosine in the N-terminal of Sprouty proteins. This phosphorylated tyrosine residue can bind and sequester Grb2, impairing the recruitment of SOS to RTKs (Gross *et al.*, 2001). The precise mechanism of Sprouty proteins in the MAPK signalling however, has yet to be fully elucidated

Example: Sprouty proteins regulating receptors with tyrosine kinase activities

Sprouty was first discovered in *Drosophila* as a 63kDa protein that inhibited fibroblast growth factor (FGF) receptor signalling (Hacohen *et al.*, 1999) in tracheal branching. Subsequent studies revealed that there were four human Sprouty homologues, all considerably smaller than *Drosophila*'s possessing two conserved sequences – a highy homologous Cys rich C-terminus, and a short stretch in the N-terminus centred around a phosphorylated Tyr (pTyr). Sprouty has no intrinsic enzymatic activity of its own, but its expression pattern during embryonic development synchronises with known sites of RTK signalling and localises to the membrane through the Cys rich C-terminus (Kim and Bar-Sagi, 2004). Prturbation of the RTK signalling that result from the expression of Sprouty proteins produces a broad spectrum of effects that range from alterations to developmental fate and changes in cellular homeostasis (Kim and Bar-Sagi, 2004).

Upon receptor stimulation, the essential tyrosine within the SH2-like binding motif of Sprouty becomes tyrosine phosphorylated. This phosphorylation event has been linked to the ability of Sprouty to downregulate ERK phosphorylation. Combined evidence points to two areas of action within the ERK/MAPK pathway: upstream of Ras and upstream of Raf-1

(Hanafusa et al., 2002; Lao *et al*., 2006). Under normal circumstances, EGF signalling triggers a negative regulatory response to downregulate the receptor from the cell-surface via c-Cbl mediated ubquitination. Sprouty through its binding to the conserved tyrosine in the N-terminus, competes with EGFR for binding to Cbl, and prevents it from interacting with, and thus downregulating the EGFR (Fong *et al.*, 2003).

1.5.2 The role of function-modifying proteins in regulating receptor activity

Besides down-regulation, there are also certain proteins whose nature is to modify the function of their binding partners. The ubiquitous 14-3-3 protein isoforms are probably one of the best characterized as effecting both positive and negative regulation. In the ERK/MAPK pathway, Ras activity is upregulated by its binding with the catalytic domain of 14-3-3 proteins, but suppressed through association with the regulatory domain (Yaffe, 2002). Other proteins possessing such modulatory roles include the Regulator of G-protein signaling (RGS) protein, primarily to enhance the GTPase activity of an activated G-protein, and calcium binding proteins S100 and calmodulin (CaM) that the ability to be regulated by intracellular Ca^{2+}.

Example: Calmodulin regulation of ion channel receptor activity

Various stimuli, such as changes in membrane polarization or small receptor ligands may trigger the opening of calcium channels, resulting in the influx of Ca^{2+} ions into the cytosol. Many forms of synaptic plasticity are initiated by an increase in intracellular calcium ions (Ca^{2+}) which functions as a second messenger for activity dependent, synapse specific changes. The

approximately 100-fold increase in free Ca^{2+} concentration during synaptic transmission from NMDAR and AMPAR allows Ca^{2+}-binding proteins to trigger response mechanisms that are able to integrate and transmit this signal coherently to downstream processes.

The key Ca^{2+} regulator protein in the brain is calmodulin (CaM), a 16.8kDa highly conserved ubiquitous protein whose sequence is identical among vertebrates. Many proteins, including ion channel receptors, alter their activity in response to changes in free Ca^{2+} levels, but are not able to bind Ca^{2+} ions by themselves. CaM, when bound to several ion channels crucial for synaptic plasticity, confers Ca^{2+} sensitivity to its binding partners through a conformational change (Fig 1.8). CaM's regulatory role in mediating the Ca^{2+} signal is highly localized and so entrenched in regulating ion channel receptors' functions that it has been suggested to to be classified as a channel subunit (Saimi and Kung, 2002). The importance of CaM is reflected in the extraordinarily high concentrations ranging from 10μM to 100μM in neurons and lethality if its gene is deleted (Xia and Storm, 2005).

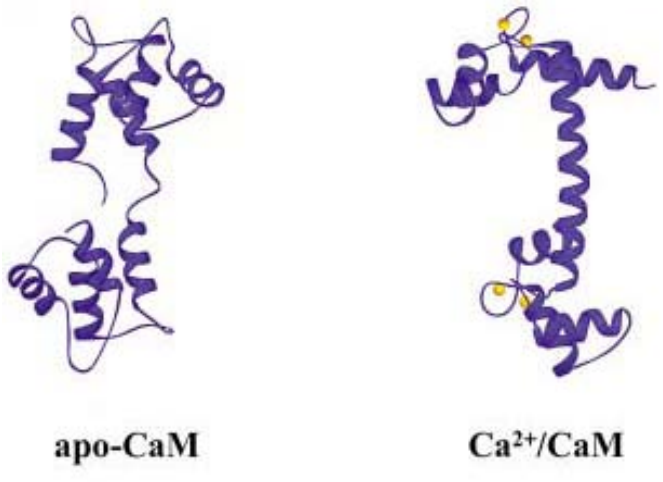


Fig. 1.8. Ribbon presentations of CaM and CaM in complex with target peptides. CaM is colored blue and Ca^{2+} ions are yellow. The N- and C- lobes are at the ends of the molecule connected a central α-helical linker in Ca^{2+}/CaM. The linker collapses in apo-CaM. Structural data were taken from the Protein Data Bank, accession codes: apo-CaM (1CFD) and Ca^{2+}/CaM (1CLL) (Vetter and Leclerc 2003).

Several hundred Ca^{2+}-binding proteins have been identified to contain the EF-hand Ca^{2+} binding motif. This motif comprises about 30 amino acids and consists of a helix-loop-helix where the two helices are arranged similar to the extended thumb and index finger of a hand: it is commonly called the EF-hand domain. In almost all Ca^{2+}-binding proteins such as CaM, two EF-hand domains are in close proximity forming an EF-hand pair.

CaM is a dumbbell shaped molecule with two EF-hand pairs: the N- and C- lobes that are arranged one at each terminus connected by a long linker (Fig. 1.8). Each lobe binds two Ca^{2+} ions, and is connected to each other by an α-helical linker that bends with changes in intracellular Ca^{2+} concentration (Vetter and Leclerc, 2003). The lobes share 48% sequence identity and 75% similarity, a difference that is reflected in 10-fold higher Ca^{2+}-binding affinity of the C-lobe. The average dissociation constant of Ca^{2+} from CaM is 15μM in the absence of other proteins. The affinity for Ca^{2+} increases when CaM is complexed with a target protein and, with the exception of neurogranin and neuromodulin, Ca^{2+} recruitment enhances CaM's affinity for its targets (Olwin and Storm, 1985). These changes in affinity occur because Ca^{2+} binding exposes a hydrophobic patch, the main site of interaction between CaM and its targets (Xia and Storm, 2005).

Recent observation have shown that CaM is an important regulator of many different ionic currents by binding to and modulating the functions of channel receptors, such as the rod cGMP gated cation channel, NMDAR, calcium activated potassium channel and the voltage-gated calcium channel (VGCC) (Schumacher *et al.,* 2001; Petegem *et al.,* 2005) (Fig 1.9). CaM acts as an intracellular calcium sensor for these channels when bound to them, thus translating Ca^{2+} signals into cellular responses according to the state of the cell.

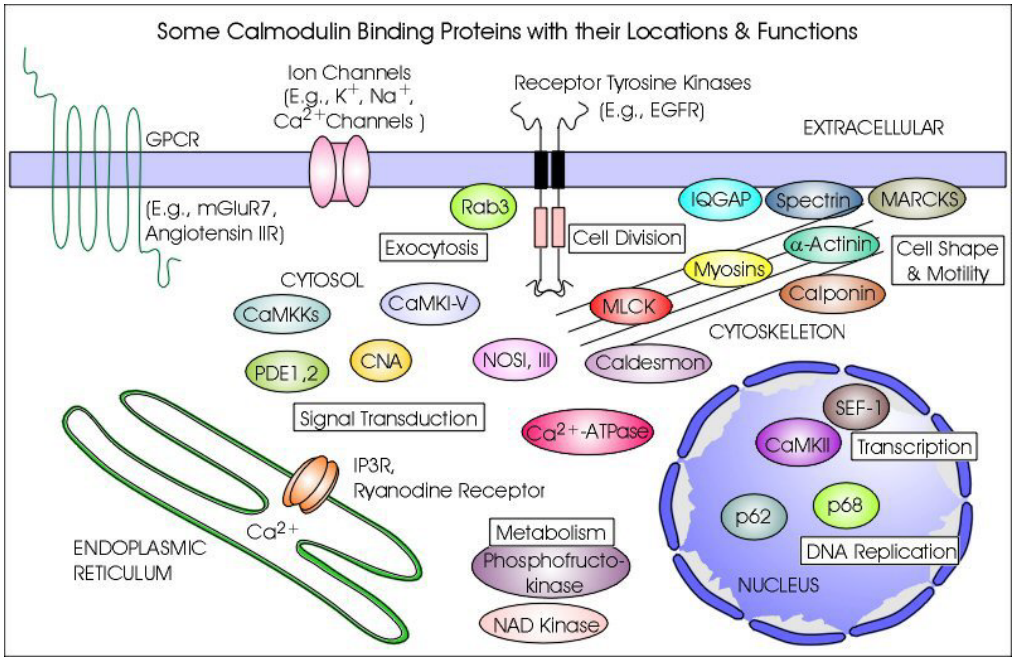


Fig. 1.9. Summary of CaM binding proteins (O'Day, 2003). CaM target proteins are varied in function and localisation, ranging from structural proteins to transcription factors.

1.5.3 The role of post-translational modifications in regulating receptor activity

Post-translational modifications (PTMs) are covalent processing events that change the properties of a protein by proteolytic cleavage or by the addition of a modifying group to one or more amino acids (Fig. 1.10). PTMs of a protein can determine its activity, state, localization, turnover, and interactions with other proteins (Mann and Jensen, 2003). Proteolytic cleavage can relieve auto-inhibition, or release a protein for folding and function after secretion. Modifications like hydroxylation, acetylation, methylation, ubiquitination and phosphorylation are frequently utilised for controlling protein interactions by completing substrate recognition sites of a target protein (Table 1.1). Kinase cascades are turned on and off by the reversible addition and removal of phosphate groups (Cohen, 2000), while ubiquitination marks targeted proteins for destruction at defined time points (Tyres and Jorgensen, 2000).

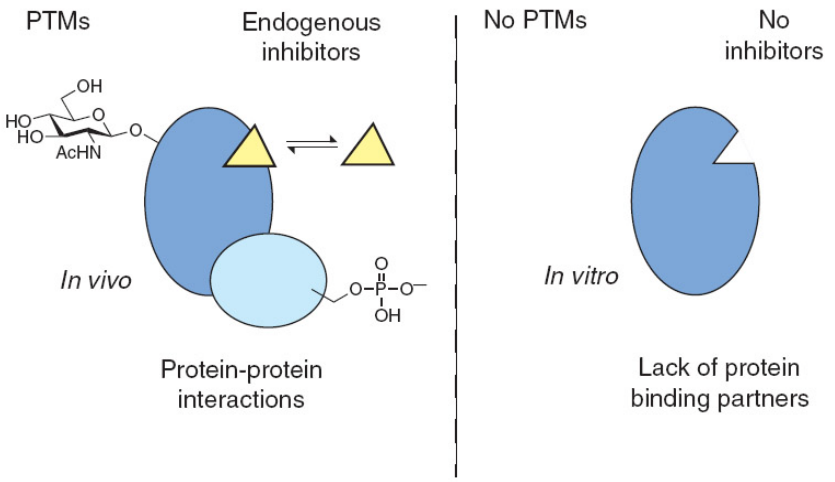


Fig. 1.10. Comparison of the *in vivo* (with PTMs) and *in vitro* (without PTMs) states a protein experiences, highlighting several regulatory modes conferred by post-translational modifications (Saghatelian and Cravatt, 2005).

Table 1.1. List of common and important PTMs (Mann and Jensen, 2003).

PTM type	ΔMass[a] (Da)	Stability[b]	Function and notes
Phosphorylation			Reversible, activation/inactivation of enzyme activity, modulation of molecular interactions, signaling
pTyr	+80	+++	
pSer, pThr	+80	+/++	
Acetylation	+42	+++	Protein stability, protection of N terminus. Regulation of protein–DNA interactions (histones)
Methylation	+14	+++	Regulation of gene expression
Acylation, fatty acid modification			Cellular localization and targeting signals, membrane tethering, mediator of protein–protein interactions
Farnesyl	+204	+++	
Myristoyl	+210	+++	
Palmitoyl	+238	+/++	
etc.			
Glycosylation			
N-linked	>800	+/++	Excreted proteins, cell–cell recognition/signaling
O-linked	203, >800	+/++	*O*-GlcNAc, reversible, regulatory functions
GPI anchor	>1,000	++	Glycosylphosphatidylinositol (GPI) anchor. Membrane tethering of enzymes and receptors, mainly to outer leaflet of plasma membrane
Hydroxyproline	+16	+++	Protein stability and protein–ligand interactions
Sulfation (sTyr)	+80	+	Modulator of protein–protein and receptor–ligand interactions
Disulfide bond formation	–2	++	Intra- and intermolecular crosslink, protein stability
Deamidation	+1	+++	Possible regulator of protein–ligand and protein–protein interactions, also a common chemical artifact
Pyroglutamic acid	–17	+++	Protein stability, blocked N terminus
Ubiquitination	>1,000	+/++	Destruction signal. After tryptic digestion, ubiquitination site is modified with the Gly-Gly dipeptide
Nitration of tyrosine	+45	+/++	Oxidative damage during inflammation

[a]A more comprehensive list of PTM Δmass values can be found at: http://www.abrf.org/index.cfm/dm.home
[b]Stability: + labile in tandem mass spectrometry, ++ moderately stable; +++ stable.

Example 1: Phosphorylation

One of the most common modifications to proteins is the phosphorylation of serine, threonine and tyrosine residues. About 30% of all cellular proteins are phosphorylated at any time; abnormal phosphorylation is now recognized as a cause or consequence of many disease states (Cohen, 2002). Phosphorylation and dephosphorylation is catalysed by protein kinases and protein phosphatases respectively, in order to modify the function of a protein by modulating its biological activity. Phosphorylation events function to stabilise a protein or mark it for destruction, facilitate or inhibit movement between subcellular compartments, as well as initiate or disrupt its binding to substrates and binding partners.

Iakoucheva *et al.,* 2004 found that the majority of protein phosphorylation sites are contained in intrinsically disordered protein segments, such that a phosphorylation site should be embedded into a generally hydrophilic and conformationally flexible sequence environment. The simplicity, flexibility and reversibility of phosphorylation, coupled with readily available sources of ATP as a phosphoryl donor may explain why phosphorylation has been evolutionarily selected as the most common regulatory mechanism by eukaryotic cells.

Example 2: Ubiquitination

Definitive inhibition is most commonly mediated by activation-deactivation protein degradation. The degradation of cellular signalling proteins following their ubiquitination plays a critical role in controlling multiple physiological processes. To terminate a signal, activated proteins are removed from the membrane by endocytic mechanisms and either recycled or degraded, generating a refractory period before the next signal can be transmitted. In cells, this

process is largely regulated by ubiquitination, where ubiquitin moieties are covalently attached to defined lysine residues within the target protein. Ubiquitination is mediated through the sequential enzymatic cascade involving ubiquitin activating (E1), conjugating (E2), and ligating (E3) proteins. Additional ubiquitin molecules can be added to the originally ubiquitinated lysines to generate ubiquitin chains. The type and length of chain generated dictates the mechanism by which the protein is then downregulated. Monoubiquitination directs a protein to internalization and endosomal sorting, while presence of multi-ubiquitin chains on proteins targets them for degradation by the 28S proteasome (Weissman, 2001).

1.6 The importance of protein – protein interaction domains and motifs

Over the last decade, we have gained considerable insights into the mechanisms by which signals are accurately conveyed from receptors at the plasma membrane to targets in the cytoplasm and nucleus. At the heart of the abovementioned interactions, specificity and regulatory processes is a recurring theme of molecular recognition.

Proteins recognize and bind to each other through stretches of amino acid sequences built into certain architectural folds known as domains. Each domain typically comprises 50 - 200 amino acids and folds into a discrete entity (Fig. 1.11), with its N- and C- termini juxtaposed in space and away from the interaction region (Schlessinger and Lemmon, 2003). This enables it to be inserted in tandem with other domains on the same protein without interference to its own binding (Pawson and Nash, 2003). A domain usually recognizes a core determinant, with flanking or noncontiguous residues providing additional contacts and an element of selectivity. They can be broadly classified into four groups – those that bind (1) short peptide motifs, (2)

associate with other domains, (3) recognise non amino acids or (4) bind post-translationally modified amino acids.

Fig. 1.11. A list of protein interaction domains (Adapted from Pawson Lab; pawsonlab.mshri.on.ca). Each domain is schematically represented by a different icon and not reflective of the true domain architecture.

Interaction domains may target proteins to a specific subcellular location, provide a means for recognising post-translational modifications or chemical second messengers, nucleate the formation of multiprotein signalling complexes, and control the conformation, activity, and substrate specificity of enzymes (Pawson and Nash, 2000). Sometimes, individual domains engage several distinct ligands either simultaneously or at successive stages of signalling.

Since the entire cellular function organized in signal transduction networks is built on a restricted group of protein domains, each class of domains must have multiple biological functions and act in a combinatorial fashion, since there is insufficient variety to have each assuming a single biological role (Pawson and Nash, 2000). The incorporation of different domains into a single polypeptide is nature's way of overcoming the limitation of a few thousand gene products having to control the entire signalling network. Multiple domains may interact with different sites on the same target, as commonly occuring in polypeptides that possess tandem SH2 domains, thereby increasing both the affinity and specificity of protein-protein interaction (Ottinger *et al.,* 1998) or engage in complex intramolecular interactions that regulate the enzymatic activities of their host protein.

Regulatory proteins also make use of such domains to their advantage. They are often constructed from domains that mediate molecular interaction or have enzymatic activity to competitively inhibit association of a protein to its substrate (Pawson and Nash, 2003). The cell therefore uses a limited set of interaction domains, which are joined together in diverse combinations, to direct the actions of regulatory systems.

1.7 Case study 1: Cbl regulation of receptors with tyrosine kinase activity

The c-Cbl proto-oncogene was first discovered through studies of hemopioetic tumours of mice infected with the Cas-Br-M retrovirus. The isolated recombinant retrovirus induced pre- and pro-B lymphomas and caused the transformation of rodent fibroblasts. The causative retrovirus was called Cas-NS-1 and its oncogene, v-Cbl, for Casitsas B-lineage lymphoma (Langdon *et al.,* 1989). Subsequent cloning of the mouse c-Cbl gene revealed that v-Cbl

encoded the first 355 amino acids of a longer 120kDa ubiquitously expressed full length protein, and overexpression of wild-type c-Cbl did not promote tumourgenesis (Blake *et al.*, 1991).

Two other mammalian homologues, b-Cbl and Cbl-c have since been identified. All three family members share a highly conserved N-terminus that corresponds to residues in v-Cbl. Originally assigned as an adaptor protein, Cbl proteins were revealed to possess E3 ubiquitin ligase activity through an adjacent RING domain that recruits E2 ubiquitin conjugating enzymes. An extreme C-terminal ubiquitin-associated domain overlaps a leucine zipper motif. Between the RING and the latter is a proline rich region for associations with SH3 domains (Fig. 1.14A).

About 150 proteins have been shown to be regulated by Cbl proteins (Fig. 1.12). More recently, it was found to play a prominent role in mediating ligand dependent downregulation of PTKs (Thien and Langdon, 2001) by targeting them to the lysozome (Shtiegman and Yarden, 2003).

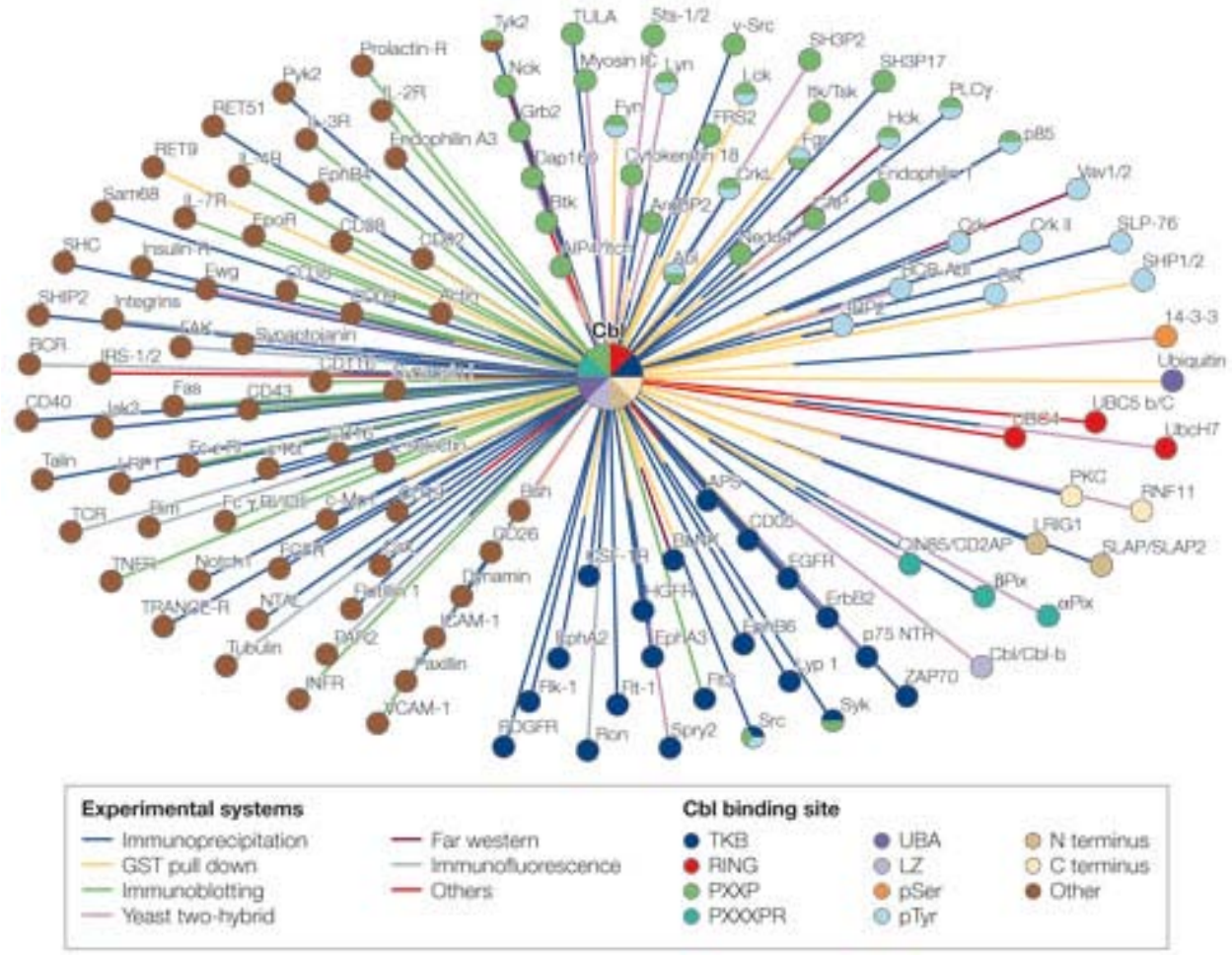


Fig. 1.12. The Cbl interactome(Schmidt and Dikic, 2005). An Osprey diagram of proteins currently known to interact with Cbl. Proteins were grouped according to the domains in Cbl to which they bind, which are represented by node colours and the colours in the centre. Methods that were used for the detection of protein-protein interactions are represented by different line colours.

1.7.1 RTK ubiquitination by Cbl in the context of EGFR and Met

In the context of signalling by the EGFR, one of the first proteins to be recruited to the phosphorylated dimerised activated EGFR is Grb2, which can then recruit Cbl proteins from the cytoplasm to the plasma membrane through interactions with the proline rich region on Cbl and the SH3 domain on Grb2. Autophosphorylation of EGFR also creates additional docking sites for the TKB domain of Cbl (Fig. 1.13). Y1068 constitutes a direct binding site for the c-Cbl

tyrosine kinase binding (TKB) domain and is required for c-Cbl mediated ubiquitination and degradation of EGFR (Levkowitz *et al.,* 1999). Mutation of residues that comprise the c-Cbl TKB domain binding site elicits stronger mitogenic signals than the wild type receptor (Waterman *et al.,* 2002). Up to 40% glioblastomas express oncogenic mutants that disrupt Cbl's interaction with EGFR. The deregulation of Src which occurs in human brain, breast, colon and lung cancers also leads to the stabilization of the EGFR by promoting Cbl ubiquitination and degradation (Bao *et al.,* 2003). Throughout its internalization route, the prevailing role of Cbl seems to be retaining receptors inside the sorting endosome by promoting and maintaining their ubiquitination (Schmidt and Dikic, 2005).

Met represents another class of RTK regulated by Cbl. The c-Cbl TKB domain also binds to a juxtamembrane tyrosine (1003) residue. This interacton is essential for ubiquitination and degradation of the receptor (Preschard *et al.,* 2001). The Y1003F mutant has a prolonged half life and is oncogenic in cell culture and tumourgenesis assays, identifying c-Cbl and ubiquitination as important negative regulators of this protein.

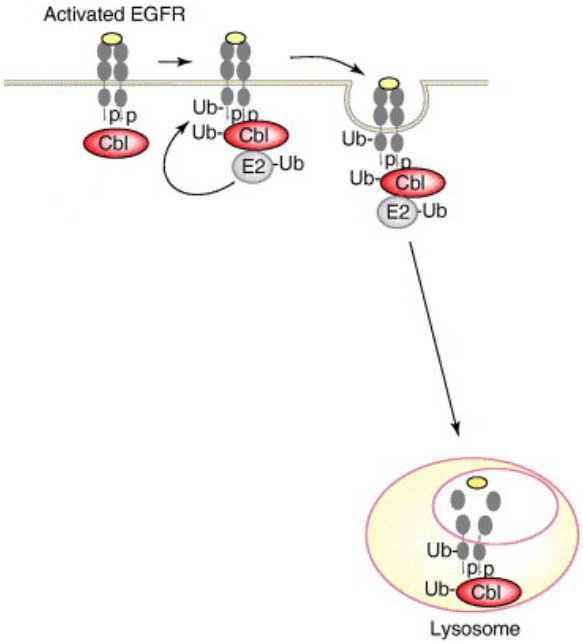


Fig. 1.13. The endocytotic and degradation pathway of EGFR via ubiquitination by c-Cbl (Ryan *et al.,* 2006). Phosphotyrosines in activated EGFR serve as sites for Cbl to dock and orientate to a position favourable for ubiquitination. Cbl recruits the E2 ubiquitin ligase to its RING domain which adds ubiquitin moieties to target lysine residues in EGFR. The whole complex becomes internalized by clarithrin mediated endocytosis and subsequently degraded in the lysozome.

1.7.2 NRTK regulation by Cbl

The relationship between Cbl and NRTKs is complex. Cbl regulates NRTK activity through ubiquitination and degradation. At the same time, it is also being regulated by NRTKs through phosphorylation. c-Cbl has 22 Tyr residues available for phosphorylation at the C-terminus and is a prominent substrate of PTKs. 700Y, 731Y and 774Y are principal phosphorylation sites (Feshchenko *et al.* 1998) by Syk and Src family kinases. Phosphorylated Cbl becomes targets for auto-ubiquitination to remove its activity, resulting in prolonged RTK signaling. However, Cbl overexpression also suppresses the activity of Syk, its homolog ZAP-70 and c-Src kinase activity in a TKB dependent manner by binding to phosphorylated tyrosines of the activated kinases (Sanjay *et al.,* 2001). The complexity of c-Cbl's regulation of Src signaling show that Cbl can have both positive roles, by functioning as an adaptor to contribute to signal transduction, as well as negative roles, by downregulating Src kinase activity or enhancing its degradation.

1.7.3 The importance of Cbl-TKB domain in PTK regulation

While there are two families of domains that bind pTyr containing sequences – the phosphotyrosine binding (PTB) domain and the SH2 domain – only about a quarter of the 79 known human PTB domains have acquired a capacity for pTyr dependent recognition (Schlessinger and Lemmon, 2003). SH2 domains, however, are wholly dedicated to recognising pTyr and thus represent the largest class of known pTyr recognition domains for primary targeting and specificity elements in tyrosine kinase signalling (Pawson *et al.,* 2001). Each SH2 domain has the ability to preferentially bind a specific phosphorylated motif (Songyang *et al.,*

1993). Since its discovery, a total of 120 different SH2 domains contained in 110 proteins from 11 functional categories have been discovered, including Cbl family of proteins.

c-Cbl interacts with a diverse array of proteins via phosphorylatable tyrosine residues at its C-terminus (Feshchenko *et al.* 1998), via its proline-rich region, and through the highly conserved N-terminal phosphotyrosine-like binding domain (Schmidt and Dikic, 2005). Although this phosphotyrosine binding domain (PTB) appeared to have several hallmarks of a PTB consensus motif, the crystal structure demonstrated that it structurally resembled an SH2 domain but required a flanking four-helix bundle (4H) and an EF-hand subdomain to accomplish binding (Meng *et al.,* 1999). Together, these three sub-domains make up the tyrosine kinase-binding (TKB) domain, which is unique to Cbl proteins and often referred to as a specialized or 'embedded' SH2 domain (Liu *et al.,* 2006). Several oncogenic RTKs have lost the ability to recruit Cbl in a TKB mediated manner. Moreover, the stability of Cbl itself is regulated by ubiquitination. Tyrosine phosphorylation of c-Cbl by c-Src promotes auto-ubiquitination of Cbl and its degradation in a proteasome dependent manner (Yokouchi *et al.,* 2001).

Binding to the TKB domain is required for the subsequent conjugation of ubiquitin through the RING domain during the ubiquitination process. The consensus binding sequence targeted by the TKB domain was originally identified by Lupher *et al.* (1997) as D(N/D)XpY, but was later experimentally refined as (N/D)XpY(S/T)XXP. This motif is common to several PTKs, such as epidermal growth factor receptor (EGFR), colony-stimulating factor 1 receptor, ZAP-70 ζ-associated protein of 70 kDa, Src and Syk, as well as members of the Sprouty (Spry) family of Ras/mitogen-activated protein kinase inhibitors. Each of these targets binds to the TKB domain following phosphorylation of a central tyrosine residue in the consensus motif and,

in most cases, their protein expression levels are subsequently downregulated by ubiquitination (Schmidt and Dikic, 2005).

In recent years, two additional consensus motifs have been identified as targeted by the c-Cbl-TKB domain (Fig. 1.14B). The first, RA(V/I)XNQpY(S/T), is a derivation of the original sequence, with conserved residues extending back to the arginine at the pY-6 position. This sequence is conserved amongst a family of adaptor proteins based on the APS (Adapter with a Plekstrin homology and Src homology-2 domains) protein (Hu and Hubbard, 2005) and, unlike the PTKs, they are not substrates for ubiquitination, but use c-Cbl as a docker protein in the insulin receptor complex (Chiang *et al.*, 2001). The second consensus sequence was identified in a study investigating the binding of c-Cbl to the c-Met receptor (Met) but, unlike the APS sequence, it bears no resemblance to the previously characterized PTK-binding motif (Peschard *et al.*, 2001). Alanine-scanning mutations identified that a DpYR motif was essential for the ubiquitination and downregulation of Met by c-Cbl (Peschard *et al.*, 2004). This motif is conserved among the Met family members, Ron and Sea (Penengo *et al.*, 2003), as well as in plexins—receptors for semaphorins that promote cell repulsion (Tamagnone *et al.*, 1999).

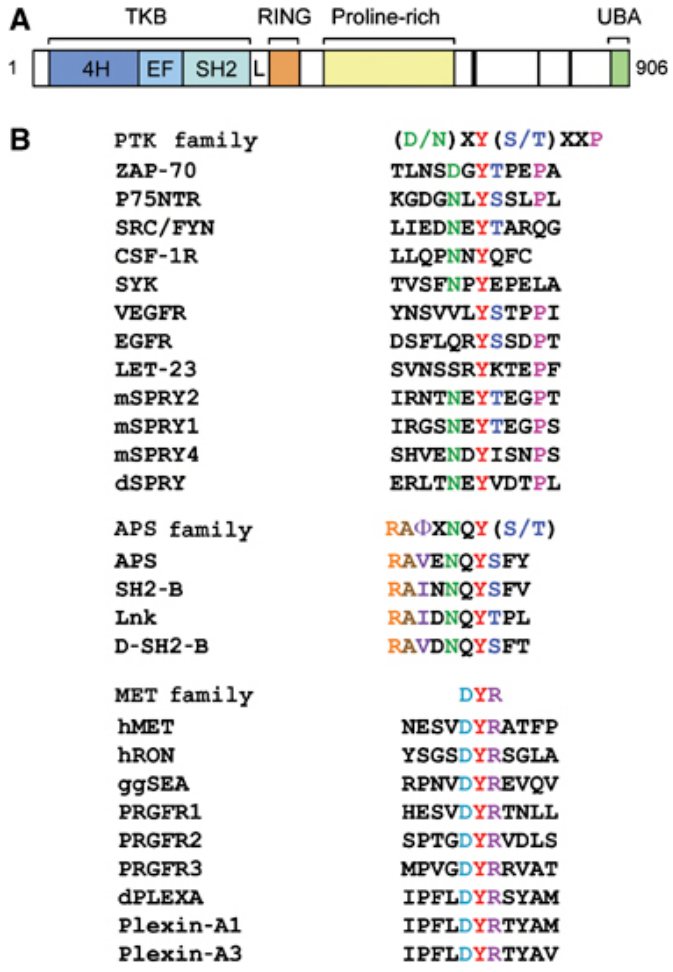


Fig. 1.14. Schematic representation of c-Cbl domain architecture and targets of its TKB domain (Ng *et al.*, 2008). (A) Structural organization of human c-Cbl. The N-terminal region contains the TKB comprising 4H, EF-hand (EF) and SH-2-like subdomains. The TKB domain is separated from the RING domain by a short linker (L) region. Cbl has a proline-rich stretch following the RING domain and a ubiquitin-associated (UBA) domain at its C terminus. (B) Sequence alignment of previously characterized c-Cbl-TKB binding sites for the PTK, APS and Met subgroups of proteins. Coloured residues refer to the conserved residues in the sequence: pY is coloured red, (pY-2)D/N is green, (pY+1)S/T is dark blue, (pY+4)P is fuschia. (pY-6)R in the APS family is orange, (pY-5)A is brown and the (pY-4)Φ represents a hydrophobic residue, coloured violet. pY flanking residues of the Met family (pY-1)D are coloured light blue and (pY+1)R magenta.

1.8 Case study 2: CaM regulation of the VGSC receptors

Besides with the voltage-gated potassium channel whose repolarization properties are modified by ERK2 phosphorylation, the voltage-dependent sodium channel (VGSC) is the other basic component of action potentials in many excitable cells. There are nine isoforms of the VGSC, Nav1.1 - Nav1.9, all of which generate the upstroke or depolarization of the action potential in various tissues (Fig. 1.5). The α-subunit of VGSC is the main component of the

channel responsible for voltage-sensitive gating and selective ion permeation (Catterall, 1995), and the function of this α-subunit can be modulated through its interaction with associated proteins in regulating the generation of action potentials.

Through sequence analyses, the C-termini of all VGSC subtypes were found to possess a conserved region for CaM interaction. As Ca^{2+} has a fundamental role in the coupling of cardiac myocyte excitation and contraction, the first evidence of CaM modulating the VGSC came from the human cardiac sodium channel subtype Nav1.5, which binds to CaM in a calcium dependent manner. This binding significantly enhances slow inactivation. Mutations that disrupt CaM binding eliminated Ca^{2+}/CaM slow inactivation (Tan *et al.*, 2002). Subsequent studies by Herzog *et al.* (2002) revealed that skeletal muscle Nav1.4 and neuronal subtype Nav1.6 are high affinity binders of CaM in the presence and absence of Ca^{2+}. The functional expression of Nav1.4 is critically dependent on the ability of the channel to bind CaM at the C-terminus, while the VGSC in *Paramecium* is dependent on CaM for activation (Saimi and Ling, 1995). Mutation of the core amino acids in the CaM binding region of Nav1.6 decreased current amplitude significantly, albeit to a smaller extent than Nav1.4. Changes in the intracellular concentration of Ca^{2+} altered inactivation kinetics of Nav1.6 currents via a CaM-dependent mechanism. Ca^{2+}/CaM slowed Nav1.6 channel inactivation by 50%, while calcium free CaM enhances the rate of inactivation. The Ca^{2+}/CaM dependent slowing of Nav1.6 inactivation kinetics could prolong action potential duration to enhance neurotransmitter release at nerve endings, and play a role in synaptic plasticity (Herzog *et al*., 2003).

1.8.1 The importance of the IQ motif in VGSC regulation

CaM binding proteins possess a basic often amphipathic stretch of residues that assumes an α-helix upon binding with CaM. These motifs contain critical hydrophobic residues at positions 1, 8 and 14 and frequently an additional hydrophobic contact at position 5. Basic amino acids are distributed throughout the motif and often flank the critical hydrophobic residues. The consensus sequence of CaM binding motif is IQxxxRGxxxR and referred to as the IQ motif, after the first two conserved residues.

Alignment of all known IQ motifs show that the residue at position 1 that may be Ile, Leu or Val. Position 2 is always Gln, 7 may be ambiguous for several residues and the 11th position can be either Arg or Lys (Table 1.2). The sequence (FILV)Qxxx(RK)Gxxx(RK)xx(FILVWY) thus represents a more generalised core IQ motif. Two or three partially conserved Ala residues may be located at the N-terminal of the primary motif. The entire region of the IQ domain comprising of core and flanking residues is 20-25 residues with no proline to form an uninterrupted seven-turn α-helix. Partial amphipathicity with a net positive charge usually between 2+ to 5+ is commonly observed, although not neccessary.

Most IQ motifs show preferential binding when Ca^{2+} is present, however, in other cases, proteins bind both in the presence and absence of Ca^{2+}. It has been suggested that IQ motifs containing a conserved Gly at the seventh position and a second basic residue do not require calcium to bind calmodulin, while binding of incomplete IQ motifs lacking the second basic residue is calcium dependent (Houdusse and Cohen,1995; Munshi *et al.*, 1996). The substitution of the 7th glycine with residues possessing bulky side chains, which occurs in approximately 50%

of IQ domains, may provide some specificity for calmodulin interactions, as might the identity of the 6 intervening amino acids in the IQ consensus sequence.

Table 1.2.
IQ motifs of established and potential CaM target proteins

Neuronal		|| | |	
Neuromodulin, bovine	ATK	IQASFRGHITRKKL	KGE
Neuromodulin IGLOO, *Drosophila*	ATK	IQAVFRGHKVRETM	KHL
	ALK	IQSTFRGHLARKLV	NKD
	ATK	IQASFRGHKTRKDA	NPE
Neurogranin, bovine	AAK	IQASFRGHMARKKI	KSG
PEP 19, human	AVA	IQSQFRKFQKKKAG	SQS
Conventional Myosins (II)			
Cardiac Myosin Beta, human	ITR	IQAQSRGVLARMEY	KKL
	LLV	IQWNIRAFMGVKNW	PWM
Cardiac Myosin Alpha , rat	ITR	IQAQARGQLMRIEF	KKM
	LLV	IQMNIRAFMGVKNW	PWM
L Type and Capacitative Calcium Channels			
Alpha-1C Subunit, human	TFL	IQEYFRKFKKRKEQ	GLV
Alpha-1C Subunit, mouse	TFL	IQEYFRKFKKRKEQ	GLV
Trp4, Bovine	ERS	IQLESRTLASRGDL	NIP
Dyneins			
Dynein-like protein, rat	VKS	IQDAIRDKKQRFSF	LGE
Dynein, D. melanogaster	VKS	VQDAIRDKKDKFNF	MGE
Plant Cyclic NMP-Regulated Ion Channel			
Arabidopsis	ACF	IQAAWRRYIKKKLE	ESL
Tobacco	ACF	IQAAWRRHCRKKLE	ESL
Plant Ethylene-Inducible CaM-Binding Proteins			
CaM and DNA Binding Protein, *Arabidopsis*	AIR	IQNKFRGYKGRKDY	LIT
	IIK	IQAHVRGYQFRKNY	RKI
ER66 Protein, tomato	AVR	IQNKLRSWKGRRDF	LLI
	IIK	IQAHVRGHQVRNKY	KNI
GTPase-activating (IQGAP) protein			
CDC42/Rac1/GTPase activating protein 1, mouse	ITK	LQACCRGYLVRQEF	RSR
	ITC	IQSQWRGYKQKKAY	QDR
	VVK	IQSLARMHQARKRY	RDR
	IIK	IQAFIRANKARDDY	KTL
Ras GRF 1 and 2			
Ras GRF-1/CDC25, human	IKK	VQSFLRGWLCRRKW	KTI
Ras GRF-2, mouse	IKK	VQSFMRGWLCRRKW	KTI
Miscellaneous Proteins			
CVT, *Branchiostoma*	ATR	IQASFRMHKNRMAL	KEK
Sperm surface protein, Sp17, human	AVK	IQAAFRGHIAREEA	KKM
Utrophin, human (one of two motifs)	EDV	LQKEVRVKILKDNI	KLL
Sodium Channel, human	AVI	IQRAYRRHLLKRTV	KQA
Inversin, human	ELR	LQIIQRERRRKELF	RKK
	AAV	IQRAWRSYQLRKHL	HLR
β-Spectrin, mouse	LFT	IQSRMRANNQKVYT	PHD
Kendrin , human (one of two motifs)	KDN	LQKELRIEHSRCEA	LLA
SIIA 1, mouse (two of four motifs)	TVV	VQALVRGWLVRKRV	SEQ
	VIF	IQRWFRKRLQRKRF	IEQ
SHA-like protein, human (one of sixteen motifs)	VIY	IQAIFRGKKARRHL	KMM
Phosphatase, PPEF-2, human	AAL	IQRWYRRYVARLEM	RRR
Retinal Degeneration Protein, D. melanogaster	AIF	IQKWYRRHQARREM	QRR
SF16 Pollen-Specific Protein, sunflower	ATK	IQAAYRGYTARRAF	RSL
Putative Ubiquitin ligase, S. pombe	SVA	VQSLSRGFLARRKF	KQD

In the crystal structures of CaM bound IQ motif complexes, the IQ domain was observed to run through the structure of CaM between N- and C-termini lobes. Exact positioning of the motif, however, differs between proteins.

The IQ motif at the C-termini of K^+ and Ca^{2+} ion channels, together with the voltage-gated sodium channels (VGSC) enables CaM to functionally modify their function (Fig. 1.15). Disruption of VGSC regulation by CaM through mutation of the IQ motif has been linked to abnormalities linked to life-threatening idiopathic ventricular arrhythmias in cardiac muscle (Veldkamp *et al.*, 2000; Wang *et al.*, 2000; Tan *et al.*, 2002). Similarly, a proposed link between increased Ca^{2+} concentration, CaM kinase II activity and myotonia suggests that skeletal muscle VGSCs may be implicated (Deschenes *et al.*, 2002). A study of an individual with autism indicated that a mutation identified in the C-terminus of human Nav1.2 might reduce the binding affinity of Nav1.2 for Ca^{2+}/CaM (Weiss *et al.*, 2003). Although the physiological relevance of altered inactivation kinetics in other VGSCs caused by disrupting CaM binding has yet to be ascertained, it remains conclusive that Ca^{2+} plays an important role through CaM-IQ motif interaction and regulates VGSCs in an isoform-specific way.

Fig. 1.15. Sequence alignment of the CaM interacting motif (IQ motif) within the C-termini of VGSC isoforms Herzog *et al.*, 2003). Identical and highly conserved residues are shaded in black and grey respectively from a schematic GST-fusion protein with the 250 amino acid long C-terminus of a sodium channel, indicating the position of the highly conserved IQ motif.

1.9 Chapter summary

The ERK/MAPK pathway is a ubiquitous serine/threonine kinase cascade directing growth, differentiation and plasticity in various tissues. It is initiated by multiple receptor inputs that converge at levels downstream of Raf through to ERK1/2. Regulating receptor activity instead of the MAP kinase module has certain advantages. Firstly, Raf, MEKs and ERKs are entities common to all ERK/MAPK pathways: regulating upstream events, such as at the receptor level, enables signalling from a certain receptors to be specifically controlled without affecting signals from other receptors. Secondly, it allows an unambiguous attenuation of an unwanted signal at its source.

Like all signalling proteins, receptors leading to the ERK/MAPK pathway require interaction with other factors be functional. Evolution has thus selected for the incorporation of molecular recognition domains within these proteins. Regulatory proteins also exploit these domains in order to latch on to substrates for their purposes. With the examples of regulatory proteins c-Cbl, Sprouty and CaM, this thesis explores the effects of regulatory strategies such as post-translational modifications, competitive inhibition and functional modifiers. Through the use of three unique classes of receptors (RTKs – receptors with enzymatic activity, NRTKs – intracellular receptors and VGSCs – ion channel receptors), this thesis seeks to address the importance of protein interaction domains in the regulation and modulation of ERK/MAPK signalling across all receptor types.

In the next chapter (Chapter II), the implication of protein phosphorylation in MAPK signal regulation by c-Cbl TKB domain is emphasized. Together with biophysical and molecular quantization on the contribution of certain residues to its binding capacity, the basis of substrate

alignment on the TKB domain is thoroughly described. The third chapter describes a calcium-dependent regulation of the voltage-gated sodium channel by calmodulin and its implication in the neuronal MAPK processes of learning and memory, an interaction mediated by the IQ motif. Both the Cbl SH2 domain and the calmodulin interacting IQ motif represent domains involved in signal regulation that are seemingly able to bind their partners in multiple ways – Cbl has an abnormally wide range of binding motifs, and IQ motifs complex with CaM in many orientations as observed from previously known structures.

Through structural characterisation of protein-peptide complexes, the aim of this thesis is to explore whether common interaction mechanisms that govern the regulatory networks of Cbl and CaM with their respective, seemingly disparate, binding motifs exist, such that they can retain their intrinsic flexibility of binding, and yet still evoke specific responses.

Chapter II

c-Cbl and Protein Tyrosine Kinase Signalling

2.1 Introduction

Although the approximatey 120 documented SH2 domains use diverse strategies to aggregate signalling complexes, there is no precedent for a given SH2 or PTB domain to associate with such apparently unrelated consensus binding motifs as seen with the TKB domain. Cbl proteins appear to have two major physiological functions—either in acting as scaffolds or in targeting proteins for ubiquitination (Thien and Langdon, 2005). These functions are reflected in the purpose of the TKB domain in (a) recognizing proteins and, in the case of ubiquitin substrates, (b) ensuring its appropriate orientation for ubiquitin conjugation. Cbl proteins have critical roles in the downregulation of both RTKs in growth and differentiation and the antigen receptors in the control of immune responses, and thus elucidating the TKB recognition sequence has been central to understanding the biochemistry of tyrosine kinase downregulation by Cbl proteins. The TKB domain is, in essence, an 'embedded' SH2 domain—a domain generally recognized to bind a 'tight' consensus motif (Liu *et al.,* 2006). The abnormal range of the experimentally observed motifs that bind to the TKB domain invites the question as to whether this domain has an uncharacteristic flexibility of binding or whether there is a hitherto undiscovered common binding mechanism that encompasses the previously observed sequences. Furthermore, with such diverse sequence recognition strategies, binding affinity to c-Cbl is likely to be varied and may lead to the selective sequestration of c-Cbl by higher-affinity targets over lower-affinity ones. This model has been suggested as a mechanism whereby Spry2, the most well-characterized of the Spry proteins, recruits c-Cbl from EGFR and results in the ubiquitination of Spry2 and preservation of EGFR on the cell surface (Wong *et al.,* 2002; Hall *et al.,* 2003; Rubin *et al.,* 2003).

2.2 Materials and methods

2.2.1 Plasmid constructs, cloning, expression and purification

The human c-Cbl-TKB domain (residues 25–351; Cbl-TKB) was cloned into pGEX4T-1 expression vector (GE Healthcare, Buckinghamshire, UK), overexpressed as a glutathione-*S*-transferase fusion protein and purified according to Meng *et al.,* (1999) with modifications. The purified protein (1 mg/ml) was kept in a storage buffer (20 mM Na-HEPES pH 7.0, 0.2 M NaCl) and used for crystallization and ITC experiments. Wild-type full-length constructs of FLAG-tagged Spry2, Spry4, FGFR1, HA-tagged c-Cbl and EGFR have been described previously (Wong *et al.,* 2001). pBABE-bleo human insulin receptor B (IR-B) plasmid 11211 was from Addgene (deposited by Dr C Ronald Kahn; Entingh *et al.,* 2003). Rat myc-tagged APS was kindly provided by Dr W Langdon (University of Western Australia). c-Met was from Origene Technologies Inc. (Rockville, MD) and subcloned into PXJ40 vector (courtesy of Dr E Manser, Institute of Molecular and Cell Biology) using *Not*I/*Not*I restriction sites. Point mutations in FLAG–Spry2 and HA–Cbl were generated by site-directed mutagenesis using *Pfu* DNA polymerase from Promega (Madison, WI).

2.2.2 Complex formation and crystallization

Tyrosine phosphorylated peptides of Human Sprouty2 49–61 ($Spry2^{49–61}$), Sprouty4 69–81 ($Spry4^{69–81}$), EGFR 1063–1075 ($EGFR^{1063–1075}$), Syk tyrosine kinase 317–329 ($Syk^{317–329}$) and Met receptor 997–1009 ($Met^{997–1009}$), (Sigma Genosys) were reconstituted in Cbl-TKB storage buffer and incubated with the purified Cbl-TKB in two-fold molar excess at room temperature

for 1 h for complex formation. The Cbl-TKB:peptide complexes were concentrated to 5 mg/ml in Amicon Ultra ultrafiltration devices (Millipore, Billerica, MA). Complexes were screened for crystals against crystal screen kits from Hampton Research, CA and Jena Biosciences, DE using hanging drop and under oil vapour diffusion methods respectively. Initial crystallization conditions were optimised by varying precipitant or additive concentrations.

2.2.3 Data collection, structure determination and refinement

Data for the Cbl-TKB:EGFR$^{1063–1075}$ complex crystals were collected with an R-axis IV^{++} image plate detector mounted on a RU-H3RHB rotating anode generator (Rigaku Corp., Tokyo, Japan). Data for all other complex crystals were collected at X29 beamline in the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY). Data were processed and scaled with HKL2000 (Otwinowski and Minor, 1997). The TKB domain of c-Cbl from the ZAP70:Cbl complex structure, PDB code 2CBL (Meng *et al.*, 1999) was used as a starting model for molecular replacement solution using the programme Molrep (Vagin and Teplyakov, 1997). Subsequent rigid body refinement reduced the R-factors close to 30%. At this stage, the calculated difference electron density map clearly showed the presence of substrate peptides. Model building and refinement were carried out in O (Jones *et al.*, 1991) and CNS (Brunger *et al.*, 1998) programmes, respectively. After several cycles of map fitting and refinement the R-values were converged.

2.2.4 Protein Data Bank accession code

Coordinates and structure factors of all of the Cbl-TKB:phosphopeptide complexes have been deposited with RCSB Protein Data Bank with codes 3BUM (Spry2), 3BUN (Spry4), 3BUO (EGFR), 3BUW (Syk) and 3BUX (Met).

2.2.5 Isothermal titration calorimetry (ITC)

Both phosphorylated and unphosphorylated peptides of Spry2$^{49–61}$, Spry4$^{69–81}$, EGFR$^{1063–1075}$, Src tyrosine kinase (Src$^{413–425}$), Syk$^{317–329}$, VEGFR (VEGFR1$^{1327–1338}$), VEGFR1$^{1327–1338}$ V1331N mutant, VEGFR1$^{1327–1338}$ L1332R mutant, c-Met$^{997–1009}$ and macrophage stimulating 1 receptor 1011–1023 (Ron$^{1011–1023}$) were titrated against 10.8μM Cbl-TKB in a VP-ITC microcalorimeter (Microcal, Northampton, UK) performed under identical conditions. All peptides were maintained in storage buffer at 22°C and titrated in 10 μl injections into Cbl-TKB also in storage buffer. The heat of dilution, determined by titrating peptide into storage buffer, was subtracted from the raw data before curve fitting and refinement. The dissociation constants of Cbl-TKB to the various peptides were determined by least squares method and the binding isotherm was fitted using Origin v7.0 (Microcal) assuming a single-site binding model. All measurements were repeated at least twice for verification.

2.2.6 Antibodies and reagents

Rabbit anti-Cbl (C-15), anti-FGFR1 (C-15), anti-Myc (A-14), anti-EGFR (1005), anti-Grb2 (C-23) anti-insulin Rβ (C-19), and mouse anti-Myc (9E10) and anti-EGFR (528) were from Santa Cruz Biotechnology (Santa Cruz, CA). Mouse anti-c-Cbl and anti-Grb2 were from Transduction Laboratories (Lexington, KY). Mouse anti-Met (25H2) was from Cell Signaling Technology (Beverly, MA) and rat anti-HA was from Roche Diagnostics (Basel, Switzerland). Mouse anti-FLAG, rabbit anti-HA, goat anti-HGFR, agarose-conjugated anti-FLAG M2 beads, recombinant human (rh) FGF-2, and goat anti-rabbit-HRP and anti-mouse-HRP were from Sigma Aldrich (St Louis, MO). rhEGF was from Upstate Biotechnology Inc. (Lake Placid, NY) and rhHGF was from Calbiochem (EMD Chemicals, San Diego, CA).

2.2.7 Cell culture and transfection

Human embryonic kidney 293T cells (ATCC, Manassas, VA) were cultured, maintained and transfected as previously outlined (Wong *et al.,* 2001; Yusoff *et al.,* 2002). For binding between WT c-Cbl and Spry2 point mutants, cells were transfected with FLAG–Spry2 and FGFR1 for 16 h and precipitated for endogenous c-Cbl. Cells were transfected with WT c-Cbl or Cbl bearing various point mutations with WT Spry2, Spry4, Met, EGFR or APS with IR-B for 16 h or 40 h (in the case of Met). For Spry2, Spry4 or Grb2 experiments, cells were stimulated with 20 ng/ml FGF-2 for 2 h. Met-transfected cells were stimulated with 100 ng/ml of HGF 30 min. EGFR-transfected cells were stimulated with 100 ng/ml of EGF for 10 min. For APS experiments, cells were stimulated as described previously (Hu and Hubbard, 2005).

2.2.8 Cell lysis, immunoprecipitation and western blotting

Cells lysis and western blotting was performed as described previously (Wong *et al.*, 2001). Immunoprecipitation was carried out using FLAG–M2 beads (Spry2), anti-HA (c-Cbl), anti-HGFR (Met), anti-myc (APS) or anti-EGFR (528).

2.3 Results

2.3.1 Purification of recombinant Cbl

The human c-Cbl-TKB domain (residues 25–351; Cbl-TKB) was cloned into pGEX4T-1 plasmid vector (GE Healthcare, Buckinghamshire, UK) and transformed into *Escherichia coli* expression strain Rosetta. 100ml of Luria-Bertani (LB) media supplemented with 100µg/ml ampicillin was inoculated with one colony of Rosetta(pGEX4T1-Cbl-TKB) and grown overnight at 37ºC with aeration. The overnight culture was amplified to 1L medium and protein expression induced with 0.1mM isopropyl-thio-galactoside (IPTG) for 16h at 20ºC during mid-log phase of cellular growth.

GST-fused Cbl-TKB was harvested by six cycles of 1min sonication with 40sec interval between cycles. Cbl-TKB in the cleared lysate was purified and cleaved on GST-sepharose (GE Healthcare, Buckinghamshire, UK) as described by Meng *et al.*, 1999. Cbl-TKB was purified to homogeneity by gel filtration using Superdex G200 column (GE Healthcare, Buckinghamshire, UK), eluting at 80ml as a single symmetrical peak (Fig. 2.2) and migrating as a single band in 12.5% denaturing polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 2.1).

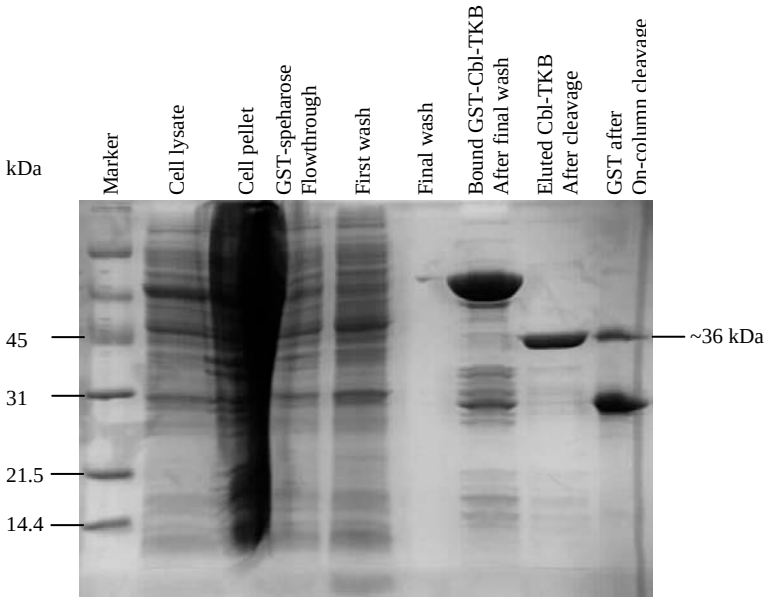


Fig. 2.1. Purification profile of Cbl-TKB. Samples from different stages of purification were collected, separated by 12.5% SDS-PAGE and visualized by Coomassie staining. GST fused Cbl-TKB was bound to GST-Sepharose and released from column by thrombin cleavage. Eluted Cbl-TKB migrated as an approximately 36kDa band.

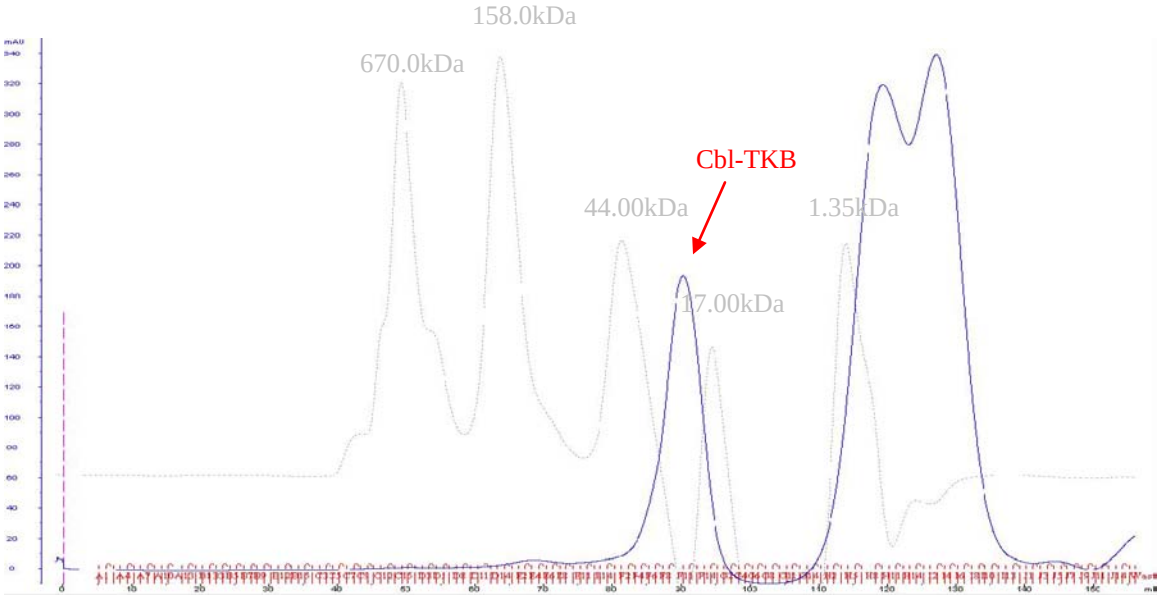


Fig. 2.2. Gel filtration profile of cleaved Cbl-TKB in a 120ml column volume Superdex-200 column. Cleaved Cbl-TKB was purified to homogeneity by gel filtration and eluted in storage buffer. Blue plot indicate Cbl-TKB gel filtration profile, grey line is gel filtration molecular size standard with molecular weight of the peaks labeled (BioRad, CA).

2.3.2 Verification of purified Cbl-TKB functional properties

Purified Cbl-TKB was incubated with increasing amounts of phosphorylated Spry2$^{49–61}$ for 1h on ice. Glutaraldehyde, a cross-linking agent whose aldehyde groups that reacts with nitrogen atoms of amino acids in close proximity, was added with stirring to a final concentration of 0.1% after 1h of complex formation. The covalently linked Cbl-TKB: Spry2$^{49–61}$ was separated on 12.5% SDS-PAGE.

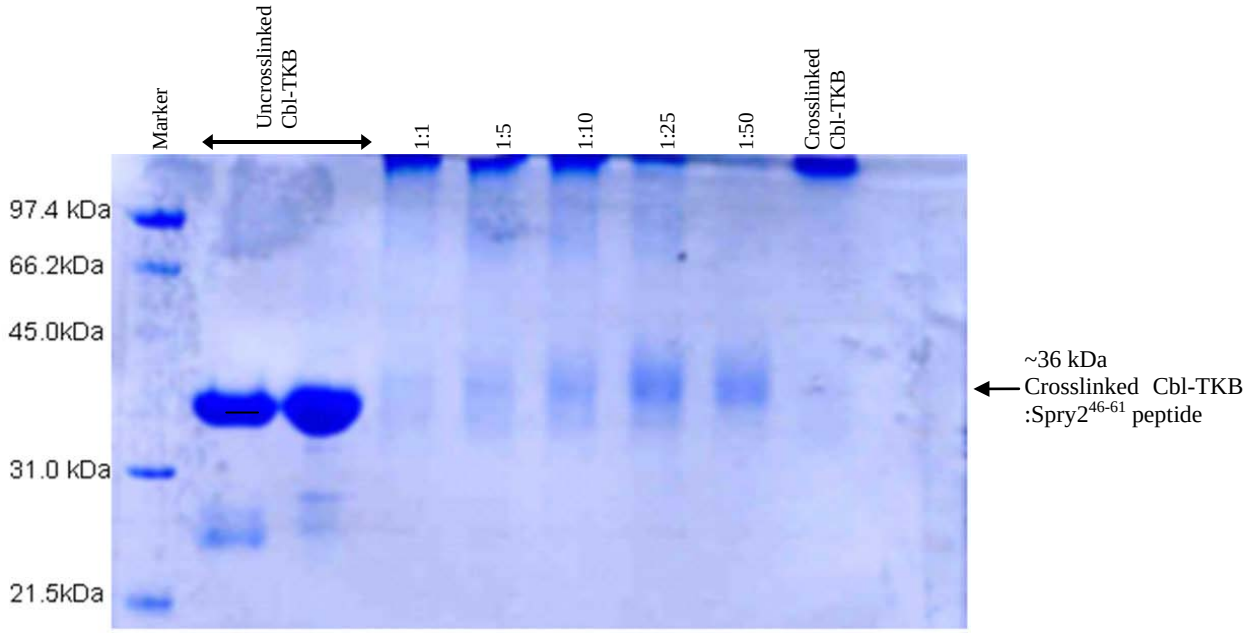


Fig. 2.3. Glutaraldehyde cross-linking of Cbl-TKB with various molar ratios of phosphorylated Spry2$^{46-61}$. Uncrosslinked Cbl-TKB (without adding glutaraldehyde) and uncomplexed, crosslinked Cbl-TKB (without Spry2$^{46-61}$ peptide) were negative controls, while Cbl-TKB was complexed with Spry2$^{46-61}$ in various molar ratios as indicated above each lane.

At low peptide concentrations, cross-linked Cbl-TKB was observed to form high molecular weight complexes that were unable to move into the separating gel, possibly due to the non-specific oligomerization of Cbl-TKB molecules. With increasing peptide concentration,

Cbl-TKB: Spry2$^{49–61}$ was increasingly observed to migrate as a band larger than the molecular weight of uncomplexed Cbl-TKB (Fig. 2.3). This observation confirms that the structure of purified recombinant Cbl-TKB was preserved, since ligand binding requires all three sub-domains of the TKB domain to be structurally intact (Meng *et al.,* 1999).

2.3.3 Crystallization

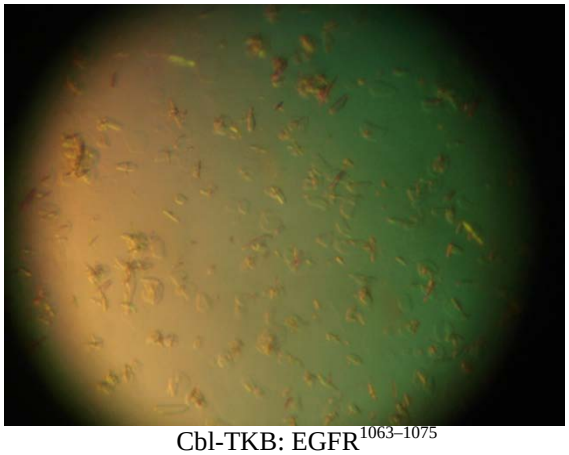
The ability of the TKB domain to bind to two different consensus motifs—(D/N)XpY(S/T)XXP or DpYR—but still appropriately align the protein for ubiquitination suggests that there would be a common binding mechanism. The motifs on Spry2, EGFR, Syk and Met that recognize c-Cbl-TKB domain (herein referred to as Cbl-TKB) have previously been identified and, as Spry4 has three of the four conserved residues, it was also explored whether Spry4 binds to c-Cbl, as results for this are currently inconclusive. Cbl-TKB was therefore concentrated to 5mg/ml and incubated with excess of phosphorylated Spry2$^{49–61}$, Spry4$^{69–81}$, EGFR$^{1063–1075}$, Syk$^{317–329}$, and Met$^{997–1009}$ peptides. Complexes were screened for crystals in 1µl protein to 1µl mother liquor hanging drops with 500µl reservoir solution.

The initial crystals obtained were too small, too flat or twinned to be of diffraction quality (Fig. 2.4). These initial conditions were optimized and finally each of the complexes crystallized in a different condition. Cbl-TKB:Spry2$^{49–61}$ complex crystals grew in 0.1M Na-HEPES pH 7.5, 75mM ammonium sulphate, 25% PEG3350 and 4% n-propanol; Cbl-TKB:Spry4$^{69–81}$ complex crystals grew in 0.2M Na/K tartrate and 20% PEG3350; Cbl-TKB:EGFR$^{1063–1075}$ crystals grew in 0.25M sodium formate and 20% PEG3350; Cbl-TKB:Syk$^{317–329}$ crystals grew in 0.1M Bis-Tris

pH 6.5, 0.12M ammonium acetate and 18% PEG3350; Cbl-TKB:Met$^{997–1009}$ crystals grew in 0.15M malic acid and 20% PEG3350 (Fig 2.5).

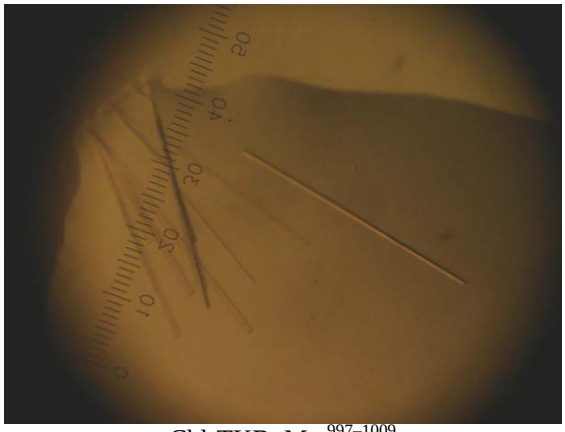
Cbl-TKB: Spry2$^{49–61}$

Cbl-TKB: Spry4$^{69–81}$



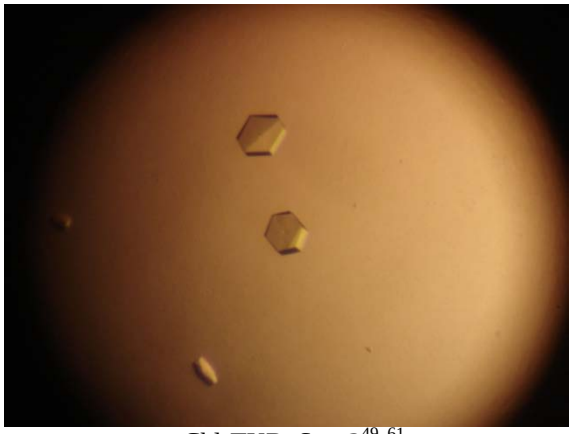
Cbl-TKB: EGFR$^{1063–1075}$

Cbl-TKB: Syk$^{317–329}$

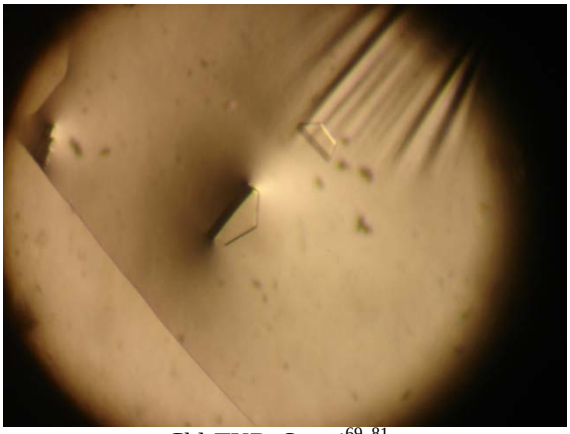


Cbl-TKB: Met$^{997–1009}$

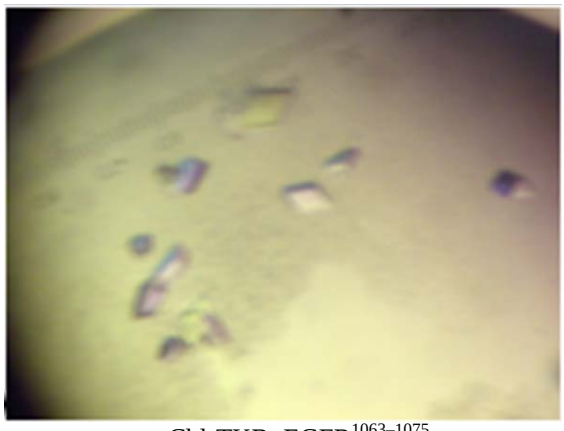
Fig. 2.4. Initial crystals from the screening of Cbl-TKB complexed with phosphorylated peptides obtained using hanging drop vapour diffusion method. These crystals were obtained from conditions in the Hampton Research Index Screen Kit (Hampton Research, CA).



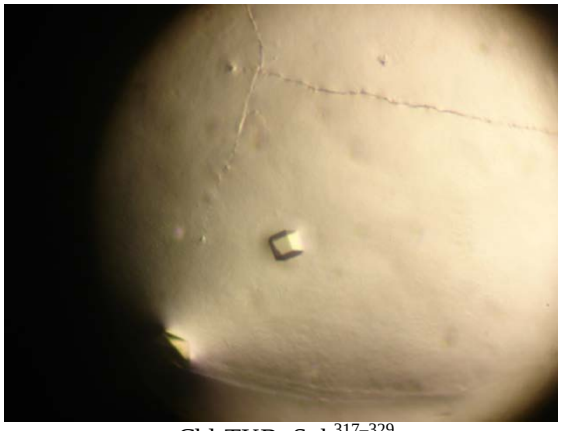
Cbl-TKB: Spry2$^{49–61}$



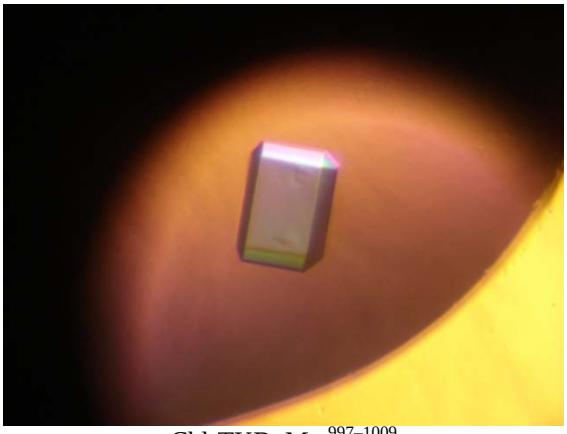
Cbl-TKB: Spry4$^{69–81}$



Cbl-TKB: EGFR$^{1063–1075}$



Cbl-TKB: Syk$^{317–329}$



Cbl-TKB: $Met^{997–1009}$

Fig. 2.5. Diffraction quality crystals of Cbl-TKB complexed with phosphorylated peptides obtained after initial crystals optimization using hanging drop vapour diffusion method. These crystals were cryo-protected by dehydration method using mother liquor supplemented with increasing concentrations of glycerol and flash frozen in liquid nitrogen before data was collected.

2.3.4 Data collection, structure determination and refinement

Except Cbl-TKB:$EGFR^{1063–1075}$, all other complex crystals were diffracted at the X29 beamline in the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY). The structures of the complex crystals between Cbl-TKB and the phosphorylated 13-mer peptides of these five proteins were determined between 1.35–2.60Å resolutions and crystallized in two different space groups (Table 2.1).

Table 2.1. Data collection and refinement statistics.

Data collection	Cbl-N Spry2	Cbl-N Spry4	Cbl-N EGFR[a]	Cbl-N Syk	Cbl-N Met
Resolution (Å)	50.00 - 1.98 (2.05-1.98)	50.00-2.00 (2.07-2.00)	50.00 - 2.60 (2.69-2.60)	50.00-1.45 (1.50-1.45)	50.00 – 1.35 (1.40-1.35)
Space group	P6	P6	$P2_1$	$P2_1$	$P2_1$
Unit cell (Å)	a= b= 122.26 c= 54.71	a= b= 122.85 c= 55.54	a= 63.86 b= 110.17 c= 55.82 β= 89.94°	a=63.79 b=104.81 c=52.78 β= 89.83°	a= 63.18 b= 104.80 c= 52.60 β= 90.41°
Molecules / a.s.u.	1	1	2	2	2
Reflections (total/unique)	528974/32750	566291/31352	59344/21945	181855/36701	930080/142157
R_{sym} (%)[b]	8.3 (34.6)	5.2 (15.1)	10.9 (33.5)	7.6 (25.4)	10.9 (23.6)
Completeness (%)[c]	99.9 (99.1)	96.3 (82.1)	91.9 (83.0)	92.7 (66.0)	95.0 (87.2)
Refinement[d]					
Resolution (Å)	20.00 – 2.00	20.00-2.00	20.00 - 2.60	20.00-1.45	20.00 – 1.35
R_{work} (%)[e]	22.40(29469)	20.13(29166)	23.04(15455)	22.41(95507)	22.52(127952)
R_{free} (%)[f]	26.24(2221)	24.24(2128)	27.48(2304)	23.84(2176)	23.98(2208)
Average B-factors (Å^2) (no. of atoms)					
Cbl-TKB	29.71(2489)	40.11(2489)	33.65(4980)	21.79(4994)	17.21(4994)
Peptide	25.94 (74)	48.82(63)	35.29(190)	33.02(162)	22.76(140)
Water	42.86(297)	53.97(355)	32.18(235)	30.71(575)	26.66(686)
r.m.s.d bond lengths (Å)	0.009	0.010	0.009	0.006	0.006
r.m.s.d bond angles (°)	1.24	1.36	1.32	1.10	1.11
Ramachandran plot[g]					
Most favoured regions (%)	89.8	90.5	81.3	89.3	89.1
Additionally allowed regions (%)	9.8	9.1	16.9	10.0	10.1
Generously allowed regions (%)	0.4	0.4	1.8	0.7	0.7
Disallowed regions (%)	0.0	0.0	0.0	0.0	0.0

Values in parentheses indicate statistics for the highest-resolution shells.

[a]NCS restraint was kept throughout the refinement of Cbl–TKB EGFR.

[b]$R_{sym} = |I_i - \langle I \rangle| / |I_i|$ where I_i is the intensity of the *i*th measurement, and $\langle I \rangle$ is the mean intensity for that reflection.

[c]Completeness = (number of independent reflections)/(total theoretical reflections).

[d]For all models, reflections with $I > \sigma I$ was used in the refinement.

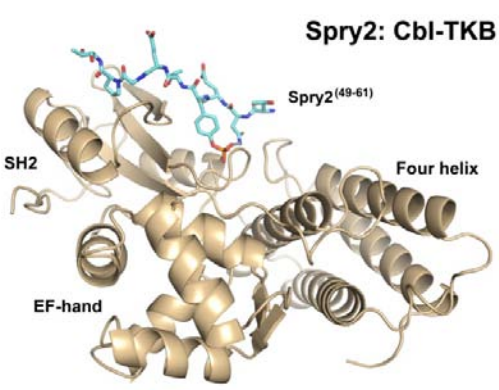
[e]R factor = $100 \times \Sigma|F_P - F_{P(calc)}| / \Sigma F_P$.

[f]R-free was calculated with approximately the same number of reflections in all the complexes' test set.

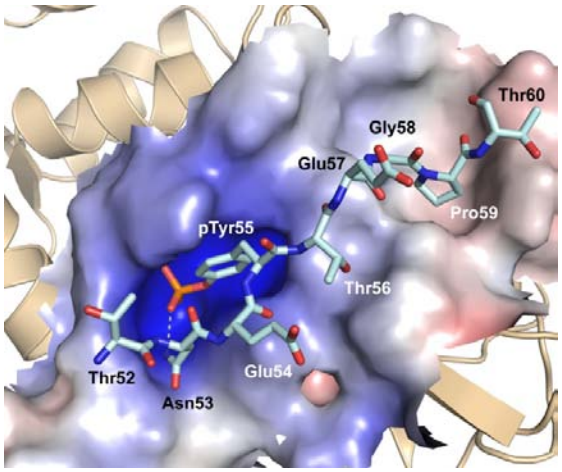
[g]Statistics for the Ramachandran plot from an analysis using PROCHECK (Laskowski et al, 1993).

Four of the five complexes resembled the binding pattern for c-Cbl with a phosphorylated peptide that has previously been identified, with several exceptions (Meng *et al.*, 1999; Hu and Hubbard, 2005) (Fig. 2.6K), whereas the Met peptide bound to Cbl-TKB in the reverse orientation, with backbone directionality extending from C- to N- terminus. This reversed binding has not been observed for the TKB domain or for any other SH2 or PTB domain to date. The Cbl-TKB:Met complex 1.35Å resolution also represents the highest-resolution structure currently available for c-Cbl. The structure of the Cbl-TKB:peptide complexes for Spry2$^{49–61}$, EGFR$^{1063–1075}$ and Met$^{997–1009}$ are shown as ribbon diagrams to illustrate the position of the peptide relative to the subdomains of the TKB domain (Fig. 2.6A, E and I). Although the main site of interaction is at the SH2 region, all three subdomains must be intact before binding occurs. In particular, the SH2 domain was found to shift and pack against the 4H upon ligand binding to complete the pTyr binding pocket (Meng *et al.*, 1999). The Cbl-TKB:peptide complexes for Spry4$^{69–81}$ and Syk$^{317–329}$ resemble the Cbl-TKB:Spry2 complex (Fig. 2.6C and G). Electrostatic surface representations for the Spry2, EGFR and Met complexes are shown in Figure 2.6B, F and J respectively and in the Figure 2.6D and H for Spry4 and Syk, demonstrating peptide binding and the surface morphology of the TKB domain. Electron density maps for each of the peptides are provided in Figure 2.7. In all of the complexes, extreme N- and C-terminal residues are absent in the electron density map and presumed to be disordered.

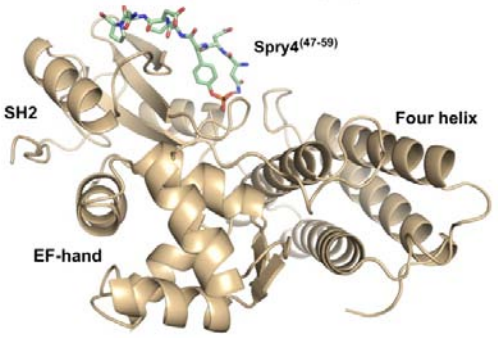
(A)


Spry2: Cbl-TKB
Spry2(49-61)
SH2
Four helix
EF-hand

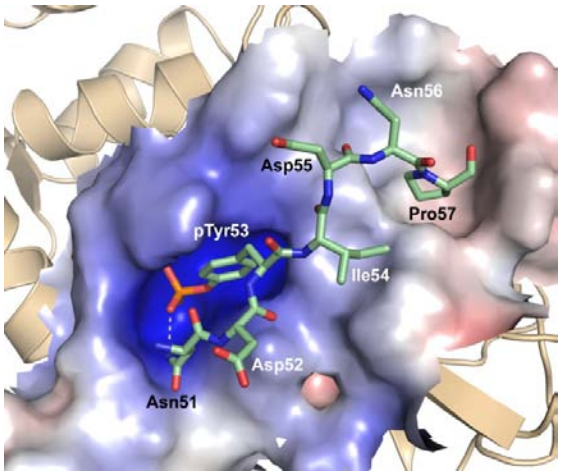
(B)


Thr60
Gly58
Glu57
Pro59
pTyr55
Thr56
Thr52
Glu54
Asn53

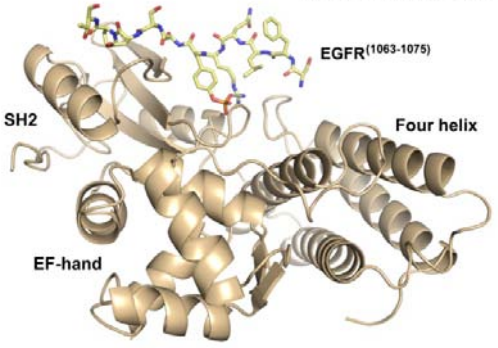
(C)


Spry4: Cbl-TKB
Spry4(47-59)
SH2
Four helix
EF-hand

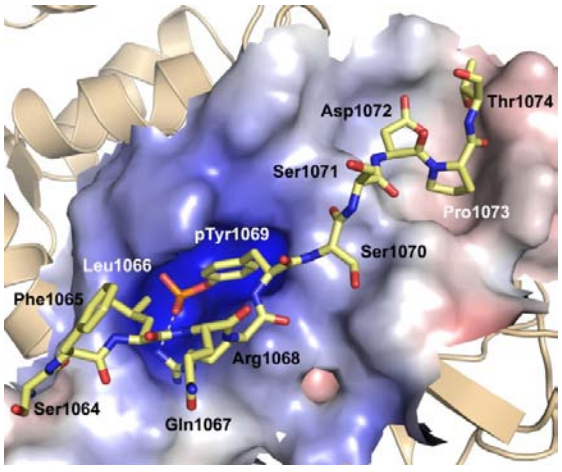
(D)


Asn56
Asp55
Pro57
pTyr53
Ile54
Asp52
Asn51

(E)


EGFR: Cbl-TKB
EGFR(1063-1075)
SH2
Four helix
EF-hand

(F)


Asp1072
Thr1074
Ser1071
Pro1073
pTyr1069
Ser1070
Leu1066
Phe1065
Arg1068
Ser1064
Gln1067

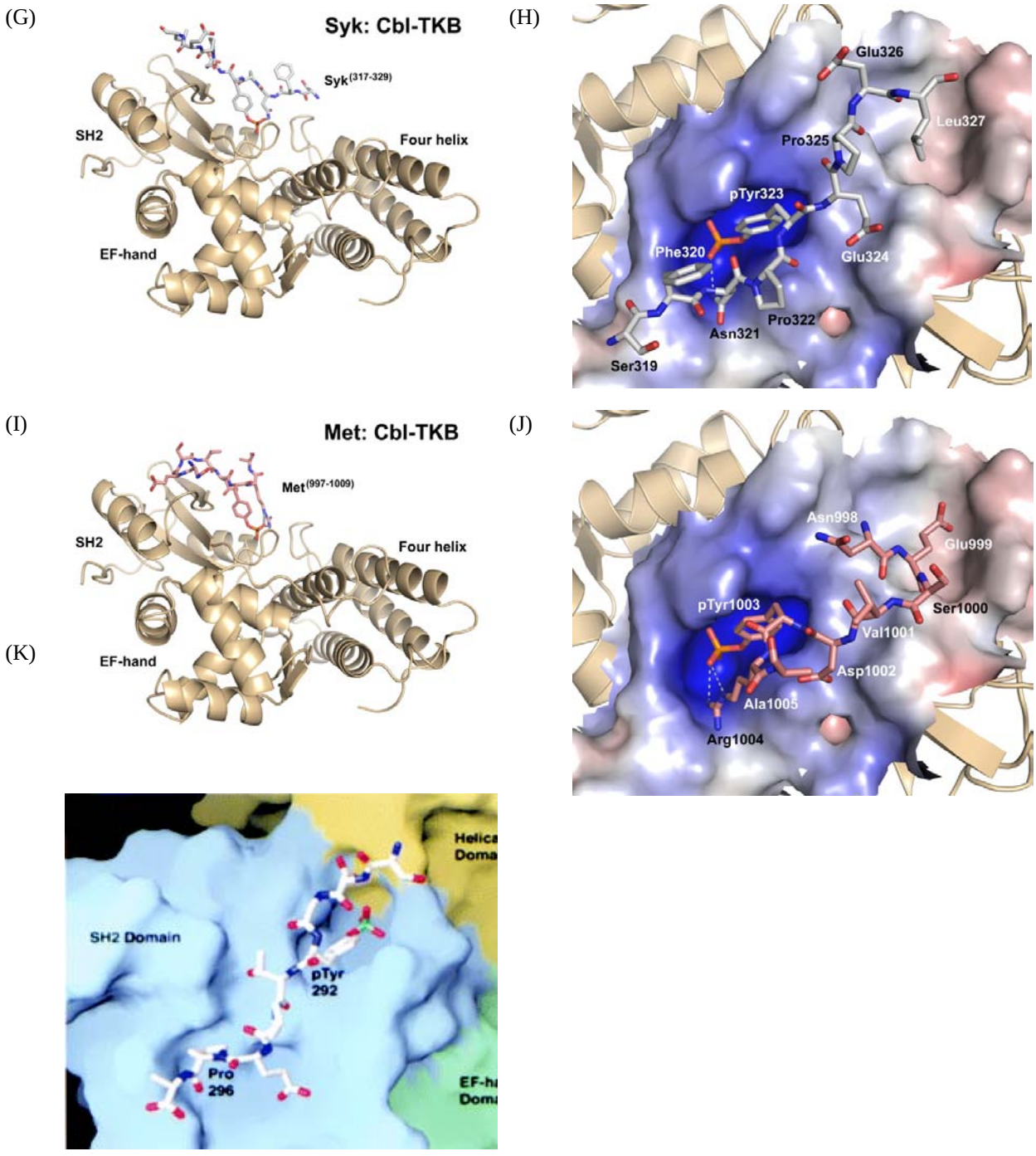


Fig. 2.6. Crystal structure and electrostatic surface representations of TKB domain complexed with Spry2, Spry4, EGFR, Syk and Met (Ng *et al.*, 2008). Crystal structures of the Cbl-TKB complexed with each of the peptides (**A**) Spry2$^{49–61}$ (cyan), (C) Spry4$^{69–81}$ (green), (**E**) EGFR$^{1063–1075}$ (yellow), (G) Syk$^{317–329}$ (grey) and (**I**) Met$^{997–1009}$ (pink), and surface representations of the Cbl-TKB complexed with (**B**) Spry2$^{49–6}$, (D) Spry4$^{69–81}$, (**F**) EGFR$^{1063–1075}$, (H) Syk$^{317–329}$, (**J**) Met$^{997–1009}$ and (K) ZAP-70 phosphopeptide (Meng *et al.*, 1999). The c-Cbl-TKB domain is shown as a ribbon diagram with helices, β-strands and loops coloured light brown. Phosphopeptides are shown as stick models. Intrapeptidyl H-bonds are shown as grey dotted lines. Residues at the extreme N and C termini of the phosphopeptides are disordered and not included. The figure was prepared using Pymol (DeLano, 2002) and the APBS plugin (Baker *et al.*, 2001).

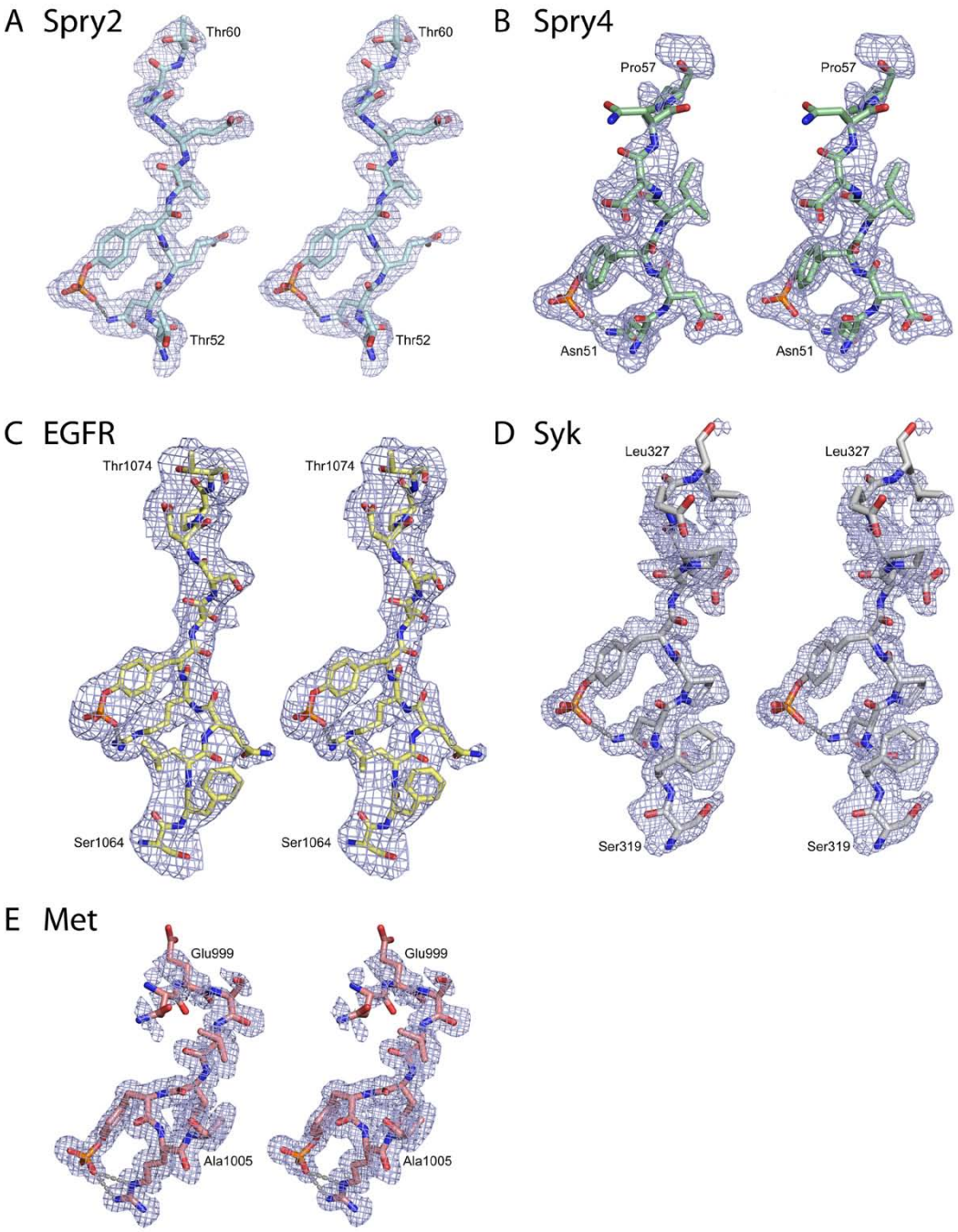


Fig. 2.7. Stereo diagrams showing omit electron density maps of (A) Spry2$^{49\text{-}61}$ (cyan), (B) Spry4$^{69–81}$ (green), (C) EGFR$^{1063\text{-}1075}$ (yellow), (D) Syk$^{317–329}$ (grey) and (E) Met$^{997\text{-}1009}$ (pink) peptides whose structures were solved in complex with Cbl-TKB (Ng *et al.*, 2008). Intra-molecular hydrogen bonds are shown as grey dotted lines. With the exception of EGFR whose map was contoured at 1.6 σ level, all other maps were contoured at level 2.0 σ and calculated phases up to their maximum diffraction resolution as stated in Table 2.1. The figure was prepared using Pymol (DeLano, 2002).

2.3.5 Conserved residues in (D/N)XpY(S/T)XXP motif contribute to binding in varying degrees

The Cbl-TKB:peptide interactions from the PTK subgroup—Spry2, Spry4, EGFR and Syk—are primarily mediated by 9–13 H-bonds (<3.2 Å) (Figure 2.8; Table 2.2). The essential pTyr residue occupies a positively charged binding pocket on Cbl-TKB and makes five to six H-bond contacts with several residues of the TKB domain (Arg294, Ser296, Cys297 and Thr298), supporting previous reports that the pTyr contributes almost 50% of the free energy of association to c-Cbl binding (Liu *et al.*, 2006). A close-up view of this interaction is shown for Spry2 and EGFR (Fig. 2.8A and C), with H-bonds represented by grey dotted lines, and an extended linear view of these interactions (Fig. 2.8B and D) with green dotted lines representing H-bonds.

The interactions mediated by the other conserved residues of the PTK subgroup were more varied. The (pY+4)Pro of Spry2, Spry4 and EGFR peptides occupies a shallow hydrophobic cleft flanked by Tyr307, Thr317 and Phe336 on Cbl-TKB. The Syk peptide lacks this interaction and binds solely to the NXpY region, possibly further stabilized by additional H-bonds with Asn79 of Cbl-TKB through its nonconserved (pY-4)Ser319, which is not observed with the other complexes. The conserved (pY+1)Ser/Thr residue of Spry2 and EGFR interacts with Gln316 of Cbl-TKB through a backbone–backbone H-bonding contact. However, this pY+1 residue was more tolerant of degeneracy than expected, as the amide N of both (pY+1)Ile54 in Spry4 and (pY+1)Glu324 in Syk also formed backbone H-bonding contacts with Gln316 of Cbl-TKB. This explains the lesser preference for conserved residues in the pY+1 position (Lupher *et al.*, 1997), although residues with large side chains are likely to have steric clashes and may not

be preferred. In all five complexes, the backbone carbonyl O of the pY-1 residue forms a H-bond with the OH of Tyr274 on Cbl-TKB, irrespective of the pY-1 residue.

The (pY-2)Asn in the peptides of the PTK subgroup formed a H-bonding contact through the N of the carboxamide side group with the carboxyl O of Pro81 on Cbl-TKB, as shown for Spry2 (Figure 2.8A and B). The (pY-2)Asn of Syk forms an additional H-bond with Ser80, which is not seen with the other peptides. Interestingly, for EGFR, which lacks the conserved (pY-2)Asn, the (pY-1)Arg appears to assume this role by forming a H-bond with the carboxyl O of Pro81 in the 4H bundle of Cbl-TKB in a manner similar to the other complexes (Figure 2.8C and D). This 'rescue'-binding mechanism of EGFR may enable it to acquire a higher binding affinity than would be expected.

Table 2.2. Hydrogen bonding interactions between Cbl-TKB and the various peptides

Cbl-TKB	Spry2$^{49-61}$	Spry4$^{47-59}$	EGFR$^{1063-1075}$	Syk$^{317-329}$	Met$^{997-1009}$
Asn79 Nδ2				Ser319 N, 3.04	
Ser80 Oγ				Asn321 Oδ1, 2.52	Arg1004 NH1, 2.62
Pro81 O	Asn53 Nδ2, 2.90	Asn51 Nδ2, 3.04	Arg1068 NH1, 2.81 Arg1068 NH2, 2.66	Asn321 Nδ2, 3.09	Arg1004 NH2, 2.85
Tyr274 OH	Glu54 O, 2.68	Asp52 O, 2.80	Arg1068 O, 2.81	Pro322 O, 2.71	Asp1002 Oδ2, 2.58
Arg294 NH1 **NH2**	 pTyr55 OH, 2.89 pTyr55 PO3, 2.91	 pTyr53 OH, 2.89	 pTyr1069 PO2, 3.03 pTyr1069 OH, 2.88	pTyr323 PO2, 3.02 pTyr323 OH, 2.93	pTyr1003 PO2, 2.97 pTyr1003 OH, 2.93
Ser296 Oγ	pTyr55 PO3, 2.76	pTyr53 PO3, 2.71	pTyr1069 PO3, 3.13	pTyr323 PO3, 2.72	pTyr1003 PO3, 2.67
Cys297 N	pTyr55 PO3, 2.98	pTyr53 PO2, 2.95		pTyr323 PO2, 2.92	pTyr1003 PO2, 3.04
Thr298 Oγ1 **N**	pTyr55 PO3, 2.55 pTyr55 PO3, 2.89	pTyr53 PO3, 2.70 pTyr53 PO3, 2.88	pTyr1069 PO3, 2.94 pTyr1069 PO3, 2.77	pTyr323 PO3, 2.56 pTyr323 PO3, 3.01	pTyr1003 PO3, 2.61 pTyr1003 PO3, 2.99
Tyr307 OH					Glu999 O, 2.53
Gln 316 O **N**	Thr56 N, 2.84 Thr56 Oγ1, 3.21	Ile54 N, 2.84	Ser1070 N, 2.75	Glu324 N, 2.85	pTyr1003 N, 2.86 Asp1002 Oδ1, 2.88
Thr317 Oγ1			Ser1070 O, 2.56		
His320 Nε		Asp55 Oδ2, 2.77			
Lys322 Nξ			Asp1072 Oδ2, 3.06		
Glu334 Oε1 ***Oε2***			Thr1074 Oγ1, 2.96		Glu999 Oε1, 3.03 Glu999 Oε2, 2.92 Glu999 Oε1, 2.84

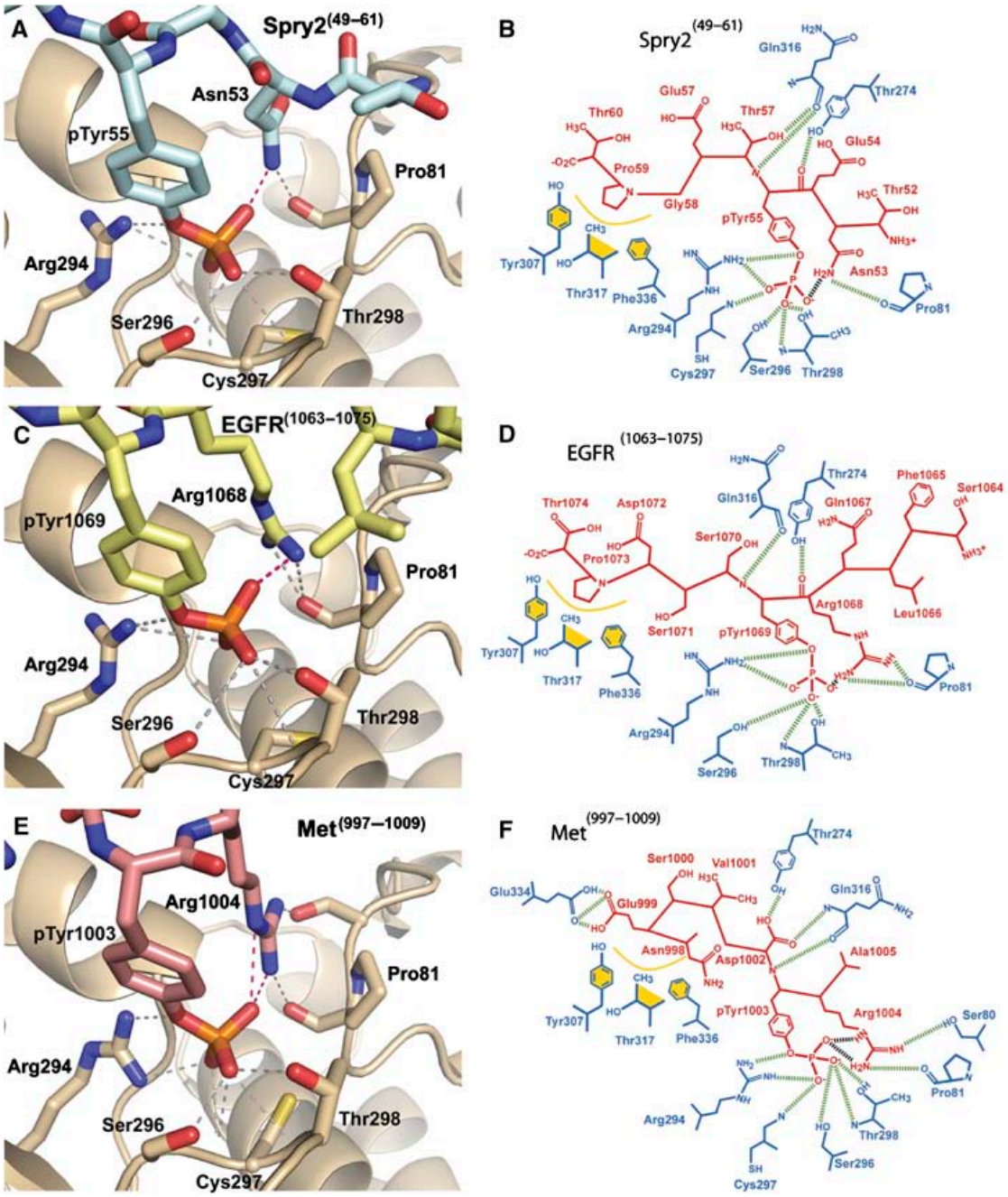


Fig. 2.8. The conserved intrapeptidyl hydrogen (H) bond and neighbouring H-bonds for the conserved pTyr of Spry2, EGFR and Met (Ng *et al.*, 2008). Close-up view of the tyrosine-binding pocket of the c-Cbl-TKB domain complexed with (**A**) Spry2$^{49–61}$ (cyan), (**C**) EGFR$^{1063–1075}$ (yellow) and (**E**) Met$^{997–1009}$ (pink) shows the conserved intrapeptidyl H-bonds between the pTyr and the (pY-2)Asn of Spry2 or the (pY-1)Arg of Met. Intrapeptidyl H-bonds are shown as dotted red lines and other H-bonds are shown as grey dotted lines. The figure was prepared using Pymol (DeLano, 2002). A schematic linear extended view of the c-Cbl-TKB domain with (**B**) Spry2, (**D**) EGFR and (**F**) Met. The phosphopeptides are coloured red and the Cbl residues are blue. Residues participating in van der Waals contacts are highlighted in gold with the interacting surfaces as gold curved lines. H-bonds are indicated by green dashed lines and the intrapeptidyl H-bond is depicted by a black dashed line between the phosphate group of the conserved tyrosine and the side chain of the Asn53 (Spry2), Arg1068 (EGFR) or the Arg1004 (Met).

2.3.6 Met binds to the TKB domain in the reverse orientation

One of the more unexpected findings was that the Met peptide bound to c-Cbl-TKB domain in the reverse orientation from C to N terminus (Fig. 2.6I and J, 2.8E and F). This is contrary to previous modelling predictions (Peschard *et al.,* 2004) and demonstrates that the conserved Asp1002 and Arg1004 residues have more than a supporting role for the pTyr in the binding of Met and c-Cbl. Similar to the other peptides of the PTK group, Met binding in the reverse direction forms six H-bond contacts from the essential pTyr residue with Arg294, Ser296, Cys297 and Thr298 on c-Cbl in a positively charged binding pocket (Fig. 2.8E and F), and similar to the binding with EGFR, the Arg1004 in Met substitutes for the missing (pY-2)Asn and forms H-bond contacts with Pro81. Arg1004 also forms an additional H-bond with Ser80, similar to that observed with the Syk complex, but not with the other phosphopeptides.

The Asp1002 residue in Met assumes the role of the conserved (pY+1)Ser/Thr in the PTK subgroup by interacting with Gln316 and Tyr274 on Cbl-TKB. In contrast to the modelling predictions by Peschard *et al.,* (2004), neither the backbone carbonyl nor amino groups of either residue interact with Cbl-TKB, and no salt bridge is formed between the side chains of Asp1002 and Arg1004 to orientate the phosphorylated tyrosine residue into its binding cleft on Cbl-TKB. Rather, the side chains of both residues are orientated into the Cbl-TKB:Met interface in the same direction as pTyr1003. The Glu999 residue of Met forms unique H-bonding contacts with Glu334 of the Cbl-TKB that is not observed in any other complex; its aliphatic part of the side chain partially occupies the hydrophobic cleft that is occupied by (pY+4)Pro residues of the PTK subgroup.

2.3.7 An intrapeptidyl hydrogen bond is conserved across all TKB domain-binding proteins

Within the peptides of Spry2, Spry4 and Syk, the formation of an intarpeptidyl H-bond was observed between the (pY-2)Asn the phosphate oxygen (represented as a red dotted line in the Spry2 peptide in Fig. 2.8A, or a black dotted line in Fig. 2.8B when bound to the TKB domain. This H-bond was reported in earlier studies (Meng *et al.,* 1999; Hu and Hubbard, 2005); however, no significance was assigned to its existence. The presence of this H-bond across all peptides with a (pY-2)Asn strongly suggests that it is evolutionarily conserved. In addition, a similar intrapeptidyl H-bond between the phosphate oxygen of the pTyr and the guanidium group of the Arg at the pY-1 position in the EGFR peptide (Arg1068) was also observed, which appears to be compensating for the loss of (pY-2)Asn (Fig. 2.8C and D). Likewise, with a reversal in the binding mode of the Met peptide, the same intrapeptidyl H-bond as seen in EGFR was also observed to form between the (pY+1)Arg1004 and the phosphate oxygen of the pTyr in the Met peptide (Fig. 2.8E and F). Taken together, these findings demonstrate that an intrapeptidyl H-bond between the pTyr and the (pY-2)Asn, the (pY-1)Arg or the (pY+1)Arg in the reverse orientation (hereafter also referred to as the adjacent Arg) is conserved across TKB domain-binding substrates, and we propose that this intrapeptidyl H-bond is essential for binding with the Cbl-TKB domain.

2.3.8 Full-length protein binding confirms that (pY-2)Asn and pTyr residues of Spry2 are indispensable for binding

To confirm and extend upon our structural studies using full-length proteins, the relative importance of each of the four conserved residues was tested by creating alanine point mutations in the TKB-binding motif of Spry2 and examining how this affects its interaction with endogenous c-Cbl under stimulated conditions (Fig. 2.9A). When compared to WT Spry2 binding, it was observed that binding to c-Cbl was disrupted with the Spry2 N53A and Spry2 Y55A mutants, reduced with the Spry2 T56A and P59A mutants and unaffected with the Spry2 E54A mutant. This supports the crystal structure observations that (pY-2)Asn and pTyr are essential for Cbl-TKB binding and suggests that Cbl's interactions with T56A and P59A are important, but not essential for binding, whereas the nonconserved Spry2 E54A mutant is not involved in binding.

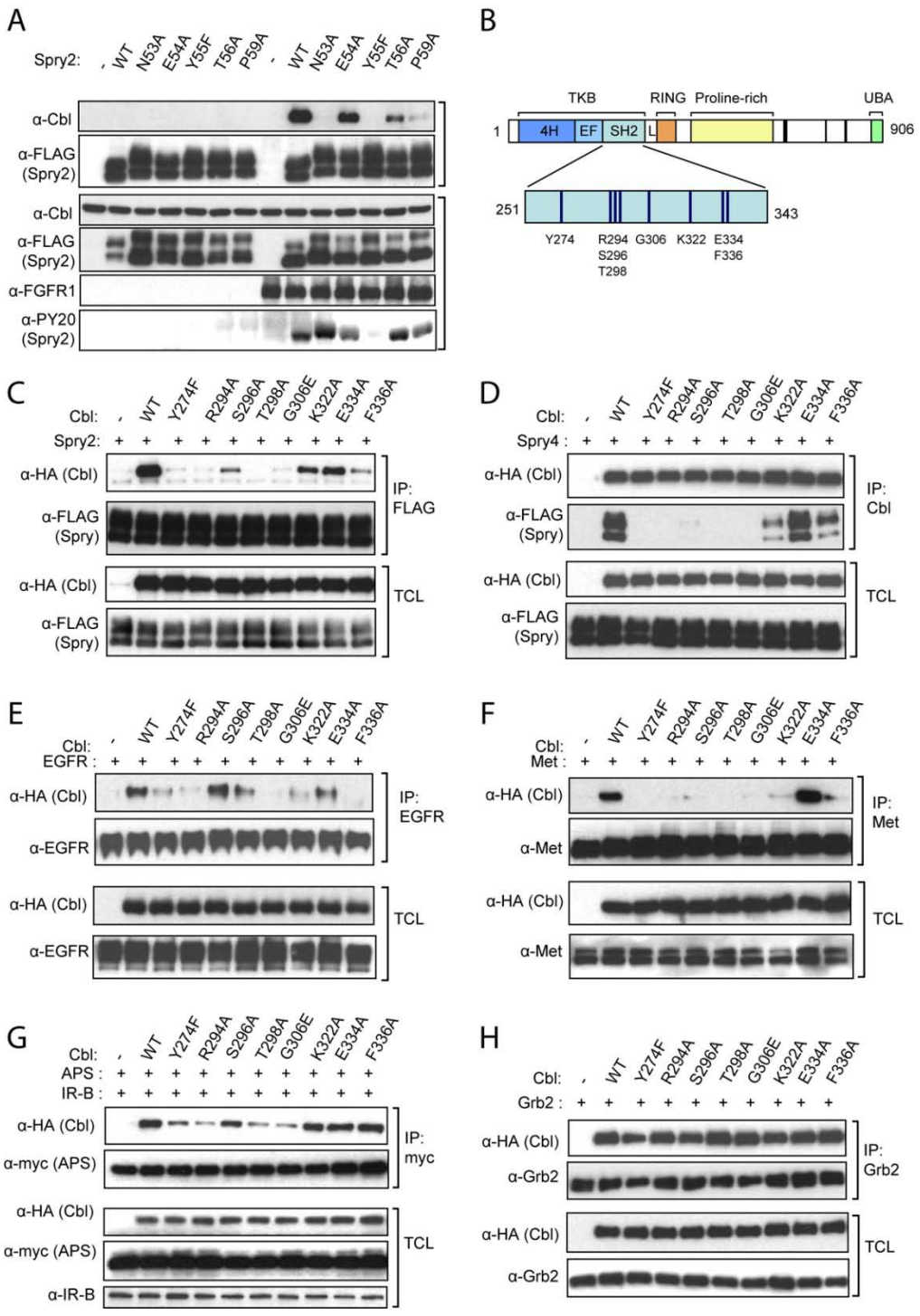


Fig. 2.9. Site-directed mutagenesis to determine the importance of conserved residues (Ng *et al.*, 2008). (**A**) 293T cells were transfected with FGFR1 and FLAG-tagged wild-type Spry2 or Spry2 bearing alanine point mutations as indicated. The cells were lysed 24 h post-transfection and FLAG–Spry2 proteins were immunoprecipitated (IP) using FLAG–M2 beads. IP and total cell lysates (TCL) were immunoblotted with anti-Cbl (C-15), anti-FLAG or anti-FGFR1 (C-15) as labelled. (**B**) Diagram of c-Cbl with an enlargement of the TKB domain. The point mutations created in the SH2 subdomain of Cbl are described. (**C**–**H**) 293T cells were transfected with HA-tagged WT c-Cbl and Cbl bearing various point mutations as described in (B), along with (C) FLAG–Spry2, (D) Spry4, (E) EGFR, (F)

Met, (G) myc-APS in combination with insulin receptor-B (IR-B) or (H) alone for endogenous Grb2. Cells were stimulated as described in the Materials and methods and lysed 24 or 48 h post-transfection. In (C) and (D), anti-FLAG and anti-Cbl IPs (C-15) and TCLs were immunoblotted with anti-FLAG and anti-HA (Cbl). In (E), anti-EGFR IPs (528) and TCLs were immunoblotted with anti-HA (Cbl) and anti-EGFR (1005). In (F), anti-HGFR IPs and TCLs were immunoblotted with anti-HA and anti-Met (25H2). In (G), anti-myc IPs (A-14) and TCLs were immunoblotted with anti-HA, anti-c-myc (9E10) and IR-β (C-19). In (H), anti-Grb2 IPs (C-23) and TCLs were immunoblotted with anti-HA and anti-Grb2.

2.3.9 Binding between full-length Cbl and its targets validates the peptide- and domain-derived structural studies

To further validate the crystal structures, key residues on c-Cbl that were shown to mediate binding in the crystal structures were mutated (Fig. 2.9B) to test their interaction with WT proteins. In all cases, binding was observed for WT c-Cbl with Spry2 (Fig. 2.9C), Spry4 (Fig. 2.9D), EGFR (Fig. 2.9E), Met (Fig. 2.9F) and APS (Fig. 2.9G). This binding was abrogated with the c-Cbl G306E mutant, a nonbinding mutation originally discovered in a genetic screen on *Sli-1*, the *Caenorhabditis elegans* homologue of c-Cbl (Levkowitz *et al.,* 1999) that was shown to disrupt the binding with full-length ZAP-70 (Meng *et al.,* 1999); included here as a negative control.

As with the G306E mutant, a phenylalaine to Tyr substitution (Y274F) in c-Cbl also caused a loss of binding to Spry2, Spry4, EGFR and Met. The OH of Tyr274 on Cbl-TKB binds to the invariant backbone carbonyl O of the pY-1 nonconserved residue of Spry2, Spry4 and EGFR or the (pY-1)Asp of Met, which is tested and confirmed by the E45A mutant that retains binding (Fig. 2.9A). The side chains of residues Arg294, Ser296 and Thr298 on c-Cbl formed H-bond contacts with the essential pTyr on all of the peptides, and this is confirmed by a loss or reduction in binding with alanine substitutions of these residues. The S296 residue, however,

appeared to be less important as some or all of the binding was retained. Phe336 participates in hydrophobic interactions with (pY+4)Pro, and an alanine substitution caused a reduction in binding between c-Cbl and all of the WT proteins, where present. (pY-4)Glu999 of Met occupies the same cleft, and an alanine mutation of its side chain causes a similar reduction in binding.

Alanine point mutations were also included for K322 and E334 as the crystal structures showed specific interactions for EGFR (Asp1072) and both EGFR (Thr1074) and Met (Glu999), respectively. However, in addition to EGFR, we observed a reduction in binding for Spry2 and Met with the c-Cbl K322A mutant. K322 is a solvent-exposed residue, and an alanine mutation may alter the surface conformation of c-Cbl and reduce its binding with proteins that are in close proximity to c-Cbl. This is possibly the case with the C termini of Spry2, Met and EGFR peptides when compared with C termini of Spry4 and Syk, which bend away from the c-Cbl's surface and are therefore unaffected. In contrast, binding was retained for all of the WT proteins with the c-Cbl E334A mutant, suggesting that the binding mediated by this residue is not essential.

Finally, to ensure that the structural changes created in the c-Cbl mutants did not influence c-Cbl's interaction with other proteins C-terminal to the TKB, such as Grb2 which binds via its SH3 domain to c-Cbl's proline-rich region, binding of these mutants to Grb2 was tested. The results confirmed that binding to Grb2 was unaffected by the mutations created in the TKB domain (Fig. 2.9H).

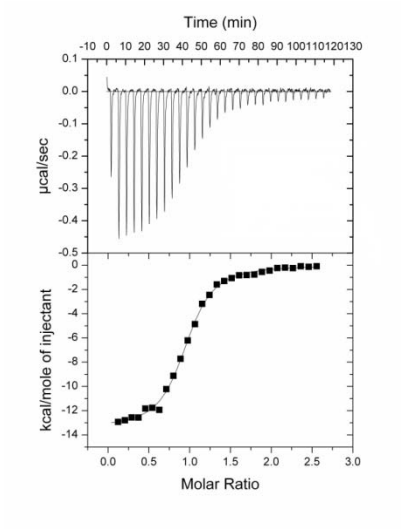
2.3.10 Isothermal titration calorimetry reveals that Spry2 has the highest binding affinity to Cbl-TKB

To validate the necessity of the intrapeptidyl H-bond, it was examined whether binding can still occur in proteins that lack either the (pY-2)Asn or the adjacent Arg, as well as the relative contributions to binding made by the conserved residues in the consensus motifs. The binding affinities of nine 13-mer phosphopeptides were determined by titration against the Cbl-TKB domain in isothermal titration calorimetry (ITC) experiments (Table 2.3): six sequences from the PTK subgroup (Spry2, Spry4, EGFR, Src, Syk, p75 NTR and VEGFR), with the exception of Spry2, other sequences lacking at least one of the conserved residues, the APS sequence and two sequences belonging to the Met subgroup (Met and Ron). Importantly, although the binding affinity for APS differed from that previously reported by Hu and Hubbard (2005) who used a pH 8.0 for their measurements, when the measurements were repeated measurements at pH 8.0 instead of pH 7.0, our findings replicated Hu and Hubbard's findings (Table 2.3; Fig. 2.10B).

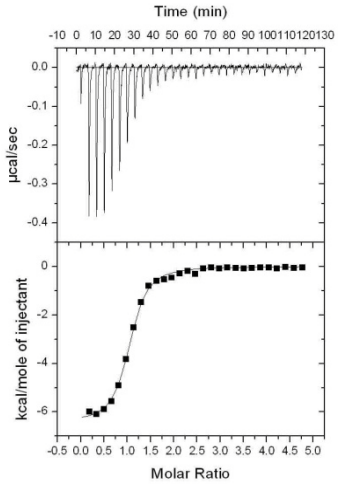
Table 2.3. Sequence, affinity and favourability of the c-Cbl TKB binding motifs.

Peptide	**Sequence**	**N**	**Ka ($x10^6$ M^{-1})**	**Kd (μM)**	**ΔH (kcal/mol)**	**TΔS (kcal/mol)**	**ΔG (kcal/mol)**
Spry2$^{49\text{-}61}$	IRNTNEpYTEGPTV	0.920 ± 0.007	3.11 ± 0.25	0.32	-13.44 ± 0.14	-4.663	-8.777
APS$^{623\text{-}632}$	RAVENQpYSFY	1.010 ± 0.007	2.54 ± 0.20	0.39	- 6.44 ± 0.07	2.217	-8.653
Spry4$^{47\text{-}59}$	SHVENDpYIDNPSL	1.060 ± 0.007	1.65 ±0.09	0.61	- 5.11 ± 0.05	3.276	-8.386
EGFR$^{1063\text{-}1075}$	DSFLQRpYSSDPTG	0.993 ± 0.01	1.14 ± 0.11	0.88	- 9.11 ± 0.15	-0.871	-8.242
Src$^{413\text{-}425}$	LIEDNEpYTARQGA	0.969 ± 0.01	0.756 ± 0.055	1.32	- 6.33 ± 0.11	1.612	-7.939
Syk$^{317\text{-}329}$	TVSFNPpYEPELAP	0.969 ± 0.01	0.673 ± 0.066	1.49	- 3.07 ± 0.08	4.811	-7.882
Ron$^{1011\text{-}1023}$	LYSGSDpYRSGLA	1.000 ± 0.01	0.644 ± 0.056	1.55	- 1.58 ± 0.03	6.257	-7.835
Met$^{997\text{-}1009}$	SNESVDpYRATFPE	1.080 ± 0.03	0.424 ± 0.053	2.36	- 1.60 ± 0.06	5.991	-7.593
Met$^{997\text{-}1009}$ D1002R, R1004D	SNESVRpYDATFPE	1.02 ± 0.01	0.436 ± 0.022	2.29	- 1.09 ± 0.02	6.522	-7.612
APS$^{623\text{-}632}$ pH 8.0	RAVENQpYSFY	0.911 ± 0.002	18.4 ± 1.1	0.054	-10.96 ± 0.04	-1.071	-9.889
EGFR$^{1063\text{-}1075}$ pH 8.0	DSFLQRpYSSDPTG	0.991 ± 0.004	2.97 ± 0.12	0.336	- 6.178 ± 0.03	2.603	-8.781
VEGFR$^{1327\text{-}1338}$	YNSVVLpYSTPPI	0	0	0	0	0	0
VEGFR$^{1327\text{-}1338}$ V1331N	YNSVNLpYSTPPI	1.050 ± 0.002	3.01 ± 0.09	0.33	- 5.003 ± 0.02	3.748	-8.751
VEGFR$^{1327\text{-}1338}$ L1332R	YNSVVRpYSTPPI	0.997 ± 0.003	1.38 ± 0.04	0.72	- 3.238 ± 0.02	5.047	-8.285

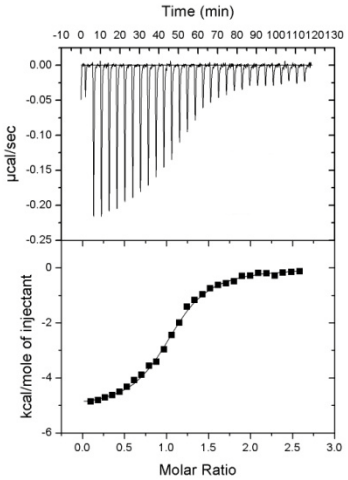
(A) ITC at pH 7.0, 200mM NaCl



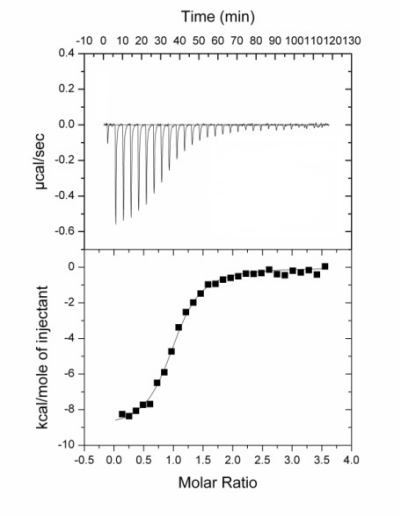
Spry2



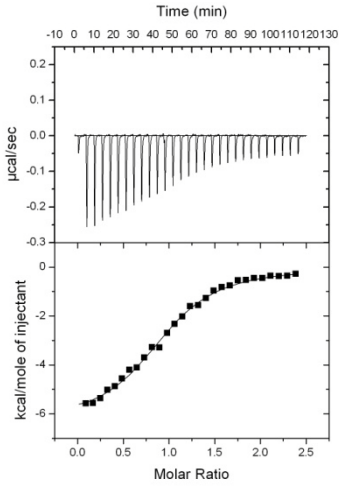
APS



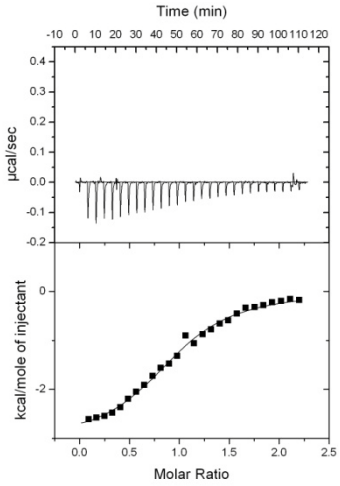
Spry4



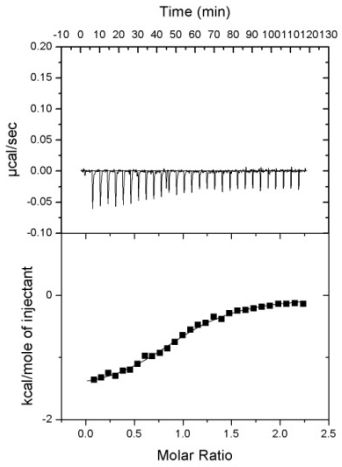
EGFR



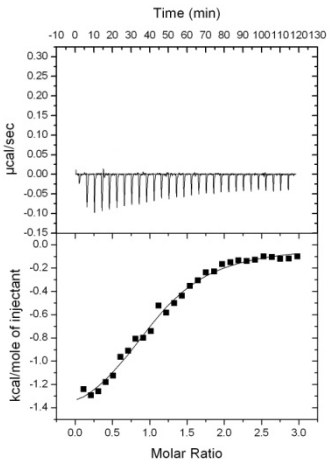
Src



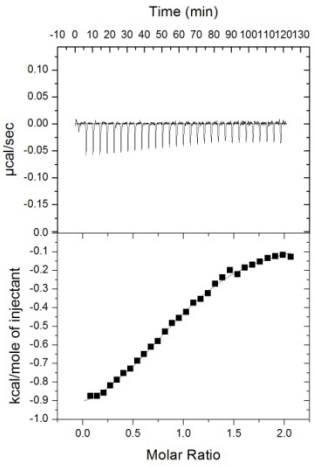
Syk



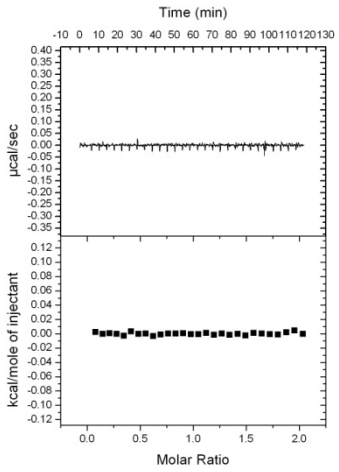
Ron



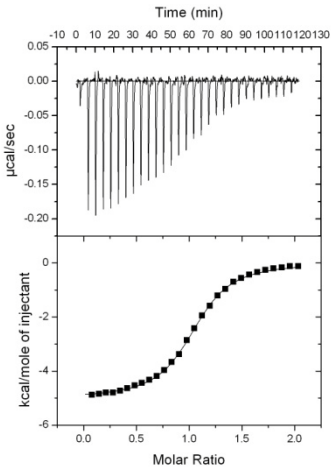
Met



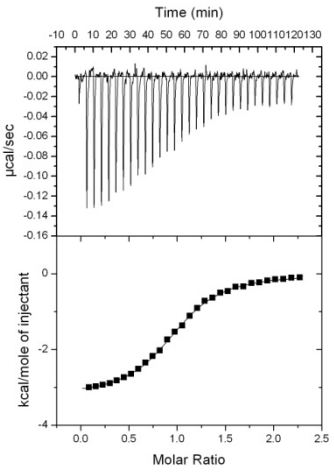
Met D1002R / R1004D



VEGFR1

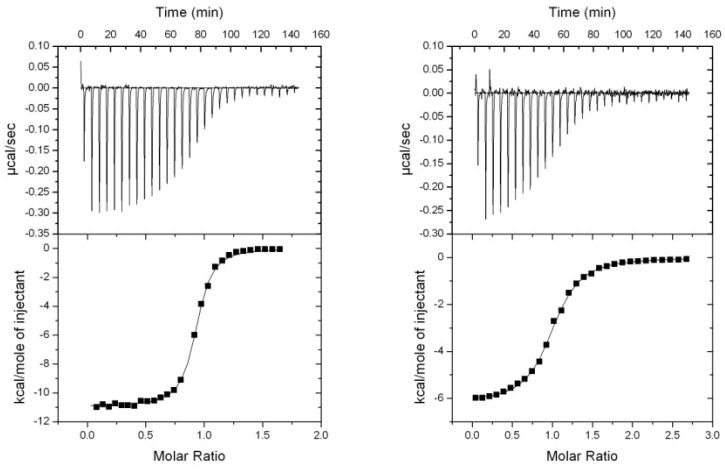


VEGFR1 V1331N



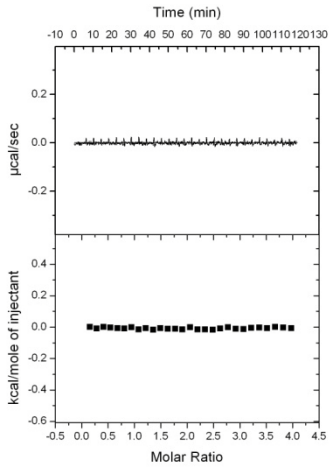
VEGFR L1332R

(B) ITC at pH8.0, 100mM NaCl (reproducing data of Hu and Hubbard, 2005)



APS EGFR

(C) Negative control using non-phosphorylated peptide



Spry2

Fig. 2.10. The binding affinities of phosphorylated Spry2$^{49–61}$, APS$^{623–632}$, Spry4$^{69–81}$, EGFR$^{1063–1075}$, Src$^{413–425}$, Syk$^{317–329}$, Met$^{997–1009}$, Met$^{997–1009}$ D1002R / R1004D, Ron1011$^{–1023}$, VEGFR$^{1327–1338}$, VEGFR$^{1327–1338}$ V1331N and VEGFR$^{1327–1338}$ L1332R phosphopeptides to the c-Cbl-TKB domain were measured by ITC and profiles shown respectively (Ng *et al.*, 2008). The figures show the injection profile after baseline correction and the bottom panels show the integration (heat release) for each injection (except the first one). The solid lines in the bottom panel show the fit of the data to a function based on a one-site binding model. The binding constants (K_a and K_d), number of binding sites (N), enthalpy (ΔH) and entropy (ΔS) changes of Cbl-TKB to the various peptides are provided in (Table 2.3). ITC experiments for APS$^{623–632}$ and EGFR$^{1063–1075}$ were also repeated at pH 8.0 according to Hu and Hubbard (2005). Non-phosphorylated Spry2$^{49–61}$ was titrated against Cbl-TKB as a negative control. The concentration of the TKB domain was 10.8 µM in all experiments.

Of the nine peptides examined, binding between the Cbl-TKB and the phosphorylated Spry2 peptide occurred with the highest affinity and the most negative Gibbs free energy change (Table 2.3), suggesting that this reaction is most favoured. In descending order of affinity, the dissociation constants (K_d) were as follows: Spry2$^{49–61}$(0.32μM) > APS$^{623–632}$ (0.39μM) > Spry4$^{69–81}$(0.61μM) > EGFR$^{1063–1075}$ (0.88μM) > Src$^{413–425}$ (1.32μM), Syk$^{317–329}$ (1.49μM), Ron$^{1011–1023}$ (1.55μM) and Met$^{997–1009}$ (2.36μM); the latter four were not significantly different from each other. Strikingly, the phosphorylated VEGFR1$^{1327–1338}$ that did not possess a conserved (pY-2)Asn or (pY-1)Arg and was unable to bind Cbl-TKB, supporting our hypothesis that the intrapeptidyl H-bond is essential for binding to the Cbl-TKB domain. The ITC profiles are shown in Fig. 2.10.

When comparing between the (pY-2)Asn and the adjacent Arg, it was observed that the Asn residue provides a greater contribution to binding affinity, as demonstrated by Spry2 versus EGFR, both of which contain the other three conserved residues, and by Src versus Ron/Met, which only have one other conserved residue. A degenerate peptide library characterization of the (D/N)XpY(S/T)XXP motif indicated that the most favourable residues for binding N-terminal of pTyr are: (pY-2)Asn, (pY-2)Asp or (pY-3)Asp (Liu *et al.*, 2002). Surprisingly, an Arg residue N-terminal of pTyr was not significantly selected for binding. It is likely that there is no preference for Arg at the pY-1 position, because an Asn provides higher affinity binding and, as their roles overlap in this context, an Asn at the pY-2 position will be preferred.

Contributing next to the binding affinity is the (pY+4)Pro residue, as shown by the higher affinity of EGFR compared with Met and Ron, and the higher affinity of Spry2 compared with Src and Syk. APS, whose affinity is just below Spry2, also lacks the (pY+4)Pro, but appears to compensate for this by using other conserved residues N-terminal to the pTyr to enhance

binding. Interestingly, the extent of the hydrophobic interaction between the (pY+4)Pro and its corresponding cleft on Cbl-TKB was directly correlated with the peptide's affinity for the latter: hydrophobic interaction was most extensive with Spry2 at this cleft, followed by Spry4 and, to a lesser extent, EGFR. The conserved (pY+1)Ser/Thr appears to contribute least to the binding affinity, as suggested by the insignificant differences in affinity between Src and Syk. All of these results support the variations in binding we observed with the full-length proteins.

The (pY-1)Asp of Ron and Met does not appear to contribute substantially to binding affinity, as these peptides have no significant differences in affinities with Syk, which has no additional residues other than the (pY-2)Asn. Interestingly, most of the interactions between Met and c-Cbl are mediated by the DpYR residues. The side chain of the (pY-1)Asp binds to the Tyr274 of Cbl-TKB, whereas in other complexes, this interaction is mediated by the backbone of the pY+1 residue (Table 2.3). Since the Met D1002A mutation cannot bind to c-Cbl (Peschard *et al.,* 2004), and since the Cbl Y274F mutant cannot bind to Met, this may indicate why the Asp residue in the Met family of proteins has been conserved.

2.3.11 The intrapeptidyl H-bond is essential for binding to the TKB domain

To further validate the importance of the intrapeptidyl H-bond, two mutant phosphorylated peptides for VEGFR1 harbouring Asn or Arg substitutions were included. These mutations would permit formation of an intrapeptidyl H-bond and facilitate binding if these amino acids were the key requirement for binding: (1) VEGFR1 with a Val1331Asn substitution at the pY-2 position, and (2) VEGFR1 with a Leu1332Arg substitution at the pY-1 position. This substitution would also endow VEGFR1 with four conserved residues, and it would therefore be

expected to bind with similar affinities to Spry2 and EGFR, respectively. Indeed, both of these substitutions facilitated peptide binding to the Cbl-TKB domain (Fig. 2.10) and, as expected, binding affinities were similar to Spry2 and EGFR (0.33 and 0.72μM respectively) were observed (Table 2.3). These findings confirm the requirement for the formation of an intrapeptidyl H-bond between pTyr and (pY-2)Asn or an adjacent Arg at the pY-1 or pY+1 position when reversed, and also substantiate the greater contribution of Asn towards a substrate's affinity for Cbl-TKB.

2.3.12 Inversion of the DpYR motif in Met preserves binding

The finding that Met binds to Cbl-TKB in a reverse orientation, with (pY+1)Arg replacing the role of (pY-2)Asn to form an essential hydrogen bond with phosphotyrosine, lead to the hypothesis that the Met peptide would still be able to bind Cbl-TKB if the DpYR motif was flipped to RpYD, provided that the adjacent residues do not pose any steric hindrances. Indeed, ITC of phosphorylated $Met^{997-1009}$ D1002R / R1004D shows that the binding affinity of this Met double mutant was not significantly different from the wild type $Met^{997-1009}$ sequence (Table 2.3), suggesting that residues flanking the DpYR motif are not necessary for binding to Cbl-TKB, although this hypothesis remains to be confirmed by structural studies.

2.4 Discussion

The aim of this study was to address two major questions: (1) Why does the TKB domain recognize an unusually broad set of sequence-recognition motifs and (2) could variations in the binding affinity of TKB domain-binding proteins result in the selective sequestration of c-Cbl by higher-affinity binding proteins. In addressing the latter, results suggest that differences in binding affinity do exist between different TKB-binding proteins. We cannot confirm, however, whether the sequestration of c-Cbl is physiologically important. Given a favourable cellular localization, it is possible that Spry2 with its higher binding affinity, could influence the binding of c-Cbl to particular RTKs (e.g. EGFR) or other TKB-binding proteins. In addressing the first question, two novel and unexpected elements have been revealed: (1) binding is essentially achieved through (D/N)XpY, RpY or pYR (in the reverse orientation) through the formation of an intrapeptidyl H-bond, and (2) the Cbl-TKB domain can bind to substrates in two orientations.

The unique diversity of the TKB domain in its recognition of three consensus sequences is an unusual characteristic for an SH2 domain. These five complexes demonstrate some consistency in binding between the Cbl-TKB targets, and together indicate that the TKB domain recognizes a less-radical set of motifs. The intrapeptidyl H-bond between pTyr and (pY-2)Asn, (pY-1)Arg or (pY+1)Arg in the reverse is necessary for the directed, 'hand-in-glove' fit of the peptide into a positively charged pocket of the SH2 region of the Cbl-TKB domain (depicted in Fig. 2.11A–D for Spry2 and Met), whether before binding to c-Cbl or concomitantly. Although this intrapeptidyl H-bond was identified in both of the previously characterized c-Cbl structures (Meng *et al.*, 1999; Hu and Hubbard, 2005), proper consideration was not attributed to it due to a lack of supporting evidence to highlight its significance. When combined with these five structures, we are able to assign importance to the intrapeptidyl H-bond as a rule for binding and

an event crucial for the directed positioning of the substrate on c-Cbl. Furthermore, with the ITC characterizations of the three VEFGR peptides, we show convincing evidence that binding to c-Cbl cannot occur without this intrapeptidyl H-bond, despite the presence of other conserved residues (depicted with a model diagram in Fig. 2.11E–G). These other conserved residues—(pY+1)Ser/Thr, (pY+4)Pro and (pY-1)Asp—tweak the binding affinity, and such as in the case of APS, high-affinity binding can be compensated for by these other residues. Substrate binding is further stabilized in concert with the 4H subdomain and EF-hand, particularly the 4H subdomain, which contributes to binding via the Pro81 residue to the conserved (pY-2)Asn or compensating Arg (Fig. 2.11), and structural evidence suggests that these sub-domains form an intimate, multifaceted association with the SH2 subdomain, linker sequence and RING domain in its interaction with the appropriate E2 enzyme, creating what appears to be a precise, rigid structure.

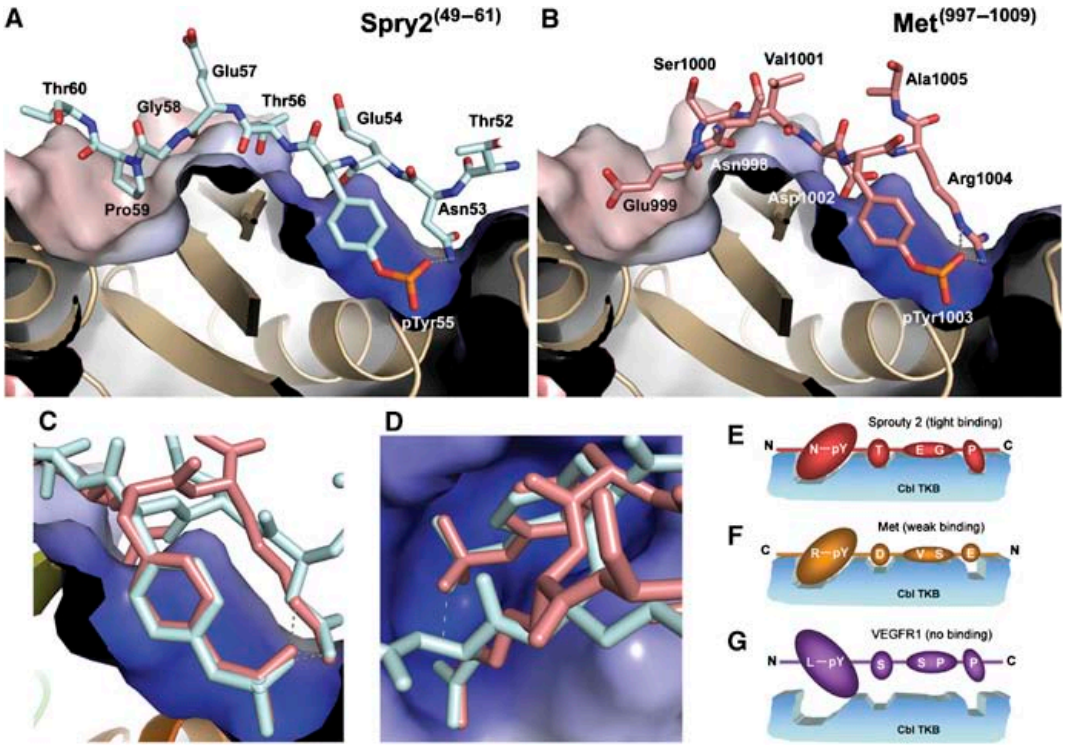


Fig. 2.11. Orientation of the intrapeptidyl hydrogen (H) bond within the binding pocket of the Cbl-TKB (Ng *et al.*, 2008). A cross-sectional perspective of the binding topograph in Figure 2B and F illustrates the fit of the phosphopeptides on the surface of the Cbl-TKB domain for (**A**) Spry2$^{49–61}$ (cyan) and (**B**) Met$^{997–1009}$ (pink). (**C**, **D**) Peptides superimposed from two different angles. (**E**–**G**) Schematic diagrams illustrating the importance of the intrapeptidyl H-bond between pTyr and (E) (pY-2)Asn for Spry2 or (F) (pY+1)Arg for Met in the reverse orientation. In the absence of an intrapeptidyl H-bond (G), pTyr is disorientated and unable to dock into the pTyr-binding pocket on Cbl-TKB, as indicated by VEGFR.

A comprehensive search of the known SH2 domain structures available in the protein data bank resulted in only four complex models where a pTyr formed a H-bond with the side chain of another residue in the peptide (Gunther *et al.*, 2003; Cho *et al.*, 2004; Frese *et al.*, 2006). For the first case, the H-bond was not conserved among the different peptides complexed with the SH2 domain of Gads (Cho *et al.*, 2004). In the Nck1-SH2:phospho-Tir and Nck2-SH2:phospho-Tir complexes, a degenerate peptide library scan showed a preference for hydrophobic residues over the interacting His at that position (Frese *et al.*, 2006). The last complex model between the SH2 domain of PI3 kinase and the phospho-platelet-derived growth

factor receptor peptide identified the preference for negatively charged residues N-terminal to the pTyr that were not essential for binding nor bound to its SH2 domain in the same manner as the conserved (pY-1)Asp binds to the TKB domain of c-Cbl in the Met family of proteins (Gunther *et al.*, 2003).

Similarly, a search performed on known structures of PTB-binding peptides bearing the conserved D/NXpY and complexed with PTB domains also revealed that none of these binding conformations had the signature H-bond between the pTyr and the (pY-2)Asn. To the best of our knowledge, there are no conserved intrapeptidyl H-bonds between pTyr and its neighbouring Asn/Arg residues for SH2 and PTB domain-binding proteins. The formation of the intrapeptidyl H-bond in the D/NXpY, RpY or pYR (in the reverse) motifs is likely to be an essential specificity determinant for the recognition and docking of c-Cbl to its activated (phosphorylated) target proteins, whether before or during binding, and thus represents a mechanism that is unique to Cbl.

The most surprising finding from the crystal structure analyses is the reversed binding of the Met phosphopeptide. From these results, it would be presumed that other members of the Met subgroup also bind in the reverse direction. Interestingly, many of these family members have a proline residue at the pY-4 position (Fig. 1.14), which would align with the (pY+4)Pro residue in members of the PTK subgroup in the reverse orientation, and presumably increase their binding affinity to c-Cbl over Met.

There are several analogous situations where ligand binding can occur in two orientations of a specific domain. The SUMO-binding amino-acid sequence motif of RanBP2 and PIASX-P have been shown to bind to the surface of SUMO-1 in two different orientations (Reverter and

Lima, 2005; Song *et al.,* 2005) and similarly, the SH3 domain of Src can bind to proline-rich peptide ligands in two orientations (Feng *et al.,* 1994). Furthermore, the PH domain of ArhGAP9 accommodates both PI(3,4)P2 and PI(4,5)P2 at 180° rotations via a shallow, noncanonical binding pocket (Ceccarelli *et al.,* 2007). To our knowledge, however, there is no precedent for an SH2-domain-binding protein to bind to an SH2 domain in two orientations, and raises the question whether this may also occur. However, as the TKB domain accomplishes binding in conjunction with the 4H subdomain, its absence in other SH2 domains may restrict them from having such flexibility of binding.

The reverse binding of Met and the previously unidentified importance of the (pY-1)Arg in EGFR suggests that there is a larger cohort of Cbl-TKB domain targets, particularly those with an RpYD or RpYS/T motif. Searching the human proteome, nearly 300 signal transduction proteins containing the RYD sequence have been identified, and while it still remains to be determined whether these proteins are substrates for c-Cbl, this emphasizes the likelihood that other unrevealed c-Cbl targets exist. In addition, the reversed binding of Met may reflect the need for Cbl to bind to targets in two orientations for steric reasons, as the DpYR residues in Met lie very close to the plasma membrane in the juxta-membrane domain. This reversed binding presumably allows the appropriate distance for the subsequent ligation of ubiquitin moieties. In contrast, the reversed binding may also allow c-Cbl to bind to proteins without resulting in their ubiquitination, expanding Cbl's role as an adapter protein in signalling complexes other than the APS family of proteins. Together, these findings provide not only further clues to Cbl's role in signalling but also the incentive to revisit previously identified binding motifs of other proteins, with the potential of redefining the apparently established paradigms of modular signaling.

Chapter III

CaM Regulation of

Voltage-gated Sodium Channels

3.1 Introduction

The regulation of VGSC by CaM was surprising but not unprecedented, as CaM appears to be an important regulator of several different ionic currents. Saimi and Cheng (2002) proposed that CaM itself could be considered a channel subunit. Interestingly, all previous accounts of Ca^{2+} influence on ion channels involve proteins that modify but are not directly involved in generating and propagating the action potential. The association of CaM for VGSC is the first description of such a basic and essential channel for the action potential being regulated by Ca^{2+}. Furthermore, all VGSC isoforms possess the conserved IQ domain for CaM binding, but are not regulated in the same way (Deschenes *et al.*, 2006). Although the common feature is coupling to cytoplasmic Ca^{2+} regulation, it is not understood why the need for certain VGSCs to be Ca^{2+} dependent.

In the early 1990s, a model was proposed in which an auto-inhibitory domain becomes displaced from the active site upon binding of CaM, leading to the activation of the enzymes. However, with the determination of an increasing number of CaM-complex structures, it became evident that there is no general model for binding, since each complex represent a different binding mode (Vetter and Leclerc, 2003). Using two VGSC isoforms from skeletal muscle and neurons, Nav1.4 and Nav1.6 respectively, we propose to characterize and compare the association of CaM with these IQ motifs of high sequence identity. This will be towards understanding (1) the mechanism by which VGSC IQ motifs bind to CaM but retain flexibility to regulate the Na^{+} current in different ways, and (2) how differences in binding affinity translate to the need for Ca^{2+} in various isoforms.

3.2 Method and Material

3.2.1 Plasmid constructs, cloning and expression

Wild type full length mouse (*Mus musculus*) Calmodulin isoform 1 (CaM) cDNA was cloned into pET3a plasmid (Novagen, Darmstadt, DE) between NdeI and BamHI restriction sites. Recombinant CaM was overexpressed in *Escherichia coli* expression host strain BL21(DE3) and purified according to Hayashi *et al.*, 1998 with modifications. The purified protein (1 mg/ml) was stored in buffer A (50mM Tris-HCl pH 7.5, 1mM $CaCl_2$) and used for crystallization and ITC experiments.

E88A and E128A mutations on CaM were performed according to Wang and Malcolm, 1999. Both mutants were expressed in *E. coli* and purified similar to wild type CaM.

Mouse neuronal VGSC subtype Nav1.6 cytoplasmic domain (ΔNav1.6) corresponding to residues 1517-1687 was inserted into pET25b (Novagen, Darmstadt, DE) between NdeI and XhoI restriction sites. Recombinant ΔNav1.6 was expressed as a His-tagged fusion protein (His-ΔNav1.6) in BL21(DE3) (GE Healthcare, Buckinghamshire, UK) and purified from inclusion bodies according to the protocol established by Oganesyan *et al.*, 2005.

3.2.2 Isothermal titration calorimetry

Nav1.4IQ and Nav1.6IQ were titrated against 1mM CaM in a VP-ITC microcalorimeter (Microcal, Northampton, UK) in the high and low Ca^{2+} concentrations. The peptides were reconstituted in buffers A and C at 22°C and titrated in 10 μl injections into CaM maintained in

the respective buffers. The heat of dilution was determined by titrating peptide into buffer in a separate experiment and subtracted from the raw data before curve fitting and refinement. The dissociation constants of Cbl-TKB to the various peptides were determined by least squares method and the binding isotherm was fitted using Origin v7.0 (Microcal) assuming a single-site binding model. All measurements were repeated at least twice for verification.

3.2.3 Computational modelling

In parallel to co-crystallisation experiments, the structures of Nav1.4IQ and Nav1.6IQ in complex with Ca^{2+}/CaM were modelled using DeepView (Guex and Peitsch, 1997). The initial template model was derived from a 2.0Å crystal structure of the Cav1.2 IQ domain – Ca^{2+}/CaM complex, pdb code 2BE6 (Petegem *et al.*, 2005). Positions of the side-chains were refined by energy minimisation.

3.2.4 Complex formation and crystallization

IQ domain peptides of mouse Nav1.4 1719-1742 (Nav1.4IQ) and Nav1.6 1645-1668 (Nav1.6IQ) (Sigma Genosys) were reconstituted in buffer A and buffer C and incubated with the purified CaM in two-fold molar excess at room temperature for 1 h for complex formation. The CaM: peptide complexes were concentrated to 20mg/ml and CaM: His-ΔNav1.6 concentrated to 5mg/ml using an Amicon Ultra ultrafiltration devices (Millipore, Billerica, MA). These complexes were screened for crystals using commercial screen kits from Hampton Research,

CA, Qiagen, NL and Jena Biosciences, DE using hanging drop vapour diffusion and under oil methods respectively. Initial crystallization conditions were optimised by varying precipitant or additive concentrations.

The crystals of Nav1.6IQ: Ca^{2+}/CaM complex were tested and a partial dataset was collected with a Platinum135CCD detector mounted on a Microstar-H rotating anode x-ray generator (Bruker AXS, DE).

3.3 Results

3.3.1 Cloning of CaM, ΔNav1.6

Mouse CaM (447bp) and ΔNav1.6 (510bp) cDNAs were isolated from whole mouse brain by RT-PCR. Total RNA from whole mouse brain was extracted using Trizol reageant (Invitrogen, California, USA), from which the first strand of total brain cDNAs were synthesized from mRNA using oligo-dT primer. CaM and ΔNav1.6 cDNAs with flanking restriction enzyme cut sites were amplified by PCR using their respective primers (Fig. 3.1).

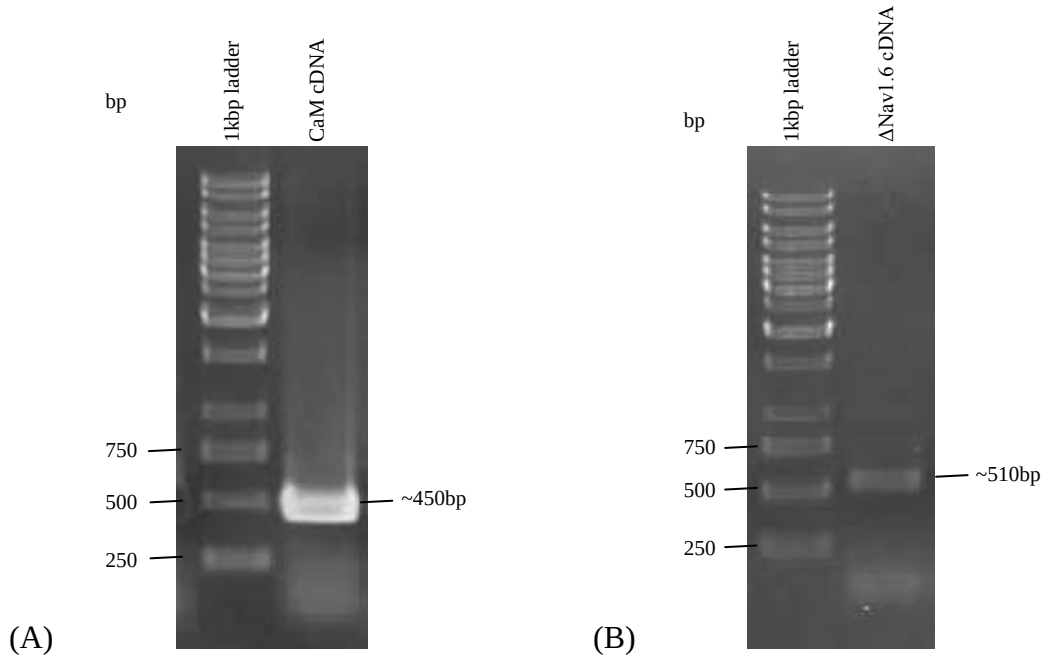


Fig. 3.1. PCR amplification of (A) CaM and (B) ΔNav1.6 cDNAs. RT-PCR was used to clone CaM and ΔNav1.6 from mouse brain, yielding fragments of approximately 450bp and 510bp that correspond to the molecular weights of these two genes.

The 450bp and 510bp fragments were gel-purified and digested with NdeI / BamHI and NedI / XhoI restriction enzymes, and inserted into pET3b and pET25b respectively. Positive

ligation clones were screened by colony PCR, and the correct DNA sequence verified by gene sequencing (Fig. 3.2).

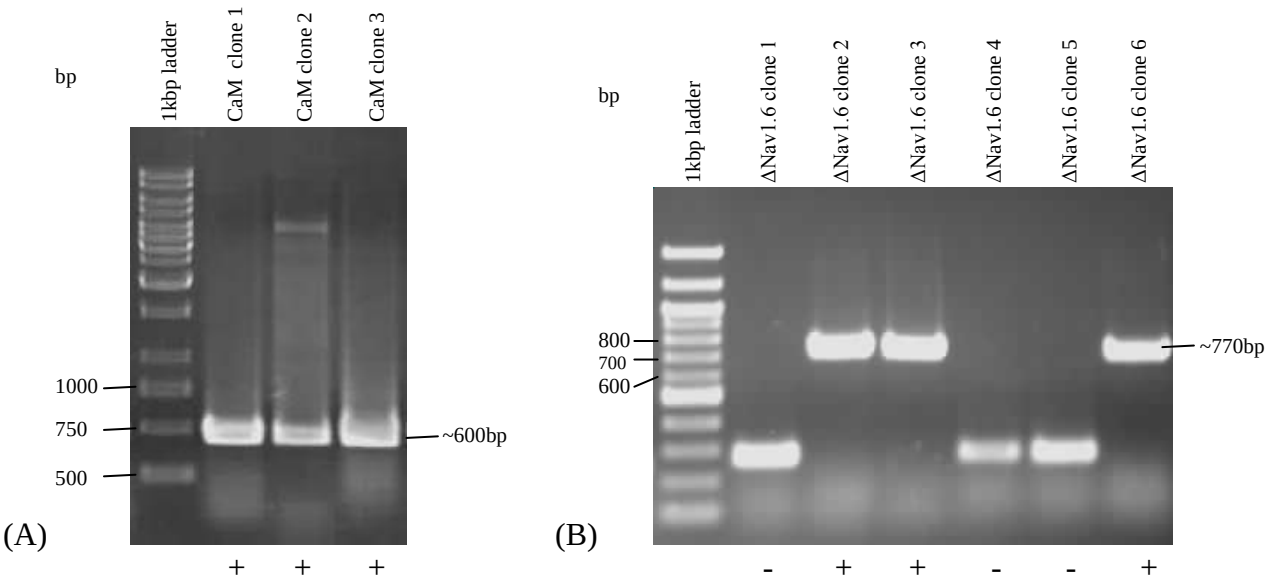


Fig. 3.2. Colony PCR of ligated and transformed (A) CaM and (B) ΔNav1.6 clones. The inserted genes of interest were amplified using vector primers pair T7 promoter / T7 terminator flanking the multiple cloning sites. The amplified fragments were thus approximately 150bp and 260bp larger than the inserted genes respectively, which correspond to the approximately 600bp fragment observed in (A) and 770bp fragment observed in (B).

3.3.2 Expression of ΔNav1.6 and association with CaM

The cloned ΔNav1.6 was transformed into *Escherichia coli* expression strain BL21(DE3). 100ml of Luria-Bertani (LB) media supplemented with 100μg/ml ampicillin was inoculated with one colony of BL21(DE3)(pET25b- ΔNav1.6) and grown overnight at 37ºC with aeration. The overnight culture was amplified to 1L and protein expression induced with 0.1mM isopropyl-thio-galactoside (IPTG) for 16h at 20ºC during mid-log phase of cellular growth.

His tagged ΔNav1.6 (His-ΔNav1.6) was harvested by six cycles of 1min sonication with 40s interval between cycles. Most His-ΔNav1.6 expressed as inclusion bodies and was pelleted in the insoluble fraction. ΔNav1.6 was successfully purified to homogeneity from the insoluble fraction by 8M urea denaturation and refolding according to the method described by Oganesyan *et al.*, 2005 (Fig. 3.3).

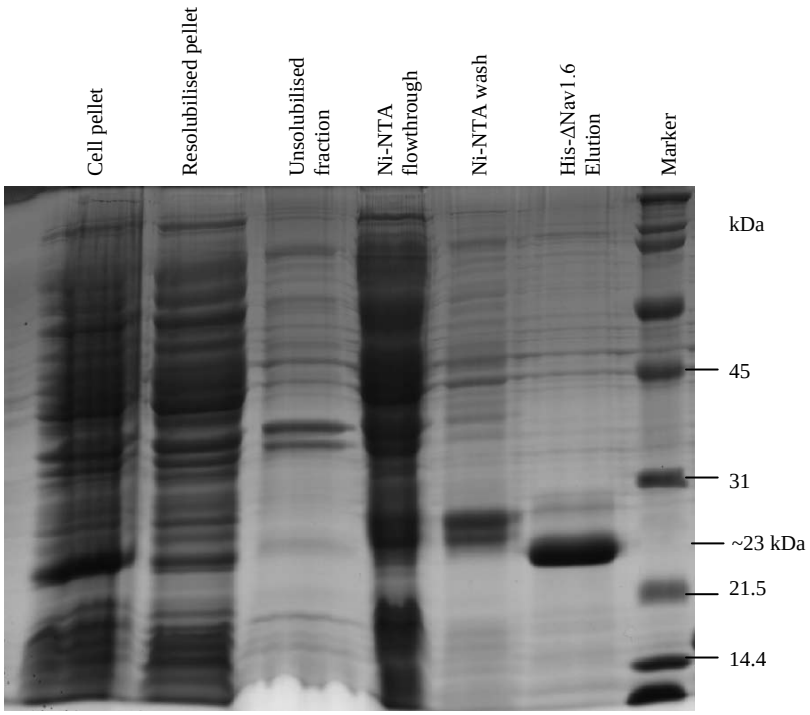


Fig. 3.3. 12.5% SDS-PAGE purification profile of His-ΔNav1.6 purified by 6M urea denaturation and refolding being visualized by Coomassie staining.

The eluted His-ΔNav1.6 from Ni-NTA was stable without any visible precipitate even after a week. It was further purified by using a Superdex-G200 (GE Healthcare, Buckinghamshire, UK) gel filtration column. His-ΔNav1.6 eluted as a homogeneous monomeric species in buffer A with 100mM NaCl and 10mM βME (Fig. 3.4).

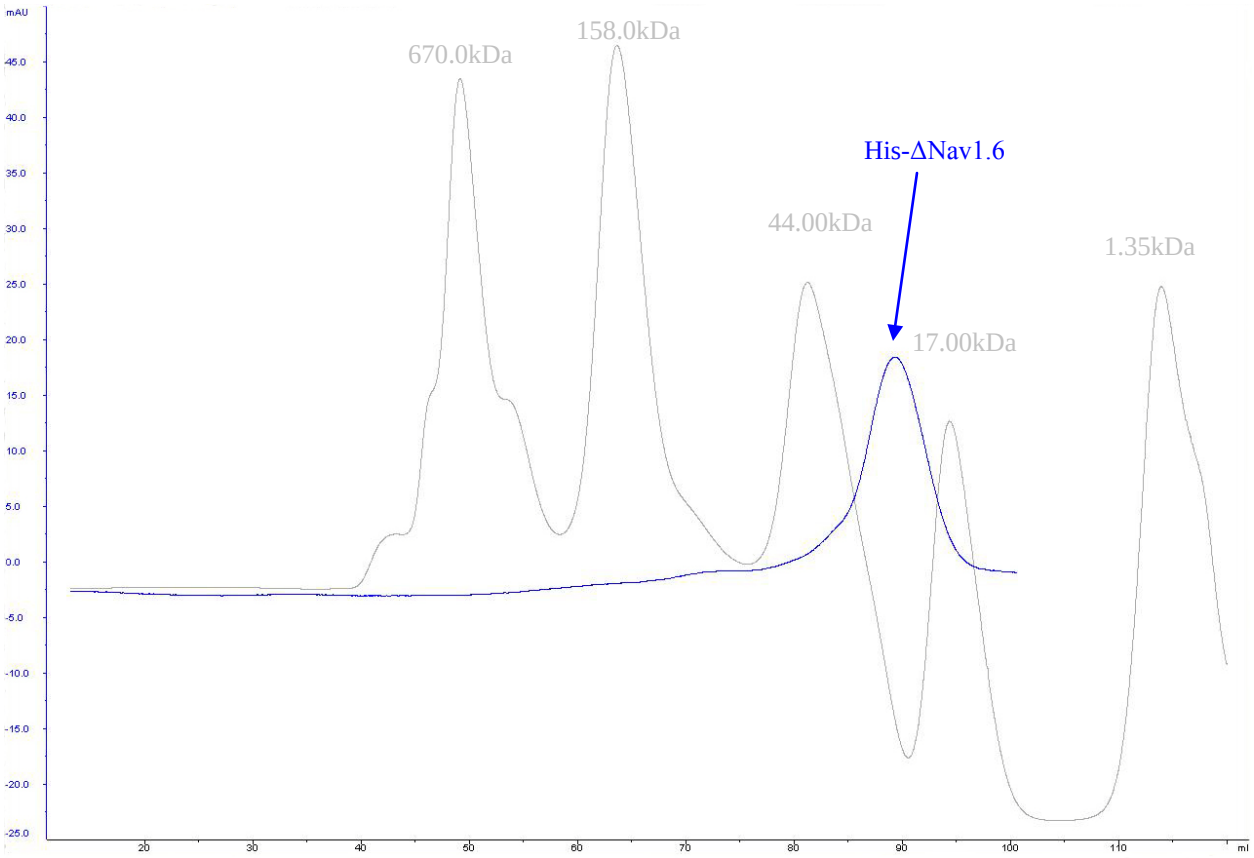


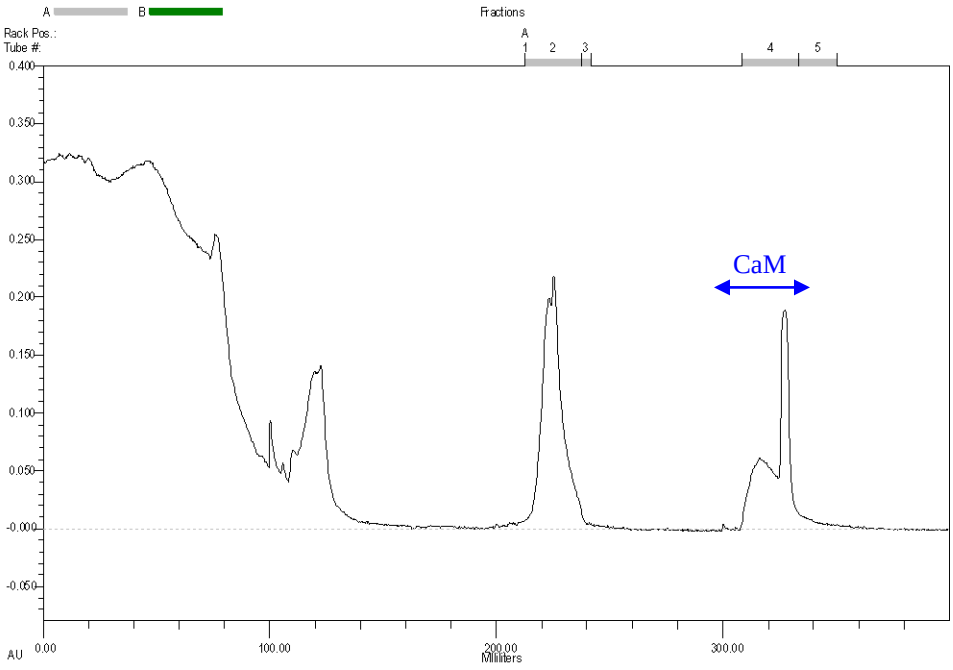
Fig. 3.4. Gel filtration profile of His-ΔNav1.6 in a 120ml column volume Superdex-200 column. His-ΔNav1.6 was purified to homogeneity by gel filtration and eluted in buffer A+100mM NaCl and 10mM βME. Blue plot indicate His-ΔNav1.6 gel filtration profile, grey line is gel filtration molecular size standard with molecular weight of the peaks labeled (BioRad, CA).

3.3.3 Expression and purification of CaM

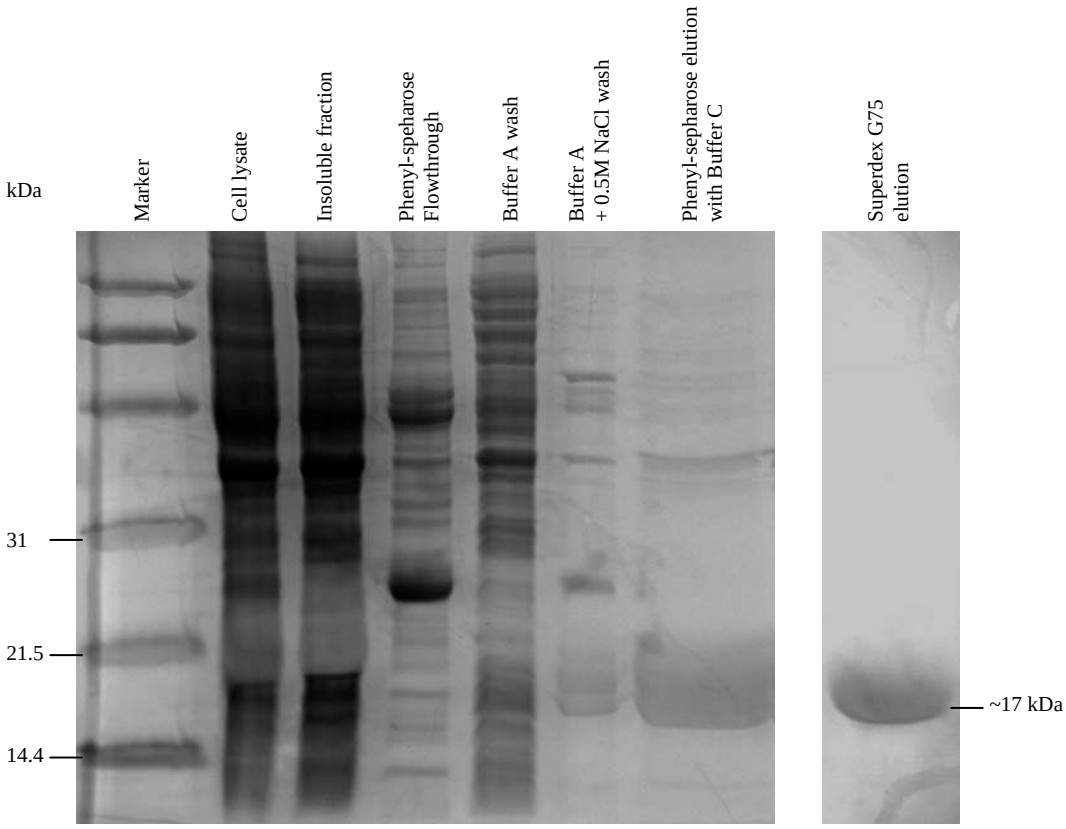
The cloned CaM was transformed into *Escherichia coli* expression strain BL21(DE3), expressed and harvested the same way as His-ΔNav1.6. CaM was purified by a phenyl-sepharose column (GE Healthcare, Buckinghamshire, UK) according to the Hayashi *et al.*, 1998 protocol. Further purification was achieved by using Superdex-G75 (GE Healthcare, Buckinghamshire, UK) gel filtration column. Purified CaM eluted as a single monomeric peak at an approximate peak volume of 80mL in the presence of 100mM NaCl, but in the absence of

salt, CaM behaved as a trimer and eluted at an approximate peak volume of 64mL from the same column (Fig. 3.5C).

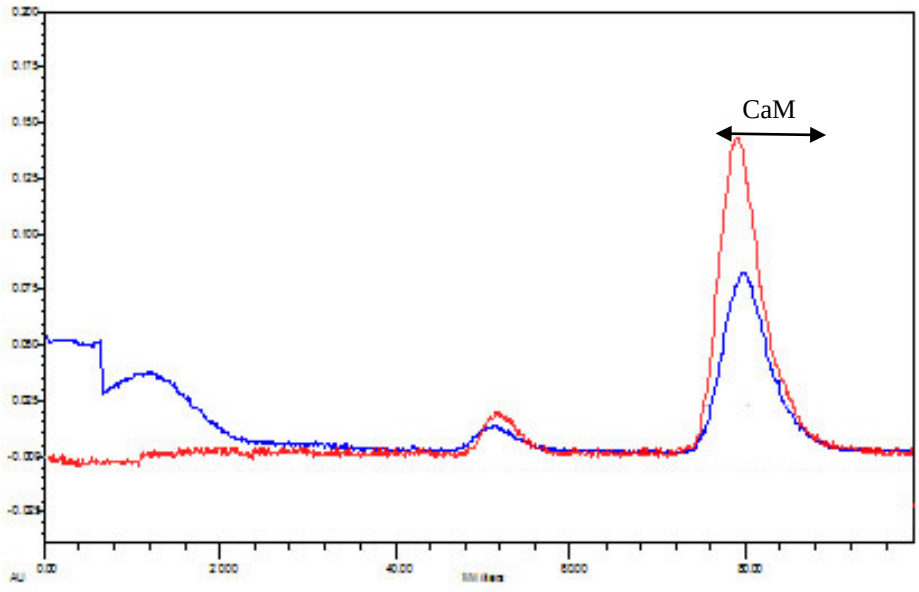
(A) Phenyl-sepharose purification profile of cleared lysate containing CaM



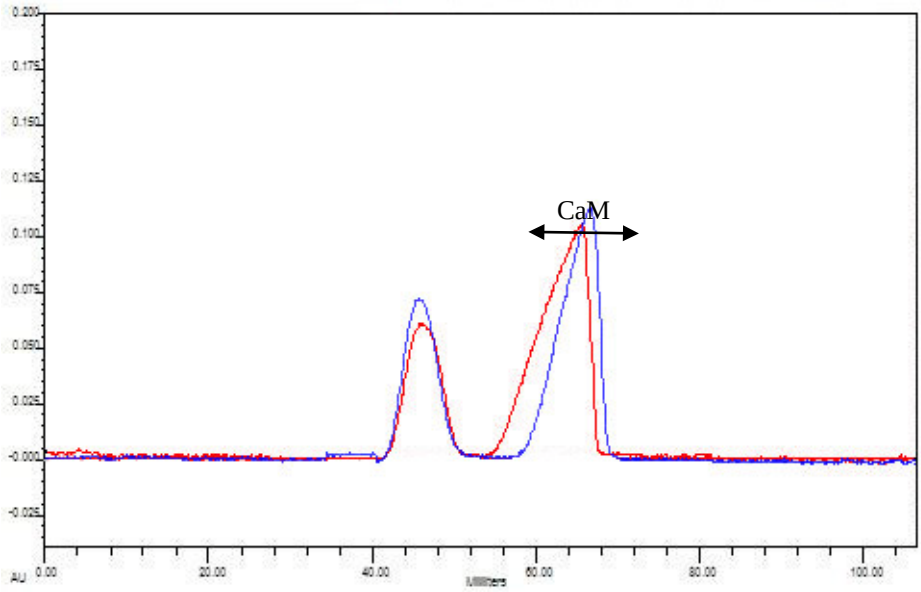
(B) 12.5% SDS-PAGE purification profile of CaM



(C) Gel filtration profiles of purified CaM



(I) CaM elution in 100mM NaCl



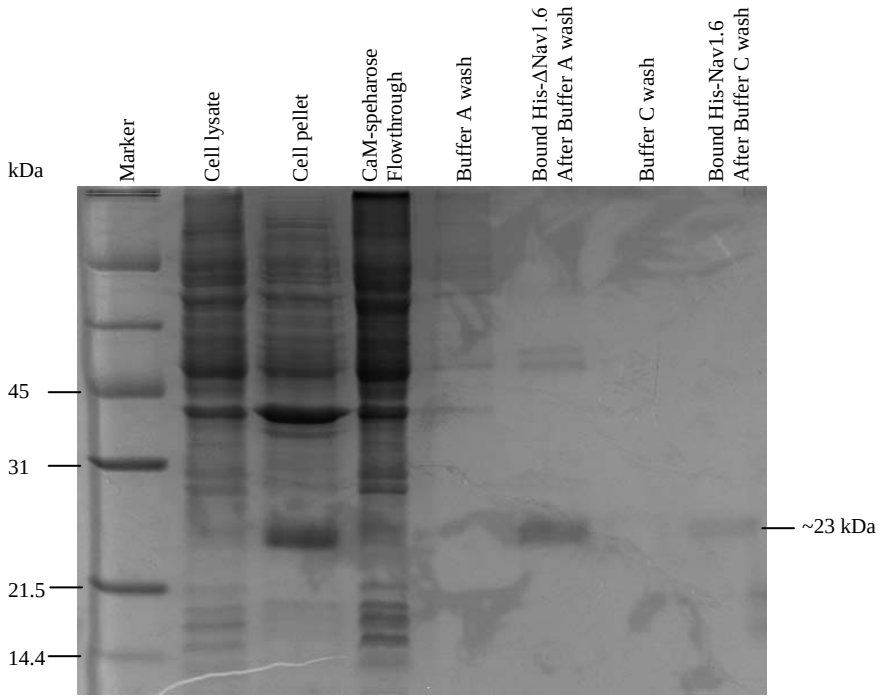
(II) CaM elution without NaCl

Fig. 3.5. Purification and preliminary characterization of CaM. (A) Phenyl-sepharose purification profile of CaM. (B) Samples were obtained at different stages of purification and separated in 12.5% SDS-PAGE. CaM eluted with little impurity phenyl-sepharose and was purified to homogeneity by gel filtration superdex G-75, migrating as an approximately 17kDa band. (C) Eluted CaM was applied to Superdex-G75 gel filtration column and eluted as a single peak of monomeric species in 100mM NaCl (I) but as a trimer in the absence of NaCl (II). Red plot represents gel filtration profile of CaM in buffer A, while blue plot represents CaM elution profile in buffer C. The CaM peak is indicated by black arrows.

3.3.4 Pull down assays and gel filtration confirm Nav1.6 binds to CaM

In order to ascertain if VGSC binds to CaM *in vitro*, resolubilised His-ΔNav1.6 was used as a representative of the two isoforms (Nav1.4 and Nav1.6) that share high sequence similarity (54% sequence identity and 80% sequence similarity) in the IQ motif. His-ΔNav1.6 was divided into two batches of identical protein concentration and bound to CaM-Sepharose (GE Healthcare, Buckinghamshire, UK). One batch was bound and rinsed in buffer A with 100mM NaCl and the other batch in buffer C with 100mM NaCl. His-ΔNav1.6 in high Ca^{2+} concentration (buffer A) showed greater binding to CaM, as indicated on SDS-PAGE (Fig. 3.6A.). This observation is typical of most CaM binding proteins.

(A) His-ΔNav1.6 bound and rinsed in buffer A+ 100mM NaCl



(B) His-ΔNav1.6 bound and rinsed in buffer C + 100mM NaCl

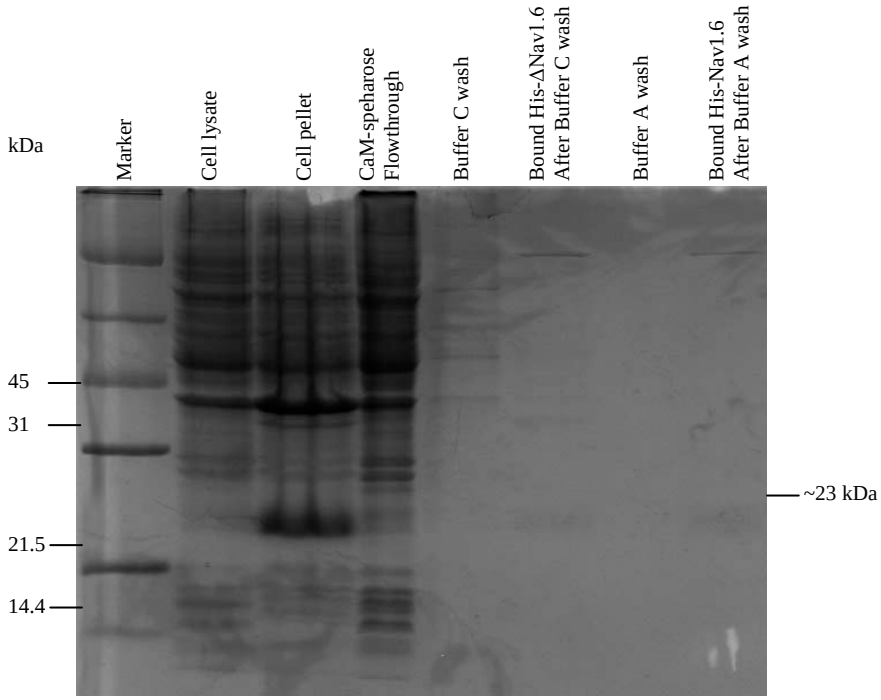


Fig. 3.6. 12.5% SDS-PAGE purification profile of His-ΔNav1.6 trapped by CaM-sepharose in (A) high Ca^{2+} concentration and (B) low Ca^{2+} concentration, visualized by Coomassie staining. The presence of Ca^{2+} enhances

binding affinity of His-ΔNav1.6 for CaM-speharose, shown by the ability of the latter to retain His-ΔNav1.6 in buffer A but not buffer C.

Purified His-ΔNav1.6 was further incubated with Ca^{2+}/CaM in a 1:1 molar ratio and applied to Superdex-G200. The complex was eluted as a single peak at 83mL, which corresponds to a 1:1 His-ΔNav1.6: CaM complex ratio that is approximately 40kDa in size (Fig. 3.7).

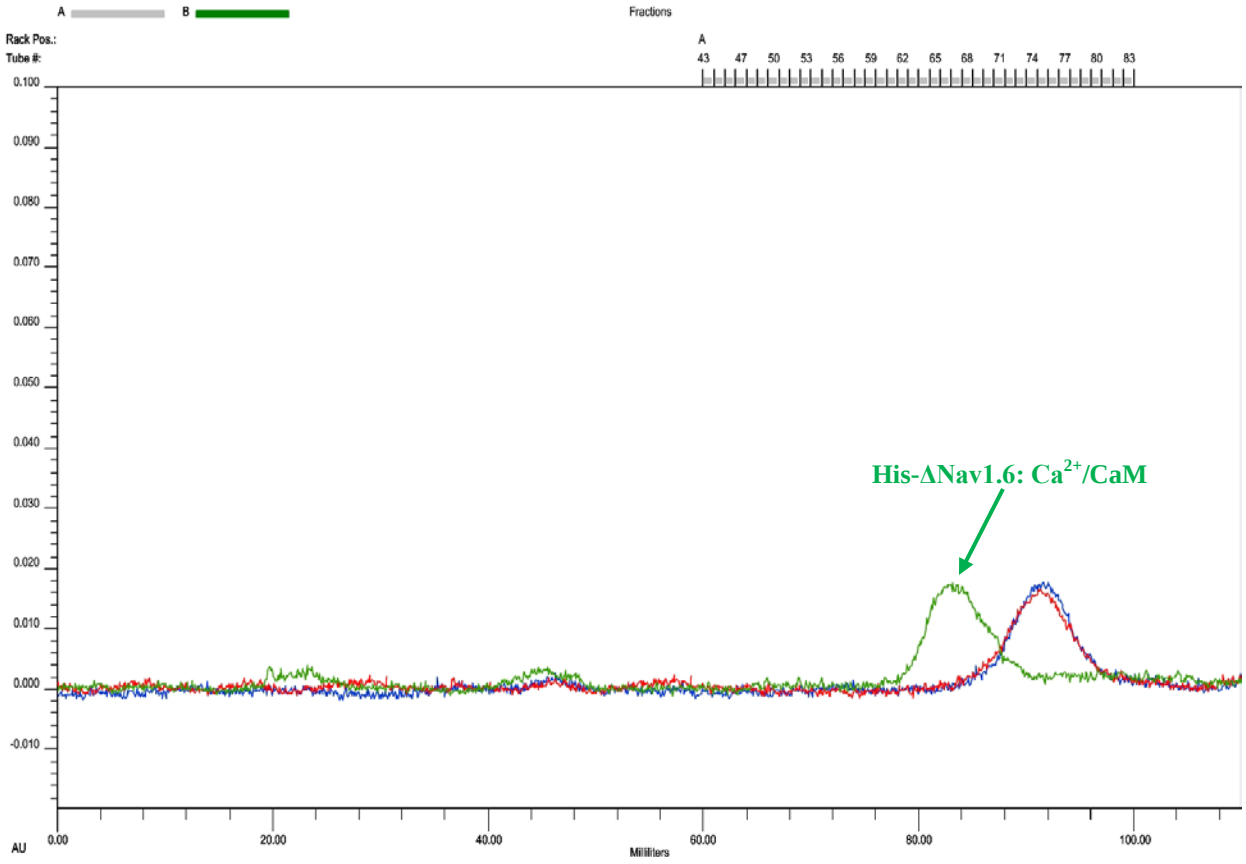


Fig. 3.7. Superimposed elution profiles of His-ΔNav1.6 (red), Ca^{2+}/CaM (Blue) and His-ΔNav1.6:Ca^{2+}/CaM complex (green) from Superdex-G200. His-ΔNav1.6 was mixed in a 1:1 ratio with Ca2+/CaM in buffer A + 100mM NaCl and 10mM βME prior to loading. Peak shift of about 10mL clearly indicates complex formation.

3.3.5 Isothemal titration calorimetry reveals that binding affinity is stronger in the presence of Ca^{2+} and NaCl

Unfortunately, binding affinity of CaM for ΔNav1.6 by ITC was unobtainable as the domain was insoluble at high concentrations required for ITC experiments. The binding affinities and influence of NaCl on VGSC isoforms IQ motif for CaM were therefore further determined by titrating the interacting peptides derived from Nav1.4IQ and Nav1.6IQ against 50μM purified CaM in the presence and absence of NaCl, along with high and low concentrations of Ca^{2+}.

Of the two peptides examined, Nav1.6IQ bound to CaM in a 1:1 ratio at a single binding site. The binding affinity of Nav1.6IQ for apoCaM and Ca^{2+}/CaM is the same in the absence of NaCl (Fig. 3.6B) but Nav1.6IQ exhibits approximately seven-fold stronger binding to Ca^{2+}/CaM compared to apoCaM and a more negative Gibbs free energy change in the presence of 100mM NaCl. These observations may be attributed to the differences in the structure of monomeric apoCaM and Ca^{2+}/CaM (Fig. 1.8; Vetter and Leclerc, 2003). However, when compared to its affinity in the absence of NaCl, binding of the IQ motif is still weaker than in the presence of NaCl.

Unlike Nav1.6IQ, the association of Nav1.4IQ for CaM was more complex. Binding kinetics for apoCaM and Ca^{2+}/CaM changed drastically in the presence and absence of NaCl. apoCaM bound to Nav1.4IQ in a 1:1 ratio that was readily fitted with a single site binding model only in the presence of NaCl, in a manner similar to Nav1.6IQ. In the absence of NaCl, the binding of apoCaM to Nav1.4IQ changed to adopt a complex pattern of two events that was reproducible and consistent over many experimental repeats – (1) a high affinity, exothermic reaction largely driven by enthalpy of 1:1 binding ratio, and (2) a lower affinity, endothermic and

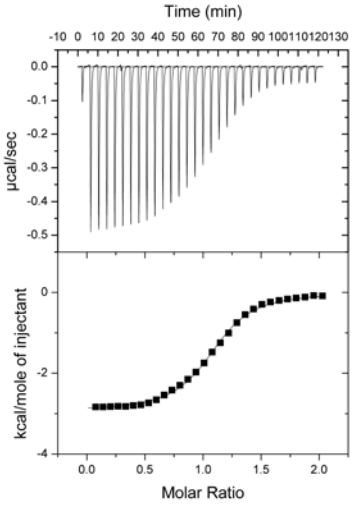
entropy driven reaction where only a third of the CaM sites were occupied, possibly due to the trimeric nature of CaM without NaCl. Notably reminiscent of Nav1.6IQ, the second yielded affinities and Gibbs free energy changes that were not significantly different from Ca^{2+}/CaM: Nav1.4IQ without salt. We thus conclude that the second reaction is the association of Nav1.4IQ to apoCaM.

In the presence of NaCl, the Ca^{2+}/CaM: Nav1.4IQ reaction does not tend to saturation even when peptide to CaM ratio was increased during titration. This incomplete binding is indicated by the large ΔH pleateau (~ -0.15 ucal/s) that is not caused by the heat of dilution (~0.04ucal/s) (Fig. 3.8A and C). However, as with Nav1.6IQ, affinity calculations still indicate that Ca^{2+}/CaM: Nav1.4IQ binding is tighter than apoCaM/Nav1.4IQ in the presence of NaCl. In the absence of NaCl, incomplete binding was also present but not as pronounced for Ca^{2+}/CaM: Nav1.4IQ.

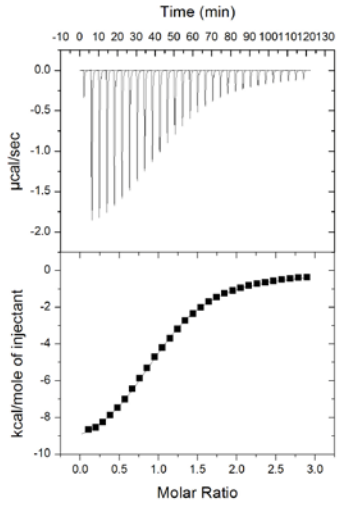
From the above discussed ITC results, the following conclusions can be drawn: (1) when NaCl is present, both peptides bind more strongly to Ca^{2+}/CaM than to apoCaM, (2) when NaCl is absent, the binding affinity of Nav1.6IQ for apoCaM and Ca^{2+}/CaM is the same, but (3) the binding dynamics of Nav1.4IQ changes. The ability of NaCl to influence binding indicates that hydrogen bonds or charge-charge interactions play an important role in the interactions between VGSC isoforms and CaM.

(A) ITC in 100mM NaCl (monomeric CaM)

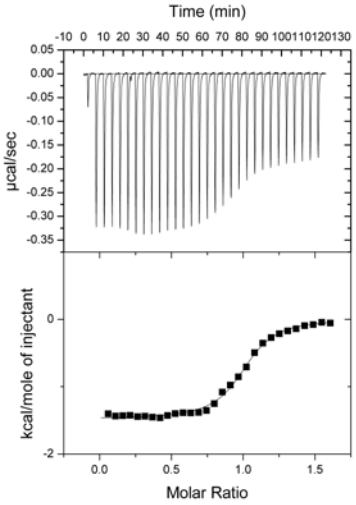
Cell	Ligand	N	Ka ($x10^6$ M^{-1})	Kd (μM)	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)
ApoCaM	Nav1.4IQ	1.08 ± 0.0043	0.650 ± 0.032	1.53	- 2.937 ± 0.016	4.899	-7.836
	Nav1.6IQ	1.03 ± 0.0042	0.121 ± 0.0025	8.26	-10.320 ± 0.058	-3.453	-6.867
Ca^{2+}/CaM	Nav1.4IQ	0.994 ± 0.0048	1.690 ± 0.14	0.592	- 1.480 ± 0.010	6.936	-8.416
	Nav1.6IQ	1.06 ± 0.0032	0.858 ± 0.033	1.17	-11.540 ± 0.047	-3.512	-8.028



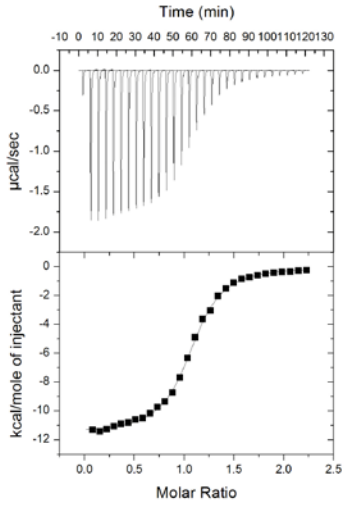
apoCaM:Nav1.4IQ



apoCaM:Nav1.6IQ



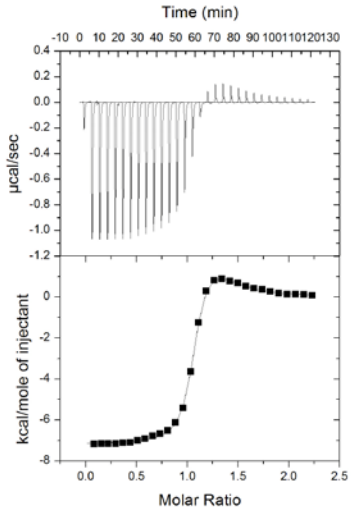
Ca^{2+}/CaM:Nav1.4IQ



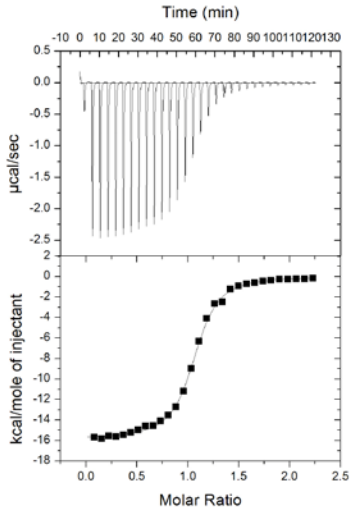
Ca^{2+}/CaM:Nav1.6IQ

(B) ITC in no NaCl (trimeric CaM)

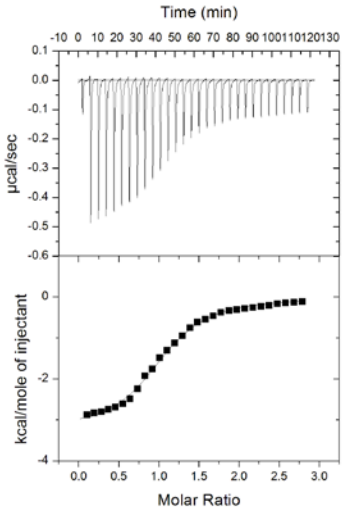
Cell	Ligand	N	Ka ($x10^6$ M^{-1})	Kd (μM)	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)
ApoCaM	Nav1.4IQ	N1: 1.03 ± 0.0052	K1: 23.0 ± 8.2	K1: 0.043	H1:-7.165 ± 0.021	TS1: 2.780	G1: -9.945
		N2: 0.358 ± 0.057	K2: 0.238 ± 0.068	K2: 4.20	H2: 2.842 ± 0.63	TS2: 10.094	G2: -7.252
	Nav1.6IQ	1.04 ± 0.0032	1.58 ± 0.083	0.633	-15.81 ± 0.071	-7.438	-8.372
Ca^{2+}/CaM	Nav1.4IQ	1.05 ± 0.011	0.248 ± 0.017	4.032	- 3.26 ± 0.046	4.044	-7.301
	Nav1.6IQ	1.06 ± 0.0028	1.58 ± 0.054	0.633	-11.18 ± 0.041	-2.807	-8.373



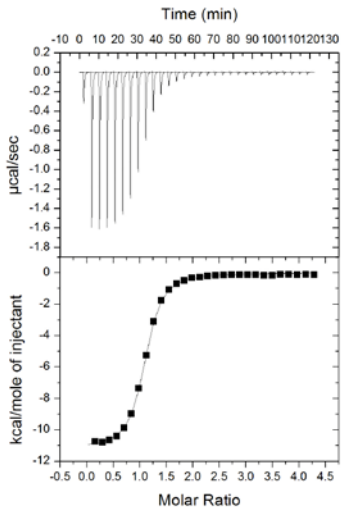
apoCaM:Nav1.4IQ



apoCaM:Nav1.6IQ



Ca^{2+}/CaM:Nav1.4IQ



Ca^{2+}/CaM:Nav1.6IQ

.

(C) ITC control by titrating Nav1.4IQ into buffer A + 100mM NaCl

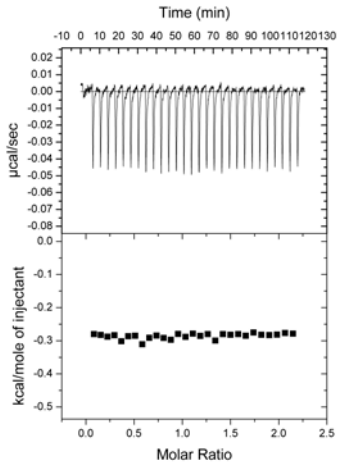


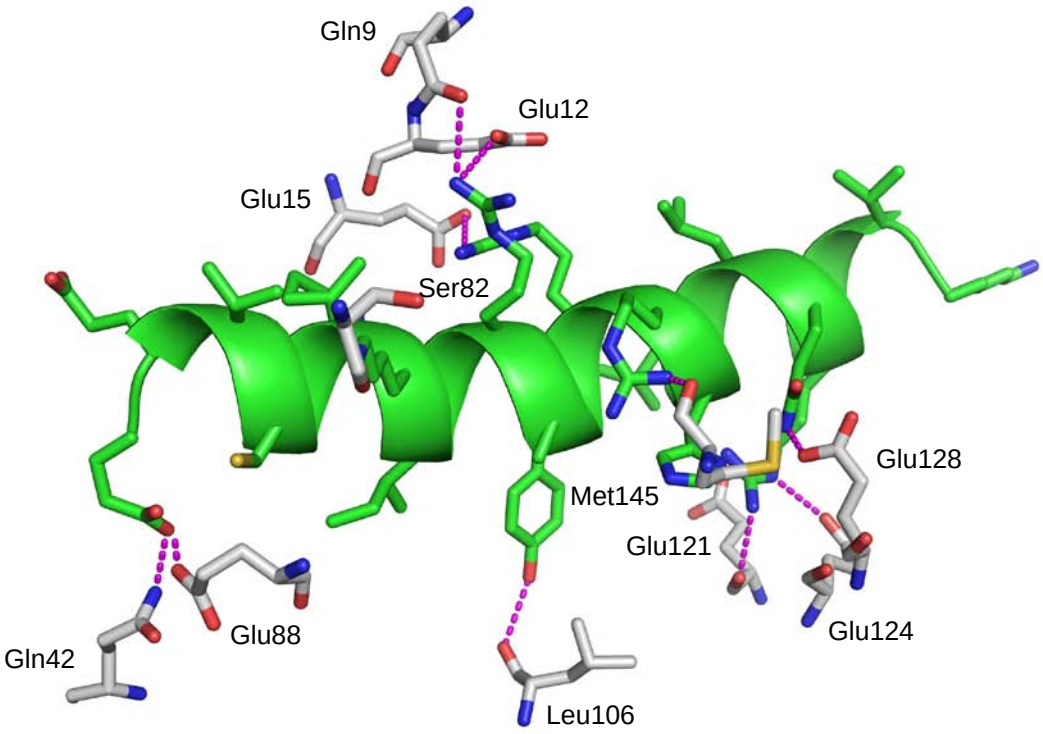
Fig. 3.8. The binding affinities of Nav1.4IQ and Nav1.6IQ peptides to CaM were measured by ITC (A) in 100mM NaCl, (B) without NaCl and (C) without CaM. The profiles are shown respectively. The binding constants (K_a and K_d), number of binding sites (N), enthalpy (ΔH) and entropy (ΔS) changes of CaM to the various peptides are provided in the tables. Figures below the tables show the injection profile after baseline correction and the bottom panels show the integration (heat release) for each injection (except the first one). The solid lines in the bottom panel show the fit of the data to a function based on a one-site binding model. Nav1.4IQ was titrated against buffer A + 100mM NaCl as a negative control and for subtraction of the heat of dilution. The concentration of the CaM was 50µM in all experiments.

3.3.6 Computational modelling and model verification by ITC

Motifs that interact with CaM generally assume an α-helix upon binding (Rhoads and Friedberg, 1997; Osawa *et al.*, 1999). Based on structural and computational analyses of IQ motifs from myosins and VGCCs in complex with CaM, it has been shown that the preferred conformation of the IQ motif is a seven-turn α-helix when bound to CaM (Hamilton *et al.*, 2000; Fallon *et al.*, 2005; Houdusse *et al.*, 2006). By combining literature and sequence analysis, we chose the 2.0Å structure of VGCC isoform 1.2 in complex with Ca^{2+}/CaM (PDB code 2BE6) as the starting template to model Nav1.4 and Nav1.6 IQ motif peptide complexes with Ca^{2+}/CaM. It is noteworthy that the VGCC shares the highest sequence similarity to the IQ motifs of Nav isoforms, and is a representative of an IQ motif from ion channel receptors. The structures of

Nav1.4IQ and Nav1.6IQ in complex with Ca^{2+}/CaM were modelled and energy minimised (Fig. 3.9).

(A) Nav1.4IQ: Ca^{2+}/CaM



(B) Nav1.6IQ: Ca^{2+}/CaM

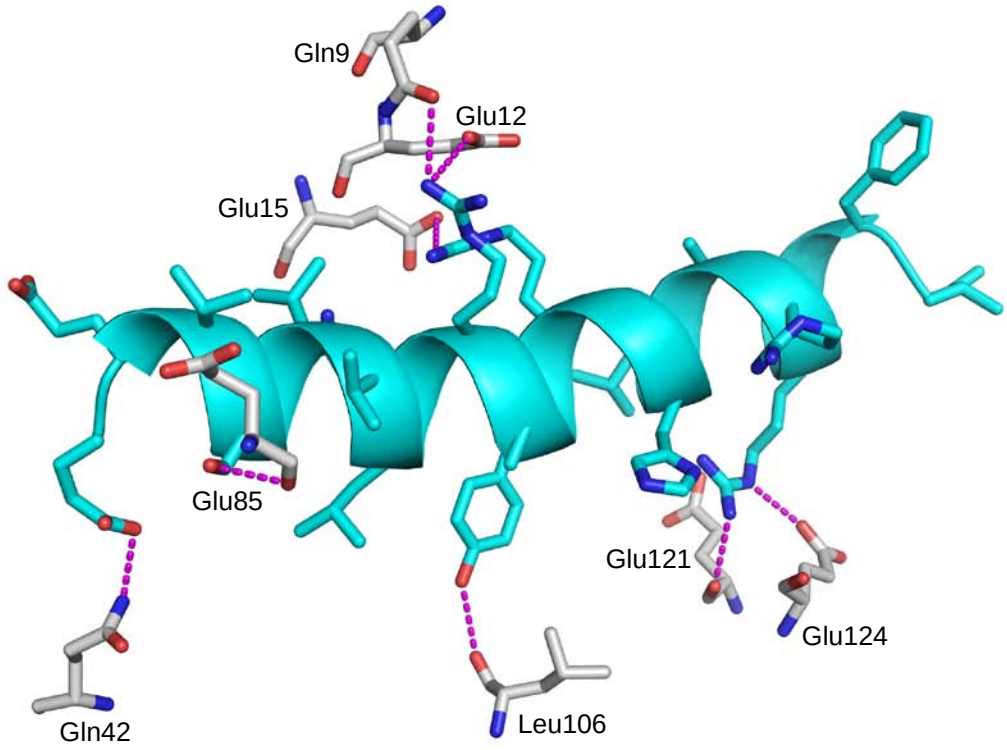
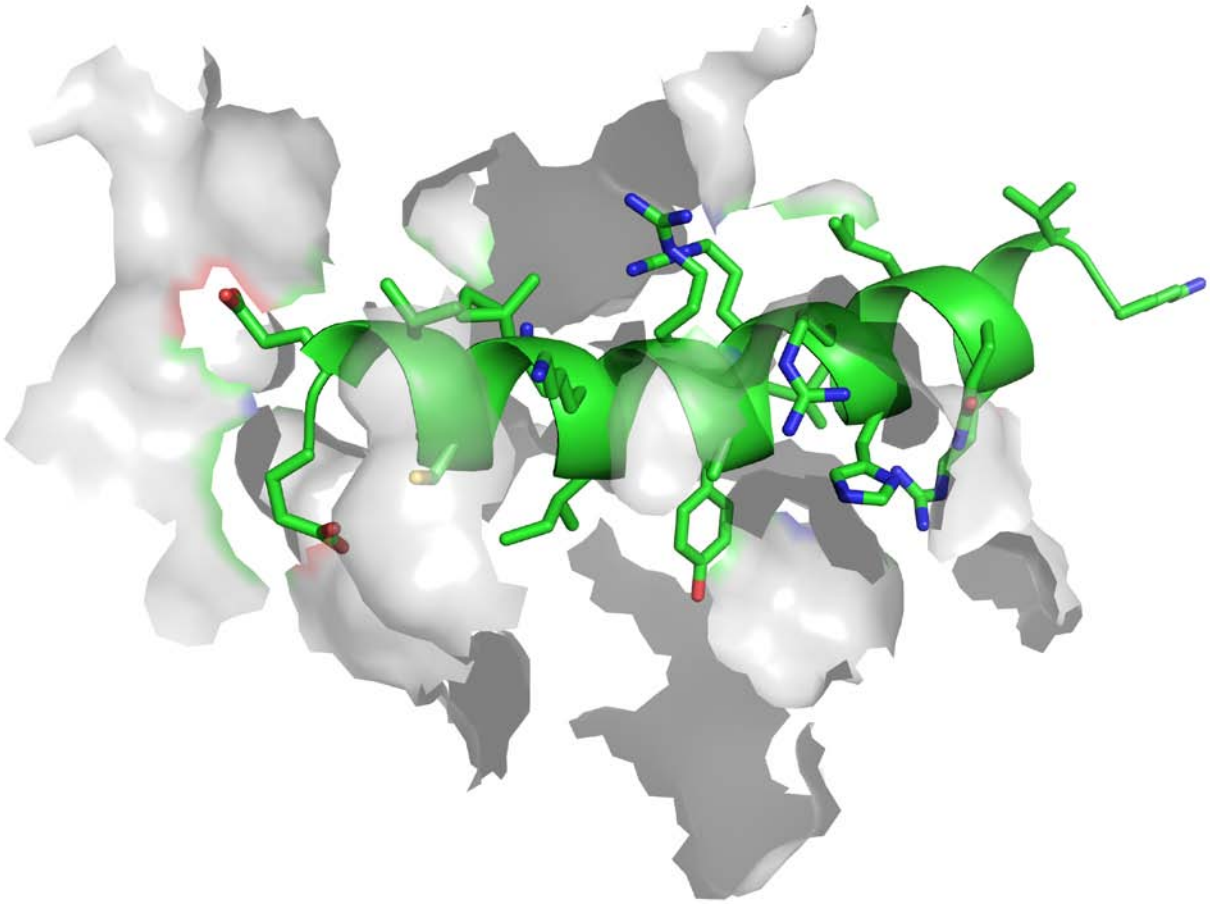


Fig. 3.9. The inferred modelled structures of (A) Nav1.4IQ and (B) Nav1.6IQ with Ca^{2+}/CaM showing potential hydrogen bonds that may govern binding. CaM residues participating in hydroen bond interactions are grey coloured and labelled. Hydrogen bonds are depicted as magenta dotted lines and the IQ motifs of Nav1.4 and 1.6 in green and cyan respectively. Figure was prepared using Pymol (DeLano, 2002).

From the inferred models, most of the residues in the substrate binding cleft of CaM interact with both Nav1.4 and Nav1.6 peptides. In particular, the side chains of Glu12, Glu88, Glu124 and Glu128 are in the vicinity of the peptide to form potential hydrogen bonding contacts (Fig. 3.9). The IQ motifs also make extensive hydrophobic interactions with CaM (Fig. 3.10).

(A) Nav1.4IQ: Ca^{2+}/CaM



(B) Nav1.6IQ: Ca^{2+}/CaM

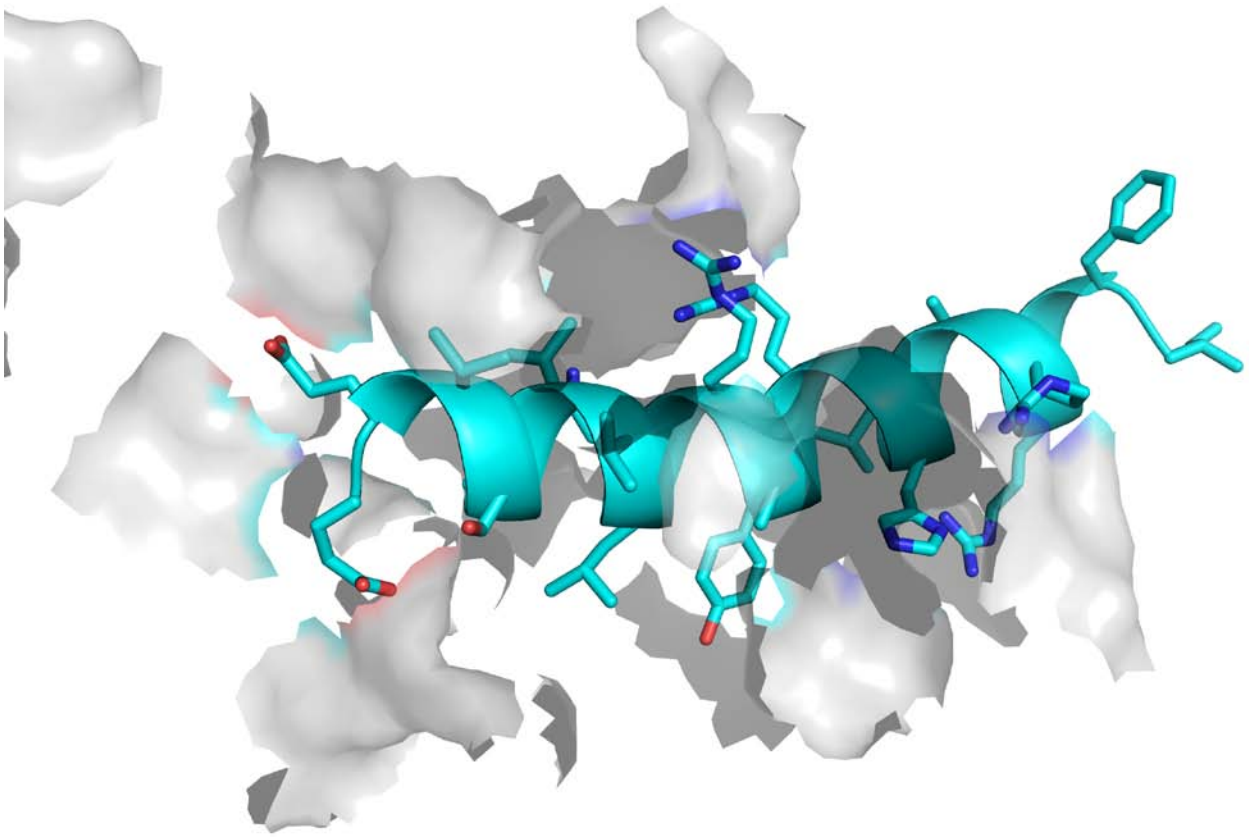
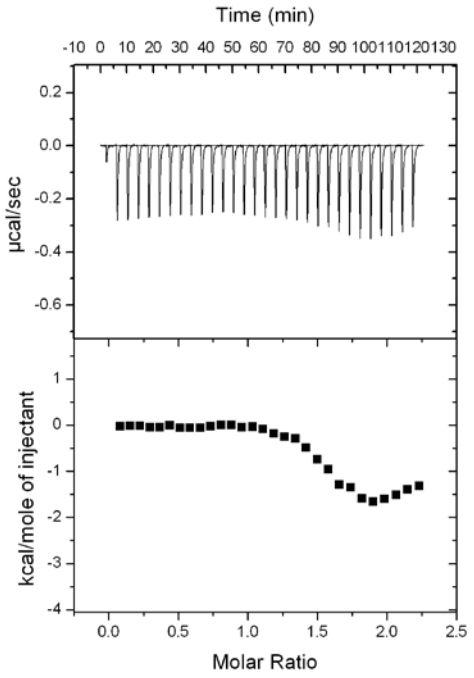
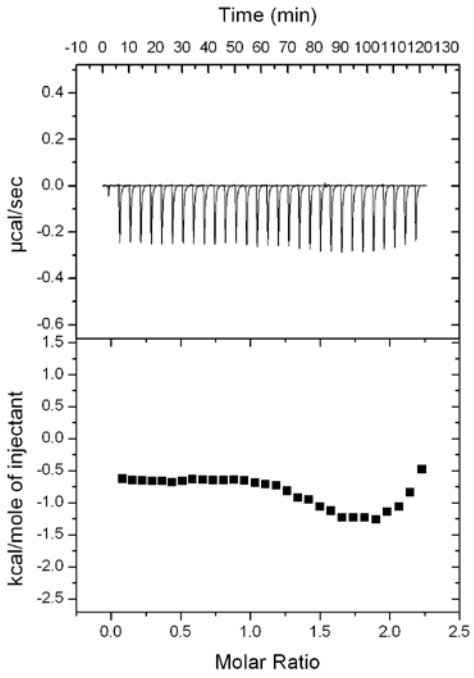


Fig. 3.10. Hydrophobic patches on CaM that form stacking interactions with (A) Nav1.4IQ and (B) Nav1.6IQ interacts with. Hydrophobic patches are shown as transparent white surfaces. The Nav1.4IQ peptide is green and the Nav1.6IQ peptide is cyan. Figure was prepared by Pymol (DeLano, 2002).

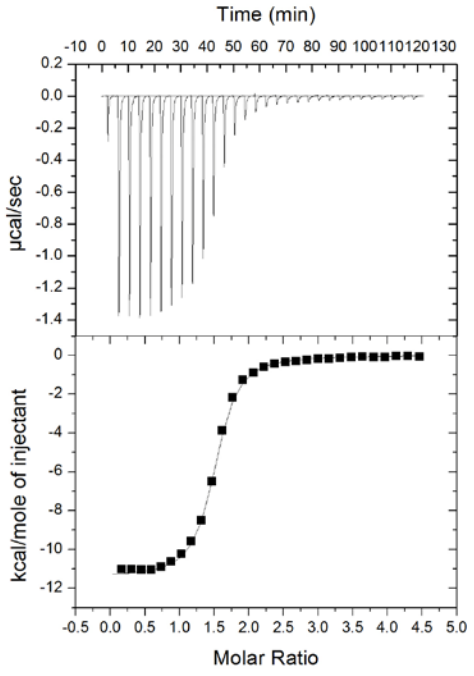
To verify that the CaM: peptide models, two residues on CaM (E88, E128) whose side-chains can potentially form hydrogen bonding contacts with Nav1.4IQ but not with Nav1.6IQ were chosen and mutated to Ala. ITC results showed an alteration in the binding profiles of Nav1.4IQ for the mutants (Fig. 3.11A and B). However, no significant change in binding affinity was observed with the Nav1.6IQ peptide (Fig. 3.11 B and C).



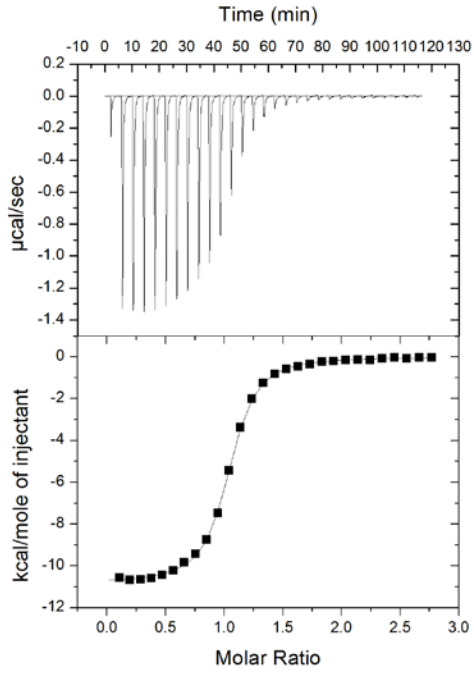
(A) Nav1.4IQ: Ca^{2+}/E88ACaM



(B) Nav1.4IQ: Ca^{2+}/E128ACaM



(C) Nav1.6IQ: Ca^{2+}/E88ACaM



(D) Nav1.6IQ: Ca^{2+}/E128ACaM

Fig. 3.11. The binding affinities of Nav1.4IQ and Nav1.6IQ peptides to CaM mutants were measured by ITC in bffer A. (A) Nav1.4IQ: Ca^{2+}/E88ACaM, (B) Nav1.4IQ: Ca^{2+}/E128ACaM, (C) Nav1.6IQ: Ca^{2+}/E88ACaM, (D) Nav1.6IQ: Ca^{2+}/E128ACaM. The injection profiles are shown respectively baseline correction. Bottom panels show

the integration (heat release) for each injection (except the first one). The solid lines in the bottom panel show the fit of the data to a function based on a one-site binding model. The concentration of CaM was 50μM in all experiments.

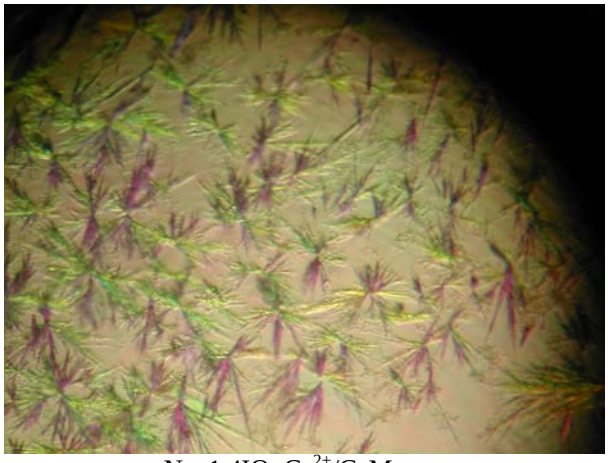
Our ITC studies on the CaM mutants clearly indicate that the inferred models are reliable. The observed change in the binding affinity with Nav1.4IQ could be a result of disrupting the formation of potential H-bonds between the peptide and CaM. However, Nav1.6IQ still binds with CaM because Glu88 and Glu128 do not participate in any hydrogen bonding contact with CaM. In order to further validate our results, we are currently in the process of obtaining co-crystal structures of these complexes.

3.3.7 Crystallization

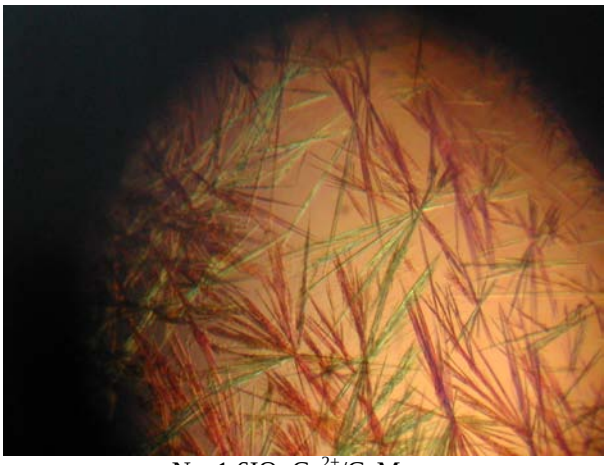
In order to confirm the Nav1.4IQ: and Nav1.6IQ: Ca^{2+}/CaM inferred models and elucidate their mechanism action, the IQ motif peptides are presently being co-crystallised with CaM. Complex crystal structures will help in understanding the differences in binding dynamics between the VGSC isoforms for CaM. To this end, Ca^{2+}/CaM was concentrated to 18mg/ml and incubated with two-fold molar excess of Nav1.4IQ and Nav1.6IQ peptides separately. Ca^{2+}/CaM was also incubated with His-ΔNav1.6 in a 1:1 molar ratio and concentrated to a maximum concentration of 5mg/ml. The crystallisation screen was performed by mixing 1μl protein with 1μl mother liquor and equilibrated against 500μl reservoir solution using hanging drop vapour diffusion method. No crystals were obtained from commercial screen kits such as Hampton Research, Qiagen and Jena Biosciences. Analysis of all the crystallisation drops showed that the protein remained soluble in most 2-Methyl-2,4-pentanediol (MPD) conditions. When compared with the conditions of previously crystallised CaM complexes, the MPD concentrations in commercial screen kits were at least 20% lower. Since the calculated pI of

CaM is approximately 4.5, the CaM: Nav1.4 and Nav1.6 complexes were further screened in specially prepared conditions with pH varied from 3.0 to 6.5 and MPD concentrations between 45 to 65% v/v.

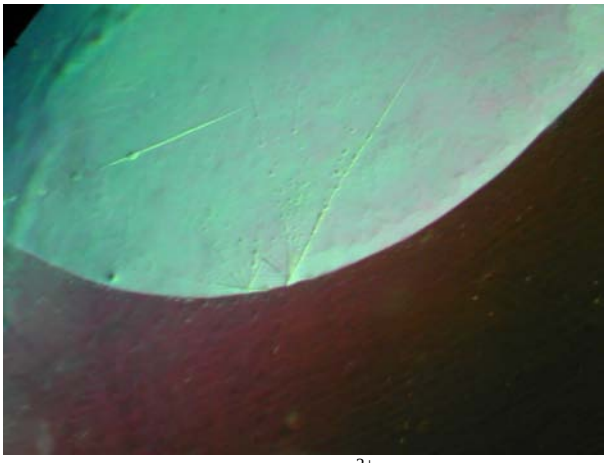
Initially, needle shaped crystals of Nav1.4IQ, Nav1.6IQ: Ca^{2+}/CaM and His-ΔNav1.6: Ca^{2+}/CaM were obtained from a condition containing 0.1M sodium acetate pH 4.5, 65% MPD and 20mM $CaCl_2$ at 4°C. These crystals were too small to be diffracted (Fig 3.12). Further optimisation marginally improved the crystal quality. Nav1.4IQ: Ca^{2+}/CaM complex crystals grew in 0.1M Bis-Tris pH 5.5, 65% MPD and 20mM $CaCl_2$ (Fig. 3.13) and is currently being optimised.



Nav1.4IQ: Ca^{2+}/CaM

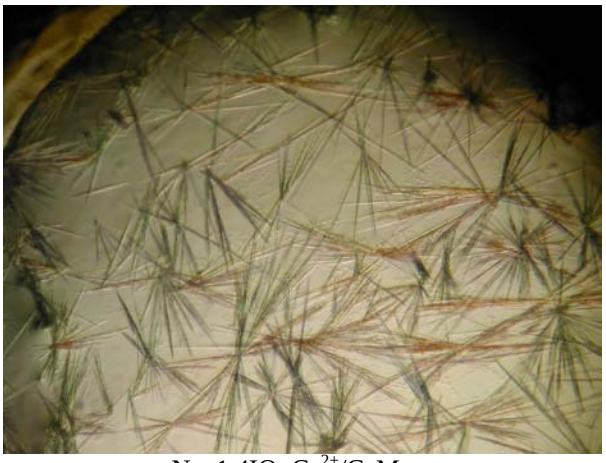


Nav1.6IQ: Ca^{2+}/CaM

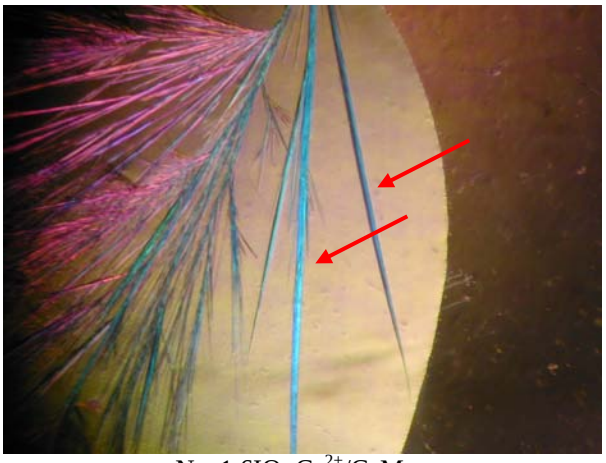


His-ΔNav1.6: Ca^{2+}/CaM

Fig. 3.12. Initial crystals from the screening of Nav1.4IQ: Ca^{2+}/CaM, Nav1.6IQ: Ca^{2+}/CaM and His-ΔNav1.6: Ca^{2+}/CaM complexes obtained using hanging drop vapour diffusion method.



Nav1.4IQ: Ca^{2+}/CaM



Nav1.6IQ: Ca^{2+}/CaM

Fig. 3.13. Current crystals of Nav1.4IQ: Ca^{2+}/CaM and Nav1.6IQ: Ca^{2+}/CaM obtained after grid optimization by the hanging drop vapour diffusion method. Crystals of Nav1.4IQ: Ca^{2+}/CaM could not be diffracted; those of Nav1.6IQ: Ca^{2+}/CaM as indicated by the red arrows were used for diffraction. These crystals were cryo-protected using mother liquor and flash frozen in liquid nitrogen before data was collected.

Similarly, crystals of the Nav1.6IQ: Ca^{2+}/CaM complex were optimised with a reservoir solution of 0.1M sodium acetate pH 4.5, 55% MPD, 20mM $CaCl_2$ and 8% n-propanol. These crystals diffracted to 3.3Å̊ (Fig. 3.15), and a partial dataset was collected. This crystallisation condition currently being optimised to improve crystal quality. To verify whether the presence of CaM: Nav1.6IQ peptide complex in the crystals, several crystals were collected from drops

and extensively washed in reservoir solution prior to its separation in a 16% SDS-PAGE under denaturing conditions. The crystals migrated as two bands, one corresponding to CaM's molecular weight of 16.8kDa and the other band smaller than the 6kDa band of a protein molecular weight standard (Biorad), presumably the Nav1.6IQ peptide which is approximately 3kDa (Fig. 3.14). Smearing of the lower molecular weight bands in the protein marker and sample lanes are most likely to be caused by the low percentage gel for this peptide. Higher percentage gels could not be used, since it would affect the migration of calmodulin. The presence of these two bands clearly indicates that the crystals obtained are the Nav1.6IQ: Ca^{2+}/CaM complex.

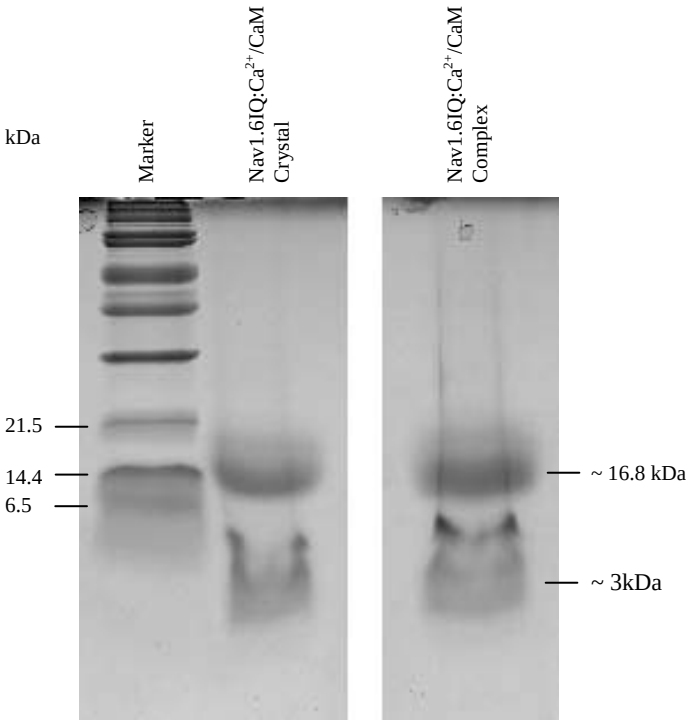


Fig. 3.14. 16% SDS-PAGE profile of Nav1.6IQ: Ca^{2+}/CaM crystals. The uncrystallised complex was included in the third lane as a comtrol.

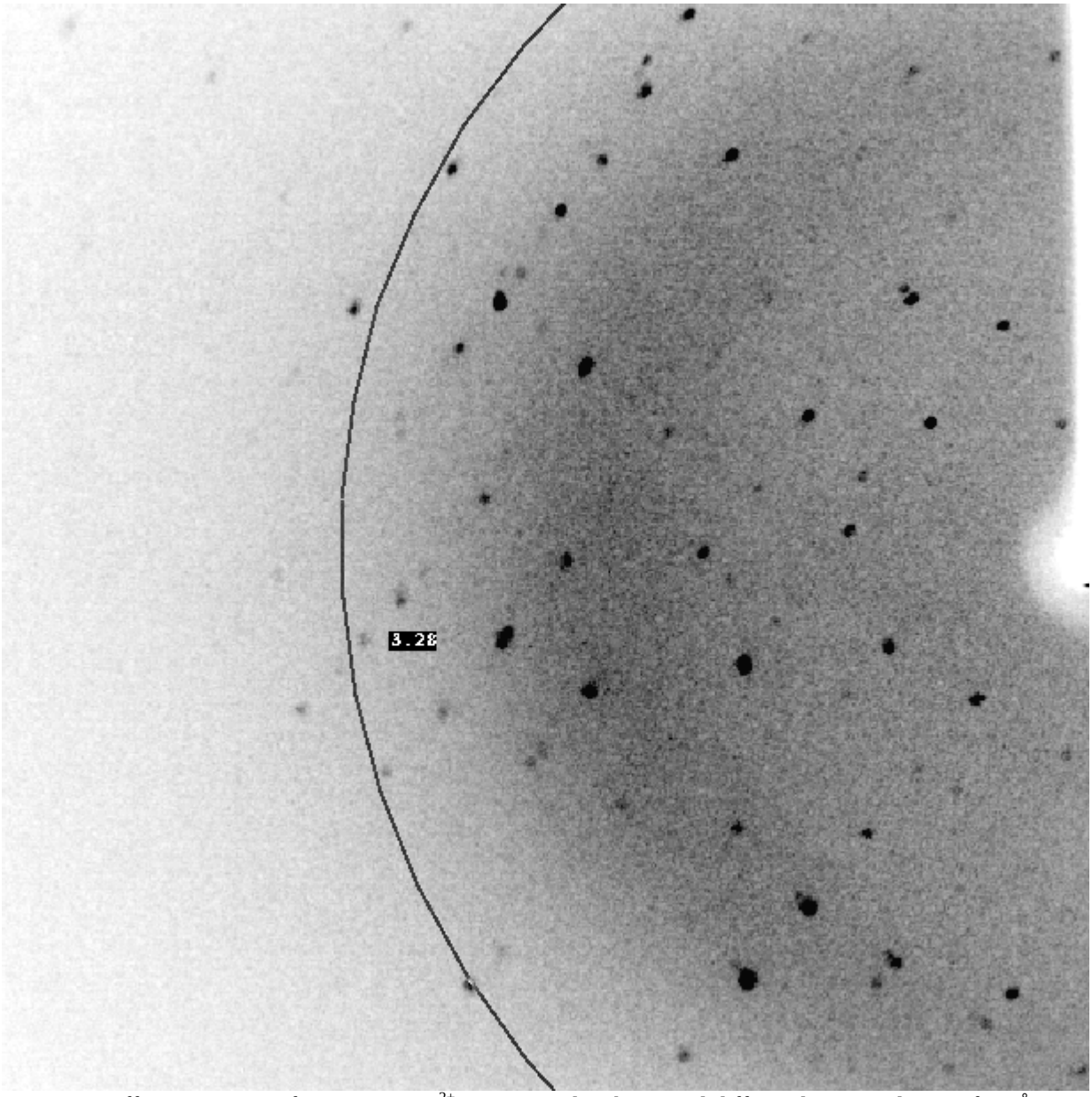


Fig. 3.15. Diffraction image of Nav1.6IQ: Ca^{2+}/CaM crystal. The crystal diffracted to a resolution of 3.3Å. Data was collected using Platinum135 CCD detector mounted on a Microstar-H rotating anode generator (Bruker, DE) given a 20s exposure.

The crystals are currently being optimised to collect complete datasets.

3.4 Discussion

Most of the IQ motifs crystallised with Ca^{2+}/CaM to date belong to the VGCC isoforms. Athough the importance of calcium regulation as a feedback mechanism in VGCC is clear; it is less obvious why Ca^{2+} should play a part in the sodium current of the action potential. The importance of Ca^{2+} regulation, underscored by the fact that all VGSC isoforms possess IQ motifs, is ironic because only certain isoforms exhibit strong affinity for CaM (Herzog *et al.*, 2002). Each VGSC isoform also has their activity differently regulated by CaM although sharing sequence similarity to a large extent (Herzog *et al.*, 2003). In this study on two IQ domains sharing high sequence identity, namely skeletal muscle VGSC isoform Nav1.4 and neuronal VGSC isoform Nav1.6 (percentage identity and similarity between IQ motifs of the two isoforms is 54% and 80% respectively), we aimed to uncover, if any, (1) the fundamental binding mechanism by which VGSC IQ motifs bind to CaM but retain flexibility to interact in different ways, and (2) whether the differences in binding affinities between isoforms represents the need for Ca^{2+} recognition during the action potential.

The ability of IQ motifs to bind CaM in several ways suggests that each protein recruits and utilizes CaM in a unique way. This is exemplified by the difference in binding dynamics between Nav1.4IQ and Nav1.6IQ. Alignment of Nav1.4IQ, Nav1.6IQ and the IQ motif consensus sequences indicate that amino acids differences between isoforms may be responsible for differences in the binding modes (Fig. 3.16). Unlike Nav1.6 that possesses all the conserved residues of the IQ motif, Nav1.4IQ does not possess a moderately conserved Gly residue in seventh position, or the highly conserved Arg/Lys residue at 11^{th} position. Non conserved residues of both isoforms share high sequence identity in amino acids up to the 12^{th} position.

Fig. 3.16. Sequence alignment of previously identified VGSC isoforms Nav1.4 and Nav1.6 IQ motifs. Coloured residues refer to the conserved residues in the sequence; residues coloured red are strictly conserved across IQ motifs, the moderately conserved G at seventh position is coloured green, while blue residues are those that are conserved between Nav1.4 and Nav1.6 isoforms but not part of an IQ motif's identity.

From the inferred model, the aliphatic part of the side chain of Arg/Lys11 is important for forming extensive hydrophobic interactions with CaM. However, the similar aliphatic part of Glu11 side chain in Nav1.4IQ does not engage with CaM and is not involved in stabilising the C-terminal part of the IQ motif. The conserved Gly7 is a surface expose residue in Nav1.6IQ. The equivalent residue in Nav1.4IQ is Arg7 which folds back to occlude a part of the CaM surface which is exposed in the presence of Gly7 instead. The residues of Nav1.6IQ fits more comfortably into the Ca^{2+}/CaM surface compared with Nav1.4IQ (Fig. 3.17). All of the above observations may have contributed to maintaining a stable Nav1.4IQ:apo/CaM complex in the absence of NaCl, and also to the unsaturated binding of Nav1.4IQ:Ca^{2+}/CaM as discussed in the previous section.

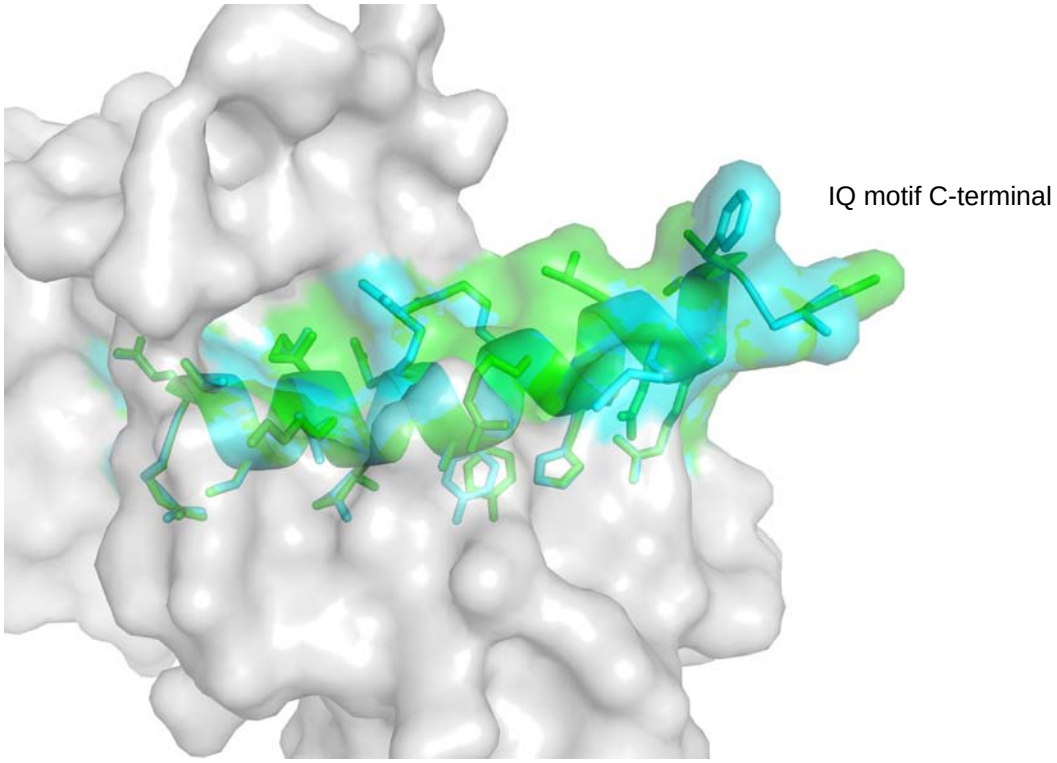


Fig. 3.17. Superimposition of the modelled Nav1.4IQ (green) and Nav1.6IQ (cyan) peptides on to CaM surface (grey). The C-terminal of Nav1.4IQ protrudes from the CaM surface to a greater extent than Nav1.6IQ, as shown by the predominance of green surfaces.

IQ motifs that lack a conserved basic residue in the 11th position have been reported to require Ca^{2+} for binding (Houdusse and Cohen, 1995; Munshi *et al*., 1996). This conserved basic residue is present in Nav1.6 but absent in Nav1.4. Indeed in conditions without NaCl, apoCaM and Ca^{2+}/CaM was observed to have identical affinity for Nav1.6 but the binding of Nav1.4 was drastically different in apoCaM and Ca^{2+}/CaM. These differences may be due to the ability of Nav1.6, but not Nav1.4, to stabilise itself on the CaM surface through hydrophobic interaction as described above (Fig. 3.17). However, the molecular details of these interactions can only be fully understood from crystal structures of these complexes.

Although both isoforms share a high sequence similarity in the IQ motif, these results indicate a hypersensitivity of CaM towards residue differences in the IQ motif of the two

isoforms which may explain why and how highly similar sequences' binding and regulation by CaM are different. Of course, the influence of residues flanking the IQ motif in CaM interaction cannot be ruled out.

We propose that in tissues where other CaM target proteins are located e.g. in the brain, there is competitive sequestration of CaM. Proteins that require calcium regulation in certain states increase their affinities over others for CaM momentarily until a time when a signal has passed, such as the increase in affinity of the VGSCs for CaM in the presence of Ca^{2+}. However, proteins that constitutively require CaM for function bind regardless of the Ca^{2+} concentration. High affinity VGSCs, such as Nav1.4 and Nav1.6 may represent a class of ion channels that require a Ca^{2+} sensor at basal level in resting state, but increasing Ca^{2+} causes a positive feedback that increases channel sensitivity for calcium through the increased sequestration of CaM. These proposals however, await to be confirmed by further structural and physiological studies.

Chapter IV

Conclusions and Future Directions

A recurring and fundamentally important theme in the study of signal transduction pathways is protein-protein interactions. Many diseased states are caused by deregulation of signalling pathways as a result of abnormal interaction and activating or inhibiting mutations in the core components of these pathways. In order to understand the essential mechanisms through which regulatory proteins perform their function, two examples of receptor protein regulators from the ERK/MAPK pathway were chosen – c-Cbl and CaM with their respective substrates of tyrosine kinases and VGSCs.

In this thesis, the binding dynamics of both systems were thoroughly described. In the case of c-Cbl, five crystal structures of Cbl-TKB complexed with the different tyrosine kinase substrates and another regulator of the ERK/MAPK pathway, Spry2, has illustrated the mode of the binding with the TKB domain. The Cbl-TKB: Met structure also represents the highest resolution Cbl atomic model to date. Elucidation of the Cbl-TKB complex structures has unexpectedly revealed the presence of an essential intrapeptidyl H-bond between pTyr and its neighbouring residue which binds to Cbl-TKB regardless of the orientation of the sequence.

The reverse binding mode of Met, together with its ability to bind to Cbl-TKB when the DpYR motif was reversed, suggests the likelihood of other Cbl targets yet to be discovered. Although not all potential targets are *in vivo* binders due to localization, temporal and steric factors, it is worth screening for such interactors on a proteomic scale, possibly through the use of microarray technology.

In depth understanding of the importance of substrate orientation in the context of Cbl role in signal transduction would require that structural characterization of Cbl-substrate complexes extend beyond peptides. Proteins harbouring the DpYR motif with a (pY-4)Pro can

be interesting candidates for future study, together with proteins whose regulation by Cbl is already established, such as EGFR and APS. Taken together, results from these analyses would greatly enhance the understanding of Cbl's mode of action as an adaptor or as a ubiquitin ligase – two distinctly different functions.

Biophysical and computational modelling data has provided interesting insights towards understanding how CaM regulates the VGSCs. However, the necessity of further work is duly acknowledged. The ITC results obtained may only be explained with a comparative study of the real molecular structures of both isoforms in complex with CaM at atomic resolution, an aim that we hope to achieve through X-ray crystallography. However, difficulty in producing diffraction quality single crystals even after extensive optimisation is the bottleneck. Nav1.4IQ: CaM would not diffract and the data collected from the Nav1.6IQ:CaM is not complete.

Improving current crystal quality is the first priority for the CaM: VGSC complexes project. As a final resort, nuclear magnetic resonance spectroscopy (NMR) will also be employed to determine the complex structures.

Two major questions were posed at the outset of this study – (1) What is the fundamental mechanism governing VGSC: CaM interaction, and (2) why do certain VGSCs require tight binding to CaM but not others. The first question may be ascertained once complex structures are known. Mutation of key residues in the full length receptor to examine its interaction with CaM may extend and confirm conclusions drawn from structural studies on the shorter IQ motif peptides. Given the present limited data, a reason to the second question may only be speculated from ITC observations. Resolving the second question requires the full understanding of the

CaM signalosome in different VGSC expressing tissues which may be feasible only through the use of high throughput proteomic technologies.

Chapter V

References

Andersen P, Sundberg SH, Sveen O, Swann JW, Wigstrom H (1980) Possible mechanisms for long lasting potentiation of synaptic transmission in hippocampal slices from guinea-pigs. J Physiol 302:463-482.

Avruch, J., Khokhlatchev, A., Kyriakis, J. M., Luo, Z., Tzivion, G., Vavvas, D., and Zhang, X. F. (2001). Ras activation of the Raf kinase: tyrosine kinase recruitment of the MAP kinase cascade. Recent Prog. Horm. Res. 56:, 127–155.

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 98: 10037–10041.

Bao J, Gur G, Yarden Y (2003) Src promotes destruction of c-Cbl: implications for oncogenic synergy between Src and growth factor receptors. PNAS 100:2438-2443.

Bergfors T (2003) Seeds to crystals. J Struct Biol 142:66-76.

Blake TJ, Shapiro M, Morse HC, Langdon WY (1991) The sequences of human and mouse c-cbl proto-oncogenes show v-cbl was generated by a large truncation encompassing a proline-rich domain and a leucine zipper like motif. Oncogene 6:653-657.

Bliss TV, Lomo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol 232:331-56.

Boxall AR and Lancaster B (1998) Tyrosine kinase and synaptic transmission. Eur J Neurosci 10:2-7.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D 54 (Part 5): 905–921.

Campbell SJ and Jackson RM (2003) Diversity in the SH2 domain family phosphotyrosyl peptide binding site. Protein engineering 16:217-227.

Catterall WA (1995) Structure and function of voltage-gated ion channels. Annu Rev Biochem 64:493-531.

Ceccarelli DF, Blasutig IM, Goudreault M, Li Z, Ruston J, Pawson T, Sicheri F (2007) Non-canonical interaction of phosphoinositides with pleckstrin homology domains of Tiam1 and ArhGAP9. J Biol Chem 282: 13864–13874.

Cestari V, Costanzi M, Castellano C, Rossi-Arnaud C (2006) A role for ERK2 in reconsolidation of fear memories in mice. Neurobiol Learn Mem. 86:133-143.

Chang LF and Karin M (2001) Mammalian MAP kinase signaling cascades. Nature 410:37-40.

Chen, C. Y. et al. Nucleolin and YB-1 are required for JNK-mediated interleukin-2 mRNA stabilization during T-cell activation. Genes Dev. 14, 1236±1248 (2000).

Chiang SH, Baumann CA, Kanzaki M, Thurmond DC, Watson RT, Neudauer CL, Macara IG, Pessin JE, Saltiel AR (2001) Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. Nature 410: 944–948.

Cho S, Velikovsky CA, Swaminathan CP, Houtman JC, Samelson LE, Mariuzza RA (2004) Structural basis for differential recognition of tyrosine-phosphorylated sites in the linker for activation of T cells (LAT) by the adaptor Gads. EMBO J 23: 1441–1451.

Cohen, P. The regulation of protein function by multisite phosphorylation—a 25 year update. Trends Biochem. Sci. 25, 596–601 (2000).

Cohen P (2002) The origins of protein phosphorylation. Nat Cell Biol 4:E127-E130.

DeLano WL (2002) The PyMOL Molecular Graphics System. San Carlos, CA, USA: DeLano Scientific.

Deschenes I, Neyroud N, DiSilvestre D, Marban E, Yue D, Tomaselli GF (2006) Isoform specific modulation of voltage-gated Na+ channels by calmodulin. Circ Res. 90:49-57.

English JD and Sweatt JD (1997) A requirement for the mitogen-activated protein kinase cascade in hippocampal long term potentiation. J Biol Chem 272:19103-19106.

Entingh AJ, Taniguchi CM, Kahn CR (2003) Bi-directional regulation of brown fat adipogenesis by the insulin receptor. J Biol Chem 278: 33377–33383.

Fallon JL, Halling D, Hamilton S, Quiocho FA (2005) Structure of calmodulin bound to the hydrophobic IQ domain of the cardiac Cav1.2 calcium channel. Structure 13:1881-1886.

Feng S, Chen JK, Yu H, Simon JA, Schreiber SL (1994) Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. Science 266: 1241–1247.

Feshchenko EA, Langdon WY, Tsygankov AY (1998) Fyn, Yes and Syk phophorylation sites in c-Cbl map to the same tyrosine residues that become phosphorylated in activated T cells. J Biol Chem 273:8323-8331.

Frese S, Schubert WD, Findeis AC, Marquardt T, Roske YS, Stradal TE, Heinz DW (2006) The phosphotyrosine peptide binding specificity of Nck1 and Nck2 Src homology 2 domains. J Biol Chem 281: 18236–18245.

Fong CW, Leong HF, Wong E, Lim J, Yusoff P, Guy G (2003) Tyrosine phosphorylation of Sprouty2 enhances its interaction with c-Cbl and is crucial for its function. J Biol Chem 278:33456-33464.

Gomperts BD (2003) Signal transduction. 3rd Ed pp283. Elsevier, USA.

Gottschalk WA, Jiang H, Tartaglia N, Feng L, Figurov A and Lu B (1999) Signaling mechanisms mediating BDNF modulation of synaptic plasticity in the hippocampus. Learn Mem, 6:243-256.

Gross I, Bassit B, Benezra M and Licht JD (2001) Mammalian Sprouty proteins inhibit cell growth and differentiation by preventing Ras activation. J Biol Chem 276:46460-46468.

Guex N, Peitsch MC (1997) Swissmodel and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18:2714-2723.

Gunther UL, Weyrauch B, Zhang X, Schaffhausen B (2003) Nuclear magnetic resonance structure of the P395S mutant of the N-SH2 domain of the p85 subunit of PI3 kinase: an SH2 domain with altered specificity. Biochemistry 42: 11120–11127.

Hacohen N, Kramer S, Sutherland D, Hiromi Y, Krasnow M (1999) Sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the Drosophila airways. Cell 92:253-263.

Hall AB, Jura N, DaSilva J, Jang YJ, Gong D, Bar-Sagi D (2003) hSpry2 is targeted to the ubiquitin-dependent proteasome pathway by c-Cbl. Curr Biol 13: 308–314.

Hamilton S, Serysheva I, Strasberg GM (2000) Calmodulin and excitation-contraction coupling. News Physiol Sci 15:281-284.

Hanafusa H, Torii S, Yasunaga T, Nishida E (2002) Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. Nat Cell Biol 4:850-858.

Hanks SK, Quinn AM, Hunter T (1991) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 241:42–52.

Herzog RI, Liu C, Waxman SG, Cummins TR (2003) Calmodulin binds to the C-terminus of sodium channels Nav1.4 and Nav1.6 and differentially modulates their functional properties. J Neurosci 23:8261-8270.

Houdusse A and Cohen C (1995) Target sequence recognition by the calmodulin superfamily: Implications from light chain binding to the regulatory domain of scallop myosin. Proc. Natl. Acad. Sci. USA 92:10644-10647.

Houdusse A Gaucher JF, Krementsova E, Mui S, Trybus KM, Cohen C (2006) Crystal structure of apo-camodulin bound to the first two IQ motifs of myosin V reveals essential recognition features. PNAS 103:19326-19331.

Hu J, Hubbard SR (2005) Structural characterization of a novel Cbl phosphotyrosine recognition motif in the APS family of adapter proteins. J Biol Chem 280: 18943–18949.

Hubbard SR and Till JH (2000) Protein tyrosine kinase structure and function. Annu Rev Biochem 69:373-398.

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z. and Dunker, A. K. (**2004**) *Nucleic Acids Res. 32*, 1037-1049.

Jaumot M and Hancock JF (2001) Protein phosphatases 1 and 2A promote Raf-1 activation by regulating 14-3-3 interactions. Oncogene 20:3949-58.

Jelinek T, Dent P, Sturgill TW, Weber MJ (1996) Ras induced activation of Raf-1 is dependent on tyrosine phosphorylation. Mol Cell Biol 16:1027-1034.

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 47 (Part 2): 110–119.

Kim HJ and Barsagi D (2004) Modulation of signaling by Sprouty: a developing story. Nat Rev Mol Cell Biol 5:441-450.

Kong-Beltran M, Seshagiri S, Zha J, Zhu W, Bhawe K, Mendoza N, Holcomb T, Pujara K, Stinson J, Fu L et al. (2006) Somatic mutations lead to an oncogenic deletion of Met in lung cancer. Cancer Res 66:283-289.

Kosako H, Gotoh Y, Matsuda S, Ishikawa M, Nishida E (1992) Xenopus MAP kinase activator is a serine/threonine/tyrosine kinase activated by threonine phosphorylation. EMBO J 11:2903-2908.

Kotlyarov, A. et al. MAPKAP kinase 2 is essential for LPS-induced TNF-a biosynthesis. Nature Cell Biol. 1, 94±97 (1999).

Langdon WL, Hartley JW, Klinken SP, Ruscetti SK, Morse HC (1989) v-Cbl, an oncogene from a dual recombinant retrovirus that induces early B-lineage lymphomas. PNAS 86:1168-1172.

Lao DH, Chandramouli S, Yusoff P, Fong CW, Saw TY, Tai LP, Yu CY, Leong HF, Graeme R. Guy (2006) A Src Homology 3-binding Sequence on the C Terminus of Sprouty2 Is Necessary for Inhibition of the Ras/ERK Pathway Downstream of Fibroblast Growth Factor Receptor Stimulation. J Biol Chem 281:29993-30000.

Lasa, M. et al. Regulation of cyclooxygenase 2 mRNA stability by the mitogen-activated protein kinase p38 signaling cascade. Mol. Cell. Biol. 20, 4265±4274 (2000).

Laskowski RA, MacAuthur MW, Moss DS, Thornton JM (1993) Procheck: a program to check the stereochemical quality of protein structures. J Appl Cryst 26: 283–291.

Levkowitz G, Waterman H, Ettenberg SA, Katz M, Tsygankov AY, Alroy I, Lavi S, Iwai K, Reiss Y, Ciechanover A, Lipkowitz S, Yarden Y (1999) Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1. Mol Cell 4: 1029–1040.

Lewis TS, Shapiro PS, Ahn NG (1998) signal transduction through MAP kinase cascades. Adv Cancer Res 74:49-139.

Liu BA, Jablonowski K, Raina M, Arce M, Pawson T, Nash PD (2006) The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. Mol Cell 22: 851–868.

Liu J, Kimura A, Baumann CA, Saltiel AR (2002) APS facilitates c-Cbl tyrosine phosphorylation and GLUT4 translocation in response to insulin in 3T3-L1 adipocytes. Mol Cell Biol 22: 3599–3609.

Lupher Jr ML, Songyang Z, Shoelson SE, Cantley LC, Band H (1997) The Cbl phosphotyrosine-binding domain selects a D(N/D)XpY motif and binds to the Tyr292 negative regulatory phosphorylation site of ZAP-70. J Biol Chem 272: 33140–33144.

Luttrell LM, Hawes BE, Biesen T, Luttrell DK, Lansing TJ, Lefkowitz RJ (1996) Role of c-Src Tyrosine Kinase in G Protein-coupled Receptor and Gβγ Subunit-mediated Activation of Mitogen-activated Protein Kinases. J Biol Chem 271:19443-19450.

Maisonpierre PC, Suri C, Jones PF, Bartunkova S, Wiegand SJ, Radziejewski C, Compton D, McClain J, Aldrich TH, Papaphosdopoulos N, et al. (1997). Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis. Science 277:48–50.

Mann M and Jensen ON (2003) Proteomic analysis of post-translational modifications. Nat Biotech 21:255-261.

Meng W, Sawasdikosol S, Burakoff SJ, Eck MJ (1999) Structure of the amino-terminal domain of Cbl complexed to its binding site on ZAP-70 kinase. Nature 398: 84–90.

Moodie SA, Willumsen BM, Weber MJ, Wolfman A (1993) Complexes of the Ras-GTP with Raf -1 and mitogen activated protein kinase kinase. Science 260:1658-1660.

Olwin BB and Storm DR (1985) Binding of calcium to calmodulin-protein complexes. Biochemistry 24:8081-8086.

Osawa M, Tokumitsu H, Swindells MB, Kurihara H, Orita M, Shibanuma T, Furuya T, Ikura M (1999) A novel calmodulin target recognition revealed by its NMR structure in complex with a peptide derived from Ca^{2+}-calmodulin dependent protein kinase kinase. Nat Struct Biol 6:819-824.

Ottinger EA, Botfield MC, Shoelson SE (1998) Tandem SH2 domains confer high specificity in tyrosine kinase signaling. J Biol Chem 273:729-735.

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 276 (Part A): 307–326.

Pakkenberg B, Pevig D, Marner L, Bundgaard MJ, Gundersen HJG, Nyengaard JR Regeur L (2003) Aging and the human neocortex. Exp Gerontology 38:95-99.

Pakkenberg B, Gundersen HJG (1997) Neocortical neuron number in humans: effect of sex and age. J Comp Neurology 384:312-320.

Pawson T, Gish GD and Nash P (2001) SH2 domains interaction modules and cellular wiring. Trends Cell Biol 11:504-511.

Pawson T and Nash P (2000) Protein-protein interactions define specificity in signal transduction. Genes Dev 14:1027-1047.

Pawson T and Nash P (2003) Assembly of cell regulatory systems through protein interaction domains. Science 300:445-452.

Pawson T and Scott JD (1997) Signaling through scaffold, anchoring ad adaptor proteins. Science 278:2075-2080.

Pearson G, Robinson F, Gibson TB, Xu B, Karandikar M, Berman K, Cobb MH (2001) Mitogen activated (MAP) protein kinase pathways: regulation and physiological function Endocrine Rev 22:153-183.

Penengo L, Rubin C, Yarden Y, Gaudino G (2003) c-Cbl is a critical modulator of the Ron tyrosine kinase receptor. Oncogene 22: 3669–3679.

Peschard P, Ishiyama N, Lin T, Lipkowitz S, Park M (2004) A conserved DpYR motif in the juxtamembrane domain of the Met receptor family forms an atypical c-Cbl/Cbl-b tyrosine kinase binding domain binding site required for suppression of oncogenic activation. J Biol Chem 279: 29565–29571.

Peschard P, Park M (2003) Escape from Cbl-mediated downregulation: a recurrent theme for oncogenic deregulation of receptor tyrosine kinases. Cancer Cell 3: 519–523.

Peschard P, Fournier TM, Lamorte L, Naujokas MA, Band H, Langdon WY, Park M (2001) Mutation of the c-Cbl TKB domain binding site on the Met receptor tyrosine kinase converts it into a transforming protein. Mol Cell 8: 995–1004.

van Petegem F, Chatelain FC, Minor Jr DL (2005) Insights into voltage-gated calcium channel regulation from the structure of the Ca(V)1.2 IQ domain-Ca(2+)/calmodulin complex Nat Struct Mol Biol 12: 1108-1115

Purcell AL and Carew TJ (2003) Tyrosine kinases, synaptic plasticity and memory: insights from vertebrates and invertebrates. Trends Neurosci 26:625-630.

Pyronnet, S. et al. Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. EMBO J. 18, 270±279 (1999).

Rahuel J, Gay B, Erdmann D, Strauss A, Gracia-Echeverria C, Furet P, Caravatti G, Fretz H, Schoepfer J, Grutter MG (1996) Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. N Struct Biol 3:586-589.

Reverter D, Lima CD (2005) Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. Nature 435: 687–692.

Rhoads AR and Friedberg F (1997) Sequence motifs for calmodulin recognition. FASEB J 11:331-340.

Rubin C, Litvak V, Medvedovsky H, Zwang Y, Lev S, Yarden Y (2003) Sprouty fine-tunes EGF signaling through interlinked positive and negative feedback loops. Curr Biol 13: 297–307.

Ryan PE, Davis GC, Nau MM, Lipkowitz S (2006) Regulating the regulator: negative regulation of Cbl ubiquitin ligase. Trends Biochem Sci 31:79-88.

Saghatelian A and Cravatt BF (2005) Assignment of protein function in the post genomic era. Nat Chem Biol 1:130-233.

Saimi Y, Kung C (2002) Calmodulin as an ion channel subunit. Annu Rev Physiol 64:289-311.

Saimi Y, Ling KY (1995) Paramecium Na+ channels activated by Ca2+-Calmodulin: calmodulin is the Ca2+ sensor in the calcium gating mechanism. J Membr Biol 144:257-265.

Sanjay et al. (2001) Cbl associates with Pyk2 and Src to regulate Src kinase activity, αvβ3 integrin mediated signaling, cell adhesion, and osteoclast motility. J Cell Biol 152:181-195.

Schlessinger J and Lemmon M (2003) SH2 and PTB domains in tyrosine kinase signaling. Sci STKE RE12.

Schrader LA, Birnbaum SG, Nadin BM, Ren Y, Bui D, Anderson AE, Sweatt JD (2006) ERK/MAPK regulates the Kv4.2 potassium channel by direct phosphorylation of the pore-forming subunit. Am J Physiol Cell Physiol 290: C852–C861.

Schmidt MH, Dikic I (2005) The Cbl interactome and its functions. Nat Rev Mol Cell Biol 6: 907–919.

Shiue L, Zoller MJ, Brugge JS. 1995. Syk is activated by phosphotyrosine-containing peptides representing the tyrosine-based activation motifs of the high affinity receptor for IgE.. J. Biol. Chem. 270:10498–502.

Shtiegman K and Yarden Y (2003) The role of ubiquitylation in signaling by growth factors: implications to cancer. Semin Cancer Biol 13:29-40.

Song J, Zhang Z, Hu W, Chen Y (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. J Biol Chem 280: 40122–40129.

Songyang Z, Shoelson SE, Chadhuri M, Gish G, Pawson T, King F, Roberts T, Ratnofsky S, Schaffhausen B, Cantley LC (1993) Identification of phosphotyrosine peptide motifs that bind to SH2 domains. Cell 72:767-778.

Sweatt JD (2001) The neuronal MAP kinase cascade: a biochemical signal integration system subserving synaptic plasticity and memory. J Neurochem 76:1-10.

Tamagnone L, Artigiani S, Chen H, He Z, Ming GI, Song H, Chedotal A, Winberg ML, Goodman CS, Poo M, Tessier-Lavigne M, Comoglio PM (1999) Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates. Cell 99: 71–80.

Tan HL, Kupershmidt S, Zhang R, Stepanovic S, Roden DM, Wilde AA, Anderson ME, Basler JR (2002) A calcium sensor in the sodium channel modulates cardiac excitability. Nature 415:442-447.

Thien CB and Langdon WY (2001) Cbl: many adaptations to regulate protein tyrosine kinases. Nat Rev Mol Cell Biol 2:294-307.

Thien CB, Langdon WY (2005) c-Cbl and Cbl-b ubiquitin ligases: substrate diversity and the negative regulation of signalling responses. Biochem J 391 (Part 2): 153–166 .

Thomas SM and Brugge JS (1997) Cellular functions regulated by Src family kinases. Annu Rev Cell Dev Biol 13:513-609.

Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. J Appl Cryst 30: 1022–1025.

Veldkamp MW, Viswanathan PC, Bezzina C, Baartscheer A, Wilde A, Balser JR (2000) Two Distinct Congenital Arrhythmias Evoked by a Multidysfunctional Na^+ Channel. Circ Res 86:E91-E97.

von Kriegsheim A, Pitt A, Grindlay GJ, Kolch W, Dhillon AS (2006) Regulation of the Raf-MEK-ERK pathway by protein phosphatase 5. Nat Cell Biol 8:1011-1016.

Wang DW, Makita N, Kitabatake A, Balser JR, George A Jr (2000) Enhanced Na^+ Channel Intermediate Inactivation in Brugada Syndrome. Circ Res 87:E37-E43.

Wang W, Malcolm BA (1999) Two-Stage PCR Protocol Allowing Introduction of Multiple Mutations, Deletions and Insertions Using QuikChange™ Site-Directed Mutagenesis. Biotechniques 26:680-682.

Warne PH, Viciana PR, Downward J (1993) Direct interaction of Ras and the amino terminal of Raf-1 in vitro. Nature 364:353-355.

Wasserman, JD and Matthew F (1998) An autoregulatory cascade of EGF receptor signaling patterns the Drosophila egg. Cell 95:355-364 (1998).

Waterman H, Katz M, Rubin C, Shtiegman K Lavi S, Elson A, Jovin T, Yarden Y (2002) A mutant EGF receptor defective in ubiquitylation and endocytosis unveils a role for Grb2 in negative signaling. EMBO J 21:303-313.

Weiss LA, Escayg A Kearney JA, Trudeu M, MacDonald BT, Mori M, Reichert J, Buxbaum JD, Meisler MH (2003) Sodium channels SCN1A, SCN2A and SCN3A in familial autism. Mol Psychiatry 8:186-194.

Weissman AM (2001) Themes and variations on ubiquitination. Nat Rev Mol Cell Biol 2:169-178.

Winzen, R. et al. The p38 MAP kinase pathway signals for cytokine-induced mRNA stabilization via MAP kinase-activated protein kinase 2 and an AU-rich region-targeted mechanism. EMBO J. 19,6742±6753 (1999).

Wong ES, Lim J, Low BC, Chen Q, Guy GR (2001) Evidence for direct interaction between Sprouty and Cbl. J Biol Chem 276: 5866–5875.

Wong ES, Fong CW, Lim J, Yusoff P, Low BC, Langdon WY, Guy GR (2002) Sprouty2 attenuates epidermal growth factor receptor ubiquitylation and endocytosis, and consequently enhances Ras/ERK signalling. EMBO J 21: 4796–4808.

Xia Z and Storm DR (2005) The role of calmodulin as a signal integrator for synaptic plasticity. Nature Rev 6:267-276.

Xu J, Kang N, Jiang L, Nedergaard M, Kang J (2005) Activity-Dependent Long-Term Potentiation of Intrinsic Excitability in Hippocampal CA1 Pyramidal Neurons. J Neurosci 25:1750-1760.

Yaffe M (2002) How do 14-3-3 proteins work? Gatekeeper phosphorylation and the molecular anvil hypothesis. FEBS Letters 513:53-57.

Yokouchi M, Kondo T, Sanjay A, Houghton A, Yoshimura A, Komiya S, Zhang H Baron R (2001) Src catalysed phosphorylation of c-Cbl leads to the interdependent ubiquitination of both proteins. J Biol Chem 276:35185-35193.

Yusoff P, Lao DH, Ong SH, Wong ES, Lim J, Lo TL, Leong HF, Fong CW, Guy GR (2002) Sprouty2 inhibits the Ras/MAP kinase pathway by inhibiting the activation of Raf. J Biol Chem 277: 3195–3201.