Structural analysis of adaptor protein CrkL and the role of CypA in *Bcr-Abl* mediated signaling

By Tamjeed Saleh

A dissertation submitted to the

Graduate School –New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate program in Molecular Biology and Biochemistry

Written under the direction of

Professor Charalampos Kalodimos

And Approved by

_____

_____

_____

New Brunswick, New jersey

October 2012

ABSTRACT OF THE DISSERTATION

STRUCTURAL ANALYSIS OF ADAPTOR PROTEIN CRKL AND THE ROLE OF CYPA IN *BCR-ABL* MEDIATED SIGNALING

By TAMJEED SALEH

THESIS DIRECTOR

Professor Charalampos Kalodimos

CrkL is a key signaling protein that mediates the leukemogenic activity of Bcr-Abl and is thought to adopt a structure that is similar to that of its CrkII homolog. The two proteins share high sequence identity and indistinguishable ligand binding preferences, yet they have distinct physiological roles. Here we show that the structures of CrkL and phosphorylated CrkL are markedly different than the corresponding structures of CrkII. As a result, the binding activities of the Src homology 2 and Src homology 3 domains in the two proteins are regulated in a distinct manner and to a different extent. The different structural architecture of CrkL and CrkII may account for their distinct functional roles. The data show that CrkL forms a constitutive complex with Abl,explaining the strong preference of Bcr-Abl for CrkL. The results also highlight how the structural organization of the modular domains in adaptor proteins can control signaling outcome.

Cyclophilin A (CypA) is a polyprolyl isomerase that is ubiquitously expressed in all human cells. It catalyzes the cis-trans isomerization of X-P motifs. Recent studies have shown that CypA is overexpressed in many human cancers. The role of CypA in this area is not well understood. We have identified a novel binding site for CypA on CrkII, an adaptor protein. Although several binding partners of CypA are known, the underlying mechanism of the CypA action and the physiological implications of the interactions have remained in most cases unknown. It is unclear whether CypA acts as an enzyme or a binding partner in mediating the biological processes. Here we show that CypA binds specifically to CrkII and modulates its level of phosphorylation by the Abl kinase. We show that the conserved proline residue in CrkII next to the tyrosine-phosphorylation site undergoes cis–trans isomerization. CypA is recruited to the site and prevents the tyrosine from becoming phosphorylated by Abl. The interaction between CypA and CrkII occurs specifically both in vitro and in vivo. This is a novel role for CypA which appears to act as a selective switch to modulate the level of phosphorylation of a signaling protein.

Acknowledgements

Thesis Director:
Prof. Charalampos G. Kalodimos

Committee Members:
Prof. KiBum Lee
Prof. Jean Baum
Prof. Raymond Birge

Lab members:
Dr. Wojciech Jankowski, Dr. Li Chen Dr. Nataliya Popovych, Dr. Dimitra Keramisanou, Dr. Ioannis Gelis Dr. Krishna Mohan Poluri, , Dr. Akash Bhattacharya, Dr. Xuinjun Ai Nandish Khanra, Soma Mandal, Pragati Sharma, Dr. Lata Chaun, Ankita Basant, Dr. Anusarka Bhaumick, Dr. Tomohide Saio, Dr. Chih-Hong Chen, Dr. Xiao Guan, Dr. Soumyasri Das Gupta, Dr. Aishwarya Ravindran, Dr. Mandar Naik, Dr.Ming-Tao Pai. Dr.Hsiu-Ru Tseng

NMR managers (New Brunswick):
Dr. Seho Kim and Dr. Nagarajan Murali

Department of Molecular Biology and Biochemistry, faculty and staff
Department of Chemistry and chemical biology faculty and staff

FAMILY and Friends.

Contents

Abstract iii

List of Figures vii

1 Cell signaling via adaptor proteins 1

1.1 Crk family of proteins 1

1.1.1 SH2 domains 3

1.1.2 SH3 domains 6

1.2 Protein complex with Crk 8

1.2.1 Paxillin 8

1.2.2 C3G 10

1.2.3 DOCK180 10

1.2.4 Crk associated kinase substrate(p130CAS) 11

1.2.5 Abl kinase 12

1.3 Structural basis of Crk function 16

1.4 The Role of SH3C domain in Crk signaling. 17

1.5 Crk proteins and disease. 20

2 Nuclear Magnetic Resonance (NMR) spectroscopy 22

2.1.1 Assignment of chemical shifts of protein spectra 22

2.1.2 Assignment of side-chains 25

2.1.3 Nuclear Overhauser Effect Spectroscopy (NOESY) 26

2.1.4 Protein deuteration 28

2.1.5 Paramagnetic Relaxation Enhancement (PRE) 29

2.1.6 TALOS 31

2.1.7 Residual dipolar coupling (RDC) 32

2.1.8 Protein structure calculation 33

2.1.9 Protein dynamics 34

2.1.10 Ligand binding studies by NMR. 35

2.2 Fluorescence Resonance Energy Transfer (FRET) 36

2.3 Isothermal Titration Calorimetry (ITC) 39

3 Research Outline 41

4 Structural determination of CrkL 42

4.1 Introduction 42

4.2 Results 43
4.2.1 Analysis of CrkL using NMR 43
4.2.2 Structure of CrkL and pCrkL 47
4.2.3 Dynamic properties of CrkL 50
4.2.4 $SH3^N$ domain in CrkL is not inhibited. 52
4.2.5 CrkL inhibition via Y207 phosphorylation 54
4.2.6 Unlike CrkII, the $SH3^N$ domain of pCrkL is not inhibited 56
4.2.7 Discussion 59
5 CypA acts as a switch for Abl phosphorylation 66
5.1.1 Introduction 66
5.1.2 PPIase family 66
5.1.3 Cyclophilins (Cyps) family 67
5.1.4 CypA 67
5.1.5 Cyclophilin and disease 68
5.2 Results 71
5.2.1 Proline induced conformational heterogeneity in CrkII 71
5.2.2 CypA binds to CrkII around G219-P220 motif 77
5.2.3 Structural basis for the interaction of CypA with CrkII 81
5.2.4 *In-vivo* localization of CypA and CrkII 84
5.3 Discussion 84
6 Materials and Methods 88
6.1 Protein Preperation of Crk fragments and Abl kinase. 88
6.2 Purification of CypA 89
6.3 NMR Spectroscopy 90
6.3.1 2D ^{1}H-^{15}N Heteronuclear (ZZ NMR experiments) 91
6.3.2 Residual dipolar Coupling. 92
6.3.3 Relaxation measurements and analysis 93
6.4 Isothermal titration Calorimetry 94
6.5 Structure Calculation 95
6.6 PRE measurements 97
6.7 Mass Spectrometry 98
6.8 Kinase assay and Western Blots 98
6.9 FRET analysis 99
7 Bibliography 102
8 Curriculum Vitae 113

List of figures.

Fig 1:1 Structure of the Crk family of proteins. 2
Fig 1:2 Figure depicts the role of SH2 domains in signaling pathways 3
Fig 1:3 Family of SH2 containing proteins. Enzymes, Adaptor proteins, Scaffolding proteins, Transcription factors and Signal Regulars contain one or more SH2 domains. [8] 4
Fig 1:4: Src-homology -2 (SH2) domain 5
Fig 1:5 Src homology -3 (SH3) domain 7
Fig 1:6 Binding partners of CrkII and CrkL. The $SH3^{C}$ domain has no binding partner. 8
Fig 1:7 Figure shows the domain organization of scaffolding protein Paxillin. Important phosphorylation sites are shown with arrows 9
Fig 1:8 Figure shows domain organization of C3G. Y504 is an important phosphorylation site. CBR –Crk Binding Region. 10
Fig 1:9.p130CAS structure and binding partners. (A)Domain organization of p130CAS.The N-terminal region binds FAK and Pyk2 kinases. (B-C) Signal transduction by p130CAS upon mechanical stress 11
Fig 1:10 Domain organization of Abl1b and *Bcr*-Abl. 13
Fig 1:11 Structural features of c-Abl and c-Src. c-Abl with kinase inhibitor PD166326 and c-Src with AMP-PMP. [66] 14
Fig 1:12 Solution structure of CrkII (left) and pCrkII(right). SH2 (red), $SH3^{N}$ (green),$SH3^{C}$ (blue). Circles depict ligand binding site. (dotted circle presumed binding site for $SH3^{C}$) 16
Fig 1:13 Prolyl *cis-trans* isomerization mediates the inhibition of the $SH3^{N}$ domain. 18
Fig 1:14 Regulation of Crk Activity. Crk exists in an equilibrium (90% inhibited:10% uninhibited) Ligand binds to the uninhibited form and drives the equilibrium towards the uninhibited form. 19
Fig 1:15 Effect of Crk knockdown by siRNA in human ovarian cancer cell line MCAS 21
Fig 2:1 ^{1}J and ^{2}J coupling constants for magnetization transfer in ^{13}C,^{15}N-labeled proteins. 24
Fig 2:2 Magnetization transfer in an HNCA(left) and an HN(CO)CA(right) experiment. 25
Fig 2:3 Magnetization transfer in a C(CO)NH experiment. 26
Fig 2:4 Magnetization transfer in a ^{15}N-NOESY-HSQC experiment. 27
Fig 2:5 Chemical structure of spin label reagent MTSL attached to a protein *via* disulphide bond. 30
Fig 2:6 Location of the dihedral angles in the polypeptide. 31
Fig 2:7 Time scales for protein dynamics and NMR techniques. 34
Fig 2:8 Intermolecular FRET: Protein X is labeled with yellow fluorescent protein (YFP) and protein Y is labeled with cyan fluorescent protein (YFP). 37
Fig 2:9 Excitation and emission profiles of donor CFP and acceptor YFP 38
Fig 2:10 Layout of an ITC machine (left). Sample titration plot from an ITC experiment (right). . 39
Fig 4:1 1H-15N HSQC spectra of CrkL domains overlayed on full-length CrkL. The cross-peaks of 44
Fig 4:2 Multi-angle laser light scattering (MALLS) shows that CrkL is monomeric in solution 45
Fig 4:3 Chemical shift differences (Δδ) plotted as a function of residue number between isolated domains and full length CrkL. 46

Fig 4:4 (a) Overlay of the structures of CrkL SH2 (this work) and CrkII SH2 domains. The DE loop, which was shown to bind to the SH3 domain of Abl, is present only in CrkII. **(b)** Overlay of the structures of CrkL $SH3^N$ (this work) and CrkII $SH3^N$ domains. Overall the structures are very similar other than some structural heterogeneity in the loops. 47
Fig 4:5 ^1H-^{15}N HSQC spectra of full length CrkL(1-303) and pCrkL..... 48
Fig 4:6 (a) Structure of CrkL. The SH2, SH3N and SH3C domains are colored green, magenta and blue, respectively. The linker regions are colored gray. The SH3C domain does not interact with the other domains. **(b)** Close-up view of the SH2-SH3N interface in CrkL. Only polar or charged residues mediate the interaction between the two domains**.** 49
Fig 4:7 Structure of pCrkL. **(a)** Structure of pCrkL. pTyr207 is shown as orange sticks. **(b)** Close-up view of the pTyr-binding site. The SH2-$SH3^N$ interface undergoes slight adjustment to accommodate the binding of the linker to SH2. 50
Fig 4:8 Plot of the R2/R1 ratio. 15N relaxation rates of the CrkL backbone as a function of residue number. The R2/R1 ratio provides information about the tumbling of the molecule, with higher values indicating slower tumbling 50
Fig 4:9 (a) Correlation times (τc) for the tumbling of CrkL. The SH2-SH3N module tumbles as a rigid unit, whereas the SH3C domain tumbles much faster and independently of the other domains. **(b)** Residues undergoing substantial μs–ms motions, as denoted by enhanced contribution to R_2 (*R*ex) values, are mapped on the structure of CrkL in red. Almost all residues located at the interface between the SH2 and SH3N domains show relatively high *R*ex values, indicating that the binding interface is dynamic. 51
Fig 4:10 Spectra showing interaction of PPII-peptide ligand is limited to the $SH3^N$ domain only. 52
Fig 4:11 ITC traces and binding isotherms of titrations performed on isolated $SH3^N$ (a), full-length CrkL(b) and pCrkL(c). 53
Fig 4:12 K_d values of PPII-peptide complexes with CrkL and CrkII fragments. Standard error was determined from three independent experiments. 53
Fig 4:13 Plot of the R2/R1 ratio of pCrkL as a function of residue number. The scheme at right shows that the SH2-SH3N module in pCrkL tumbles as a unit, as in CrkL, whereas the SH3C domain tumbles much faster and independently of the other domains. 54
Fig 4:14 Effect of Tyr207 phosphorylation on CrkL folding and its association with Abl kinase. **(a)** ^1H-^{15}N HSQC NMR spectra of the linker region of CrkL containing the phosphorylated Tyr207 (pTyr-linker) in the presence of CrkL(orange) and after the addition of catalytic amounts of AblKD and ATP-Mg2+ (blue). The pTyr-linker is 15N-labeled, whereas CrkL and AblKD are unlabeled. Asterisk denotes isotopic labeling. (b) Analysis of the NMR experiments in a shows that the pTyr-linker binds the SH2 domain of CrkL. Phosphorylation of Tyr207 in CrkLinduces the intramolecular association of pTyr207 and SH2. As a result, the pTyr-linker is displaced. (c) Pulldown of CrkLand pCrkLwith paxillin, an SH2-binding physiological partner of CrkL 56
Fig 4:15 Pull-down experiment of CrkL and pCrkL with DOCK1, an $SH3^N$ binding partner of CrkL. Membrane was blotted for DOCK1 and pTyr207. 57
Fig 4:16 (a) 1H-15N HSQC NMR spectra of free CrkL(blue), in complex with AblPxxP (orange) and after adding ATP+Mg2+ (magenta). AblPxxP is a construct that encompasses the kinase domain and the first PxxP motif that binds CrkL. **(b)** Pull down of CrkL and pCrkL with full-length Abl (form 1b) Membrane blotted for Abl and pTyr207. 58
Fig 4:17 Sequence and domain organization of CrkL and CrkII. (a) Sequence alignment of human CrkII and human CrkL. (b) Domain organization of CrkII and CrkL. Tyr207 in CrkL and Tyr221 in CrkII are phosphorylated by Abl kinase. 59

Fig 4:18 ITC traces and binding isotherms of titrations performed on isolated SH2 domain and full-length CrkL with phosphopeptide CrkL(pTyr207). pTyr207 binds isolated SH2 with a K_d ~7μM and full-length SH2 with a K_d~23 μM. 61
Fig 4:19 Binding of pTyr peptide and PPII-ligand peptide to CrkL and CrkII. (a,b) The Structure of SH2-$SH3^N$ domains in CrkL(a) and CrkII(b)The pTyr peptide binding site in CrkL is partially masked but is completely accessible in CrkII. The PPII binding site in CrkL is exposed whereas in CrkL it is masked. (c) Summary of the Dissociation constants (K_d) of pTyr-peptide and PPII ligand with CrkL and CrkII. 61
Fig 4:20 (a) Binding isotherm of isolated SH2 domain of CrkII and CrkL with pTyr peptide from FGFR peptide. Isolated SH2 domain of CrkII and CrkL have similar binding affinity. (b) The peptide binding mechanism is identical for both SH2 domains. 62
Fig 4:21 Abl Kinase docks on the $SH3^N$ domain of CrkL and phosphorylates Tyr207. Intramolecular binding of SH2 domain to pTyr207 does not disrupt Abl-CrkL complex. 63
Fig 4:22 CrkL versus CrkII signaling. Integrin activation elicits p130CAS phosphorylation by tyrosine kinases (TK), and, as a result, CrkLand CrkII are recruited. (ii) GEFs (for example, DOCK1 and C3G) associate with CrkLand CrkII via their SH3N domain, giving rise to efficient localized activation (iii) of small GTPases (for example, Rac, Rap and RRas) at the membrane. (iv) Abl-induced phosphorylation of CrkLand CrkII forces their dissociation from p130CAS and thus results in signaling suppression. The distinct structural organization of CrkLand CrkII modulates the interactions with their physiological partners to a different extent. The blue and brown shaded regions in SH2 and SH3N denote the pTyr- and PPII-binding sites, respectively 64
Fig 5:1 (a) Surface representation of CypA with ligand. The most effected residues are shown in green. **(b)**A zoomed view of the active-site of CypA bound to ligand. Arg55 forms the essential contacts with the isomeric Proline. 68
Fig 5:2 Expression profile comparing CypA expression levels in tumor tissue and the normal counterpart. Profile created from SAGE map. http://www.ncbi.nlm.nih. {Obchoei, 2009 #9196} 69
Fig 5:3 Sequence alignment of CrkII(216-231) shows absolute conservation of amino acids around Tyr221. 71
Fig 5:4. **a**. ^{1}H-15N HSQC of CrkII (1-304) **b.** ^{1}H-15N HSQC of hl-$SH3^c$-P238A. 72
Fig 5:5 (a) Strips from CCO-NH spectra showing the ^{13}C chemical shift of Pro220 carbon skeleton. The chemical shift difference between the $^{13}C_\beta$ and $^{13}C_\gamma$ nuclei for the two conformations of residue Pro220 are 9.2ppm and 4.6ppm which, on the basis statistical analysis of ^{13}C chemical shifts[191] of proline residues in proteins corroborate that the two forms are in the *cis* and *trans* conformations.(b) Overlay of ^{1}H-^{15}N HSQC of hl-$SH3^C$ (red) and hl-$SH3^C$ -Pro220 (green) 74
Fig 5:6 ^{1}H-15N HSQC of l-$SH3^C$ P238A CrkII. Zoom (right) 74
Fig 5:7 ^{1}H-^{1}H NOESY experiment of the synthetic peptide displays *cis-trans* isomerization about Pro220. The two sets of peaks, one for each conformation is easily identified. 76
Fig 5:8 (A) A titration experiment where unlabeled CypA was added to ^{15}N-labeled CrkII. Binding of CypA leads to peak broadening for the residues near the binding site. **(B)** ZZ-Exchange spectroscopy. Catalytic amounts of CypA lead to the appearance of exchange peaks shown in dotted lines. Exchange peaks appear when the rate is quite fast. **(C)** CypA interaction with CrkII. Peak broadening observed for residues that come in contact with CypA. (D) Chemical shift mapping of residues perturbed upon CypA binding. Scale Red –most peak broadening, green – moderate, orange-slight, white-unaffected. 77
Fig 5:9 (A) ^{1}H-^{15}N HSQC of ^{15}N-labeled CypA and unlabeled hCrkII at 1:1 molar ratio. Chemical shift mapping of the affected region reveals that CypA uses the active site to interact with CrkII.

(B) hCrkII-P220A mutation does not interact with CypA confirming that the Gly219-Pro220 site is indeed the binding region for CypA. 79
Fig 5:10 Spectra of CrkII-CypA Interaction 79
Fig 5:11 (A) ^{1}H-^{15}N HSQC spectra of ^{15}N –labeled CypA acquired with varying concentration of CrkII. The titration is color coded. Due to the relatively weak binding and the interaction displaying properties of fast-exchange in the NMR time scale, the dissociation constant could be extracted by fitting a plot of chemical shift change as a function of concentration of CrkII. 80
Fig 5:12 (A-B) CypA uses the active site to interact with CrkII. (**C-D)**The peptide binds deep inside the CypA active site. 82
Fig 5:13 (A) Western blot for Kinase assay performed with Abl kinase and CrkII as substrate in the presence and absence of CypA. The membrane was probed with anti-CrkII pY221 antibody to monitor phosphorylation. Quantification performed in ImageJ. In the presence of CypA there is a ~12 fold inhibition of Tyr221 phosphorylation. **(B)** Western Blot probing for pY221 phosphorylation in EGF-stimulated MDA-MB-468 cells in the presence and absence of CsA. 83
Fig 5:14(A) FRET analysis performed on HELA cells transfected with CrkII-YFP and CypA-CFP. The cell shows a maximum FRET efficiency of 70% near the membrane. (B) Treatment of cell with CsA leads to a decrease in FRET efficiency ~15%. 84
Fig 5:15 (A) ^{1}H-^{15}N HSQC of ^{15}N-labeled CrkL. (B) ^{1}H-^{15}N HSQC of ^{15}N-labeled CrkL and 2 fold molar excess of unlabeled CypA shows no interaction between these two proteins. 86
Fig 5:16 Western blot for kinase assay performed with Abl Kinase and CrkL as a substrate. The membrane was probed with anti-CrkL-Y207 antibody. Quantification performed in ImageJ. 86
Fig 6:1 Gel filtration profile of full-length CrkL and pCrkL. 89
Fig 6:2 Schematic representation of CypA catalyzed *cis-trans* isomerization. 91
Fig 6:3 A sample of the splitting observed in RDC measurements. 93
Fig 6:4 Summary table of statistics for solution structure of CrkL and pCrkL. 96
Fig 6:5 (a) Spectra of PRE effects of S20C mutation. (b) Spectra of PRE effects of 190C. 98
Fig 6:6 Equation to calculate FRET efficiency. 100

1 Cell signaling via adaptor proteins

The ability to integrate the multitude of changes occurring in the environment and organize an appropriate response is crucial for the success of an organism. Adaptor proteins are known to play a significant role in cell signaling. Devoid of enzymatic activity, these molecules can interact with multiple proteins and help localize and enhance signal transduction in the cell. Usually stimulation of cell surface receptors initiate cellular signaling that are then propagated by post-translation modifications (e.g phosphorylation) or conformational changes (G-protein receptors) leading to recruitment of other proteins. Adaptor proteins contain a variety of protein-binding modules that can connect various proteins and create a signaling complex. They form a bridge between activated receptors and specific downstream signaling proteins. Some of the adaptor proteins are crucial mediators of important cellular functions such as cell survival, apoptosis, proliferation and migration. High expression levels of these proteins often lead to tumorigenesis despite the fact that they possess no catalytic activity. Of all the adaptor proteins present in our proteome, the Crk family of adaptor proteins is the most widely studied.

1.1 Crk family of proteins

Crk proteins - a family of adaptor proteins is involved in transcription, cell growth, motility, proliferation transformation and apoptosis [1]. There are several members of the Crk family: CrkI, CrkII, CrkIII and its paralog CrkL **Fig 1:1**

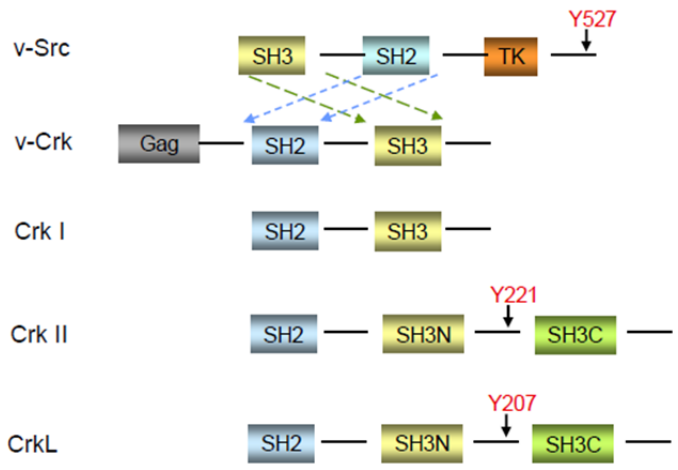


Fig 1:1 Structure of the Crk family of proteins.

Crk was originally identified as an avian sarcoma virus CT10 encoding protein, v-crk which is oncogenic and possess transforming ability.[2, 3] As Crk can activate cellular tyrosine kinases it was named *crk* (CT10 regulator of kinase). The mammalian homolog CrkII and CrkL consist of a Src homology domain 2 (SH2)domain and two Src homology domain 3 (SH3) domains. The alternative splice form CrkI contains one SH2 and only one SH3 domain and is similar to *v*-crk. [4] SH2 domains are structurally conserved and contain ~ 100 amino acids and mediate interaction with motifs that contain phosphorylated tyrosine resides allowing SH2 domain containing proteins to bind to proteins that are tyrosine phosphorylated(pTyr). SH2 domains will bind a –pYXXP- where X is any amino acid. [5]. SH3 domains are also structurally conserved and contain ~60 amino acids. SH3 domains primarily interact with proline rich sequences specifically polyprolyl type II (PPII motif) sequences. Crk SH3N domains bind to P-X-X-P-X-K motifs where X is any amino acid. [6] Even though *v-Crk* was the first identified member of the family with oncogenic potential, the cellular counterparts, CrkI, CrkII and CrkL have been studied in more detail. The modular nature of the Crk family implies that they share similar structural features. However, the biological data have been unable to elucidate a definite role for each of the

members. The Crk adaptor proteins possess no catalytic activity and wield their influence by modulating signals from various pathways including those involved in cell growth, migration, proliferation, apoptosis and gene transcription. [1, 7] More than 40 cellular proteins have been identified that bind to the SH2 and SH3N domains of Crk in the last two decades of research.

As the Crk proteins exert their influence in the signal transduction process mainly by the SH2 and SH3 domains, a brief description to these domains is provided below.

1.1.1 SH2 domains

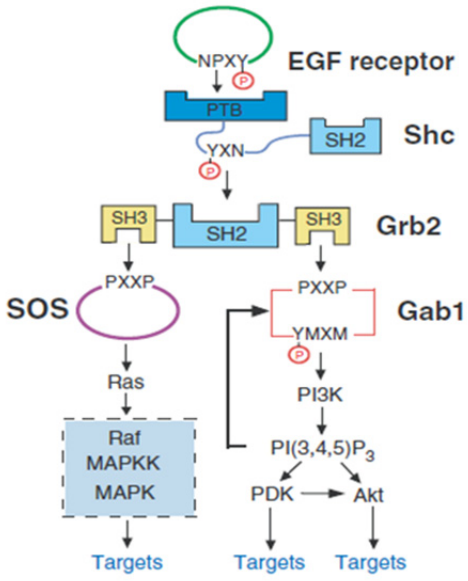


Fig 1:2 Figure depicts the role of SH2 domains in signaling pathways

SH2 domains contain ~ 100 amino acids and are known to bind phosphorylated tyrosine motifs. Unlike other binding motifs, the SH2 domain functions in the phosphotyrosine kinase

pathways in cell signaling. SH2 domains recognize the phosphotyrosine residue (pTyr) in protein sequences. This post translational modification in proteins is introduced by tyrosine kinases and is a critical mechanism used by cells to impart information.

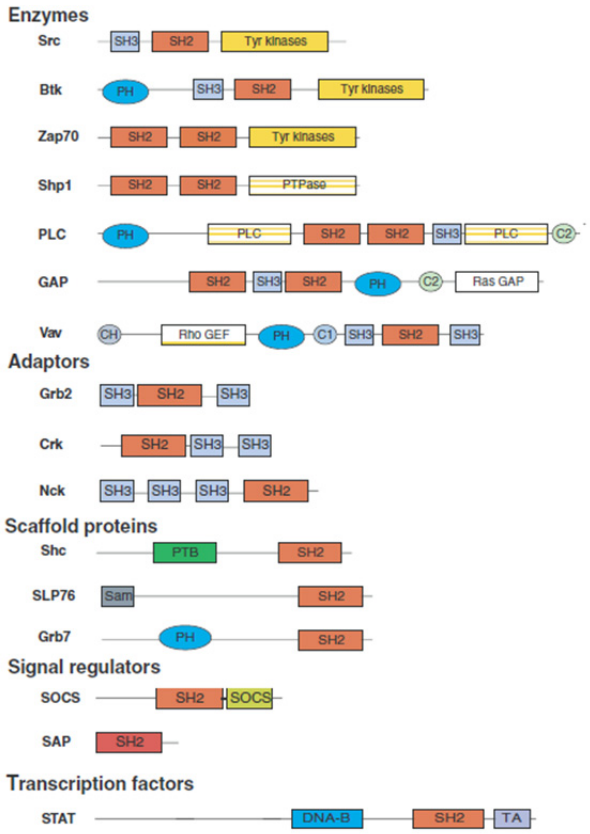


Fig 1:3 Family of SH2 containing proteins. Enzymes, Adaptor proteins, Scaffolding proteins, Transcription factors and Signal Regulars contain one or more SH2 domains. [8]

In cells, SH2 domains are frequently found accompanying other domains; kinases, SH3 domains or other SH2 domains with varying specificity. SH2 domains assist in localization of proteins and enable signaling complexes to form to expedite information transfer. The Grb2 protein contains two SH3 domains and one SH2 domain which allows for the recruitment of SOS

and Gab1 upon activation of EGFR [9]. The N-terminal SH3 domain of Grb1 interacts with the proline rich sequence of SOS [10]. Simultaneously, the C-terminal SH3 domain of Grb1 can bind to Gab1 and recruit it to the site of the activated EGRF and SOS. Phosphorylation of Gab1 creates more pTyr binding sites for the SH2 domain of PI3K. In this case, Grb2 not only acts to transmit information from the activated kinase to SOS1 and eventually the MAP kinase pathway but also recruites Gab1 to initiate the PI3K/Akt pathway. The interest in this domain has grown in recent times given that a majority of the processes in cells that involve cancer include an aberrant tyrosine kinase**Fig 1:2**.

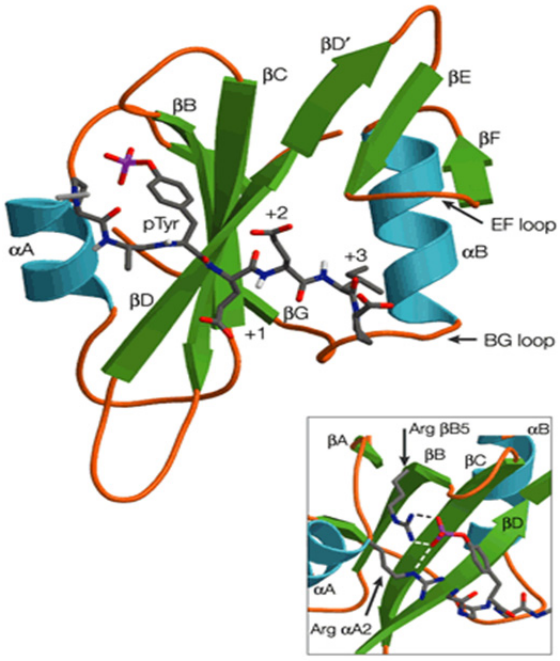


Fig 1:4: Src-homology -2 (SH2) domain

A typical SH2 domain consists of 100 amino acids that form 2 alpha helices and 3-4 antiparallel beta strands **Fig 1:4**. The pTyr moiety binds to the positively charged pocked on the surface of the SH2 domain. The specificity of this interaction is mediated by a conserved Arginine residue and a few hydrogen bonds. About 50% of the binding energy in this interaction is provided by a

few hydrogen bonds and an interaction with an invariable Arginine residue.[11] The surface of the SH2 domain also contains sites that interact with the residues following the pTyr. The ligand binds the SH2 domain in an extended conformation across the β-sheet. Peptide array analysis revealed that the 3-5 residues following the pTyr residue determine the specificity of the ligand for the SH2 domain. [12, 13]. The Src SH2 domain binds preferentially to peptides with the sequence pY-E-E-I while the PLC-γ1 prefers hydrophobic residues from pTyr +1 to pTyr+5. [14] The SH2 domain of Grb2 prefers an Asn at the pTyr+2 position.[15] Typical binding affinity for SH2 domains and peptide ligands is between 0.1-1.0μM. [16]. The SH2 domain of Crk primarily binds paxillin[17] , p130Cas,[18] Gab1 [19] and c-Cbl. [20]. The SH2 domains of both CrkII and CrkL have very similar binding specificity and prefer the motif –pY-X-X-P-. Although SH2 domains are predominantly thought to bind one ligand, (based on structural and biochemical analysis) several groups and recently reported the structures of SH2 domains interacting with two pTyr motifs simultaneously. [21, 22]. This finding may provide the basis of further research and add another layer of complexity to SH2 domain mediated signaling.

1.1.2 SH3 domains

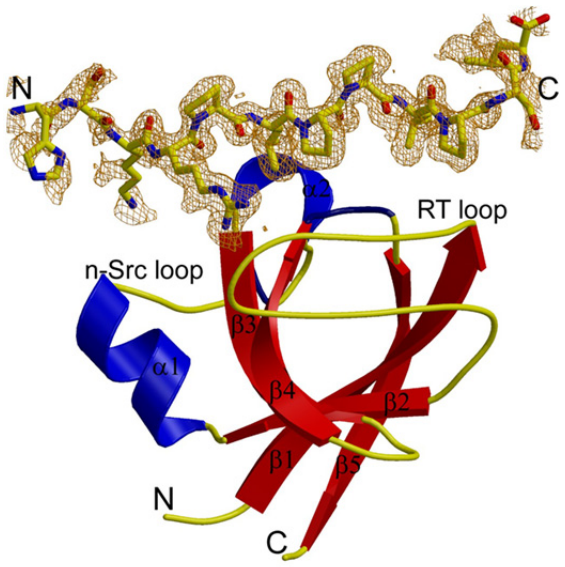


Fig 1:5 Src homology -3 (SH3) domain

The SH3 domain consists of ~ 60 amino acids and is smaller in size than the SH2 domain. With more than 300 genes containing an SH3 domain, they play an important role in protein-protein interactions. The general motif for binding includes a left handed poly-prolyl type II helix (PPII helix) with a sequence P-X-X-P.[6]. The basic fold of the SH3 domain includes five anti-parallel β-strands forming two perpendicular beta sheets.(**Fig 1:5**). The ligand binding site consists of a flat hydrophobic surface with three shallow pockets that are lined by aromatic residues.[23] Two of these pockets are occupied by two hydrophobic dipeptides and the third pocket determines specificity. The PPII helix, because of its architecture can have two orientations that are classified as Class 1 and Class II ligands. Class I ligands have the general consensus sequence of +XϕPXϕP and Class II ligands have ϕPxϕPx+ where x is any amino acid and + is a basic amino acid –commonly Arg. [24, 25]. SH3 domains display moderately lower affinity for their ligands when compared to SH2 domains (~1μM)

The first SH3 domain of Crk prefers the ligand P-X-L-P-X-K. The lysine residue is crucial as it forms polar contacts with charged residues on the SH3 and hydrophobic interactions with the

conserved tryptophan [26,27] The major interactors of Crk SH3N domain are Abl, Dock180 C3G and SOS.

Crk has been shown to bind to numerous cellular proteins and these interactions are mediated by the SH2 and SH3 domains **(Fig 1:6)**.[7] A brief description of the main interacting partners of the Crk proteins is provided below.

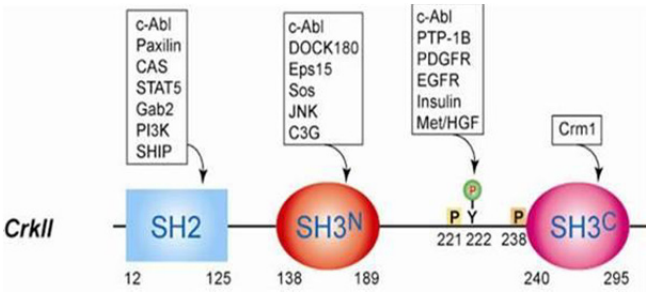


Fig 1:6 Binding partners of CrkII and CrkL. The $SH3^{C}$ domain has no binding partner.

1.2 Protein complex with Crk

1.2.1 Paxillin

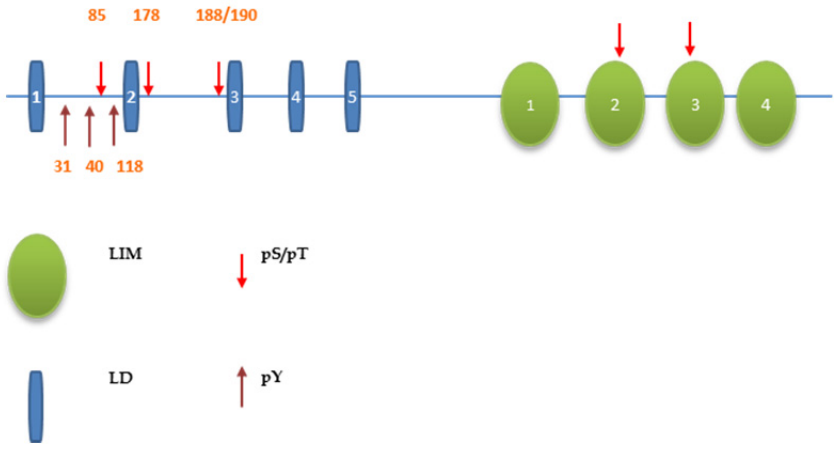


Fig 1:7 Figure shows the domain organization of scaffolding protein Paxillin. Important phosphorylation sites are shown with arrows.

Paxillin is a multi-domain scaffolding protein involved in mediating interactions of numerous proteins to control the changes in cell adhesion and cytoskeletal reorganization. Initially identified as a protein that becomes phosphorylated in chick embryonic fibroblasts transformed by *v-Src* , paxillin it has now been shown to mediate the interaction of Rho family of GTPases and play a crucial role in dynamic remodeling of the cytoskeleton. [28] Paxillin contains five leucine and aspartate rich LD motifs in the N-terminus of the protein. (**Fig 1:7**) Paxillin forms complexes with a number of kinases including p21 activated kinase (PAK), Src, JNK, ERK, CDK5 and c-Abl. Disruption of the interaction between paxillin and p95PKL (Paxillin kinase linker) prevents cell migration in wound healing assays.[29]. Phosphorylation of paxillin at Y31 and Y118 creates a docking site for the SH2 domain of CrkII. Mutation of these residues prevents the formation of paxillin-CrkII complex, a necessary step for collagen mediated cell migration. [30]. Expression of *v-Crk* results in increased level of tyrosine phosphorylation of paxillin presumably creating more docking sites for SH2 domain bearing proteins. Interestingly, paxillin also recruits PTP-PEST (a tyrosine phosphatase) *via* the LIM3 and LIM4 domains, PP2A and SHP-2. [31-34] Paxillin thus provides a unique platform where kinases and phosphatases can come together to create dynamic signaling modules.

1.2.2 C3G

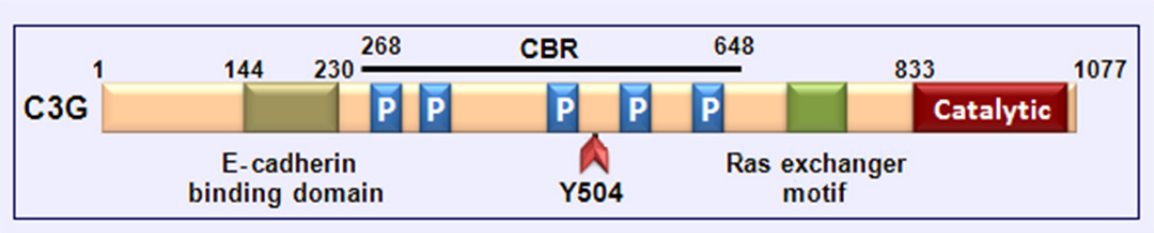


Fig 1:8 Figure shows domain organization of C3G. Y504 is an important phosphorylation site. CBR –Crk Binding Region.

C3G, a guanine nucleotide exchange factor (GEF) for Rap1 and R-Ras, was initially identified in a screen for binding partners of Crk $SH3^{N}$ domain. It is a ubiquitously expressed protein and higher levels of expression are observed in differentiated neuroblastoma cells and non-small cell lung cancers [35, 36]. The catalytic activity of C3G is required for the replacement of GDP by GTP in small GTPases like Rap-1, Rap-2 and R-Ras. [37]. Along with the catalytic region for the GEF activity, it contains proline-rich sequences (CBR) which bind to SH3 domain containing proteins like Crk and Grb2. The catalytic activity of C3G is regulated by Crk binding and the tyrosine phosphorylation at Y504 (pY504)[38] Although Src kinase has been known to phosphorylate this site, recent studies have shown that c-Abl and *Bcr-Abl* can perform this task. [39, 40]. pY504 recruites p130CAS and the C3G-CrkII complex along p130CAS localizes to the membrane and activate membrane bound Rap1. [41, 42].

1.2.3 DOCK180

DOCK180 (Downstream of Crk; DOCK1) is an unconventional guanine nucleotide exchange factor (GEF) for Rac1. It is similar to C3G in the terms of the catalytic activity but unique because it contains a DHR2/CZH2 domain instead of a Dbl homology domain. This domain is necessary for interacting specifically to nucleotide free Rho family GTPases. [43]. Although DOCK180 has poor

GEF activity, its catalytic activity is greatly enhanced when in complex with ELMO1. [44]. The DOCK180-ELMO complex activates Rac1 to mediate the optimal phagocytosis of apoptotic cells and cell migration. [45] DOCK180 interacts with CrkII via the $SH3^N$ domain and acts downstream of paxillin. [46-47] In fact, p130CAS-CrkII forms a complex with DOCK180 and promotes Rac1 activity. [48] Interestingly, the $SH3^C$ domain of CrkII plays a role in stabilizing the DOCK180-ELMO-Crk complex – one of the few examples of the $SH3^C$ playing a functional role. [49] A recent study identified a role of DOCK180 in gliobastoma, a malignant cancer of the brain. The signaling pathway includes: activation of Src kinase by EGFR which specifically phosphorylates DOCK180 at Tyr722. This leads to activation of Rac1 and gliobastoma cell migration. [50]

1.2.4 Crk associated kinase substrate(p130CAS)

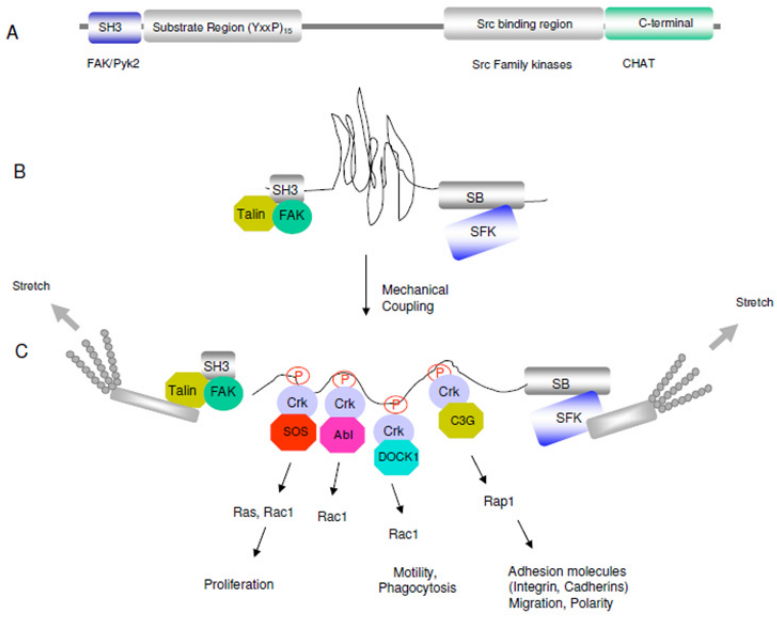


Fig 1:9.p130CAS structure and binding partners. (A)Domain organization of p130CAS.The N-terminal region binds FAK and Pyk2 kinases. (B-C) Signal transduction by p130CAS upon mechanical stress

p130CAS was identified as a partner of oncogene *v-crk* and is a substrate of *v-Src.* [18, 51] p130Cas is heavily phosphorylated in cells transformed by the oncoproteins and mutations that disrupt binding between v-crk/v-Src and P130CAS diminishes the transforming potency of the oncoproteins.[3, 52]. The 130kDa protein contains multiple –Y-X-X-P motifs that, upon phosphorylation, can act as a docking site for SH2 domain containing proteins. [53] p130CAS also contains binding site for Nsp proteins and an SH3 domain at the N-terminal that can interact with FAK and Pyk2. (**Fig 1:9**A) [54] In the cell p130CAS plays a major role in many signaling pathways as highlighted by the fact that a variety of stimulants induce tyrosine phosphorylation of p130CAS. [1] Of these, the most noteworthy is the study showing the phosphorylation of p130CAS is dependent on mechanical stretching (**Fig 1:9**) (B-C). Interestingly, this phosphorylation is performed by Src and Abl kinase and not by other members of the family. [55]The phosphorylations create docking sites for Crk which can then recruits various other proteins like, SOS, Abl and C3G.(**Fig 1:9**C) Although p130CAS has been shown to be phosphorylated at multiple sites by Src kinase [56] the inducible nature of p130CAS phosphorylation upon mechanical stretching supports earlier studies of its role in cell migration. [57] p130CAS has also been identified as being necessary for cell transformation[58] and the overexpression of p130CAS leads to chemoresistance in breast cancer cells.[59]

1.2.5 Abl kinase

Abl kinase is an important regulator of cell homoeostasis and its expression is conserved in all metazoans. The protein was first identified as a result of studies performed on the *Abelson murine leukemia virus*(A-MuLV)[60]. Mice infected with this virus developed thymic-independent lymphomas than other similar viruses like M-MuLV. A-MuLV was shown to contain a chimeric gene, the product of which was the major driving force for the transformation

observed in lymphoid cells. [61] The N-terminus of the chimeric gene is the Gag protein and the C-terminal of the protein corresponds the cellular protein c-Abl. (**Fig 1:10**)

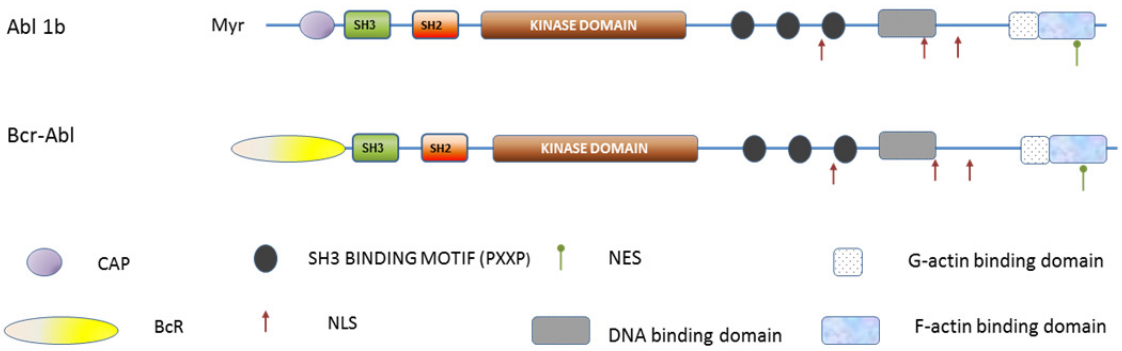


Fig 1:10 Domain organization of Abl1b and *Bcr*-Abl.

Two alternate forms of the *Abl1* transcript result in the expression of two forms of protein, Abl1a and Abl 1b. Abl 1b contains nineteen more amino acids than Abl1a and is also modified in the N-terminal by a myristoyl group.(**Fig 1:10**) Abl kinase is ~ 1100 amino acids in length and like other non-receptor tyrosine kinases of the Src and Tec family contains an SH3 and SH2 domain in the N-terminus followed by the kinase domain. However, the C-terminal region in Abl (>600 a.a) is different as it contains a number of motifs for protein-protein interaction. In this respect Abl Kinase and the related kinase Arg are unique. The number of proteins that can potentially interact with the C-terminal region of Abl include, SH3 domain containing proteins, actin, ATM, RB, p53, RNA-polII, DNA, G and F actin. (**Fig 1:10**) It also includes a number of NLS and one NES sequence. Knock out of the Abl gene results in post-partum mortality lymphipenia and osteoporosis. [62]. c-Abl responds to DNA damage and promotes apoptosis. The role for Abl in the cytoplasm is more diverse and mainly involves morphogenesis.

Given the number of interactions that this protein can make it is clear that Abl kinase plays a central role in cell signaling or acts as an integrator of various signaling inputs and orchestrates a cellular response.

In addition, the research into the regulation of Abl has been fueled by the discovery of *Bcr-Abl,* a chimeric protein is the major driving force in chronic myeloid leukaemia (CML) and in some cases acute lymphocytic leukaemia. [63] Similar to *v-Abl*, *Bcr-Abl* also displays high kinase activity and can transform Rat-1 fibroblasts and B-cell lymphocytes in culture.[64, 65] The common feature between these two variants that is different from cellular Abl kinase is a modified N-terminal region. The kinase activity of Abl is therefore closely tied to the architecture of the N-terminal region of the protein.

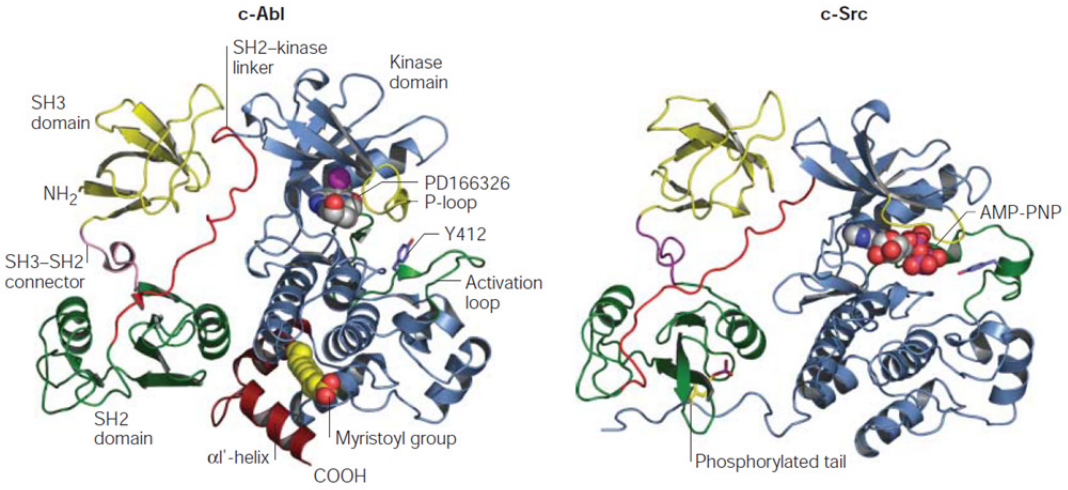


Fig 1:11 Structural features of c-Abl and c-Src. c-Abl with kinase inhibitor PD166326 and c-Src with AMP-PMP. [66]

The crystal structure of the N-terminal portion of Abl kinase helped explain the nature of auto-inhibition.[67] The SH3 and SH2 domain bind to the site opposite the active site and hold the enzyme in a closed conformation. The SH3 domain forms specific interactions with a proline rich

sequence connecting the SH2 domain and the kinase domain (SH2-kinase linker; shown in red **Fig 1:11**). The SH2 domain forms extensive contacts with the C-terminal lobe of the kinase. The general organization of this inhibited form of Abl is similar to the structure of Src kinase except that in Src Kinase the SH2 domain is locked in place by the intra-molecular interaction with a pTyr at the C-terminal. (**Fig 1:11**) Abl 1b contains a myristoyl cap (not present in *Bcr*-Abl and Abl1a)(shown in yellow) that is known to dock into the C-lobe of the kinase. The myristoyl group plays a dual role: in the localization of c-Abl in the cell to the ER membrane,[68] and suppression of the transforming activity[69]. Structurally, the burial of the myristoly group, the cap region, and the organization of the SH3 and SH2 domain is essential for c-Abl to adopt the inactive conformation. [67, 70]

The Crk family of proteins is a major binding partner and substrate of Abl kinase. CrkII was identified as a binding partner of Abl via the –P-X-X-P motifs present at the C-terminal of the kinase domain. **Fig 1:10**.[71] Co-expression of Abl and Crk not only leads to an elevated phosphorylation level of Crk but also of Abl. [72, 73] Crk phosphorylation at Tyr221 by Abl kinase inactivates CrkII and in the process dissociates signaling complexes such as p130CAS/Crk/DOCK180[74] and the p130CAS/Crk/C3G complex.[75] Tyr221 phosphorylation by Abl impairs cell adhesion [76] blocks migration retards invasive behavior in human pancreatic cancer cells[77 78] by preventing the activation of GTPases like Rap1 and Rac1. [79]

1.3 Structural basis of Crk function

The binding activity of SH2 and $SH3^N$ domains are necessary for the transforming activity displayed by Crk.[3] Functionally, CrkI is more transforming than CrkII and CrkL [4 80]. This is explained by the presence of a negative regulatory site in CrkII and CrkL. The residue Tyr221 is part of the sequence -YAQP-(221-224) and upon phosphorylation creates an internal binding site for CrkII SH2. [81, 82].

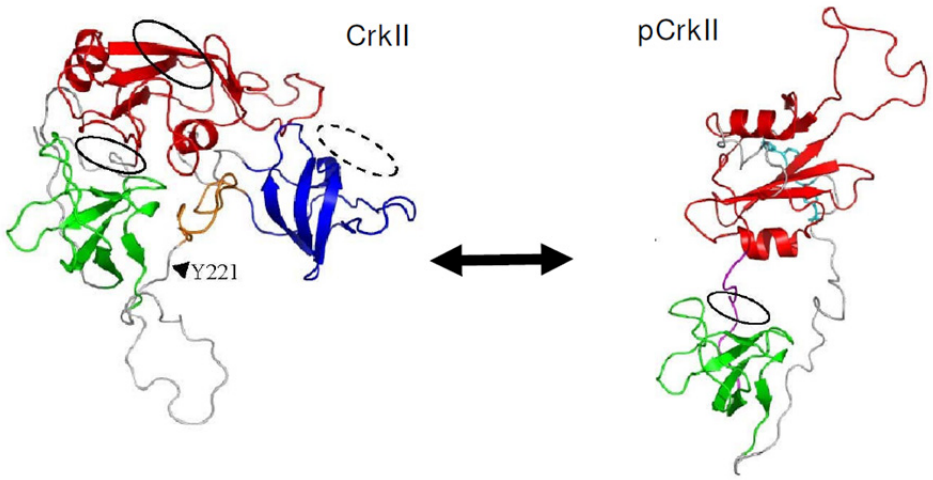


Fig 1:12 Solution structure of CrkII (left) and pCrkII(right). SH2 (red), $SH3^N$ (green),$SH3^C$ (blue). Circles depict ligand binding site. (dotted circle presumed binding site for $SH3^C$)

The structure shows that this intra-molecular interaction prevents the SH2 domain from interacting with other binding partners. It also masks the access of ligands to the $SH3^N$ domain as the linker region (Arg122-Glu133) blocks access to the binding surface of $SH3^N$. (**Fig 1:12**). The unphosphorylated form of CrkII is a compact structure that is stabilized by the hydrophobic patches of the SH3 domains, part of the linker region between the two SH3 domains (224-237) and electrostatic interactions between residues in the RT loop of $SH3^N$ and basic residues of SH2. (**Fig 1:12**). The structure reveals that the ligand binding site for the SH2 domain is solvent

exposed but the $SH3^N$ domain experiences partial inhibition. The SH2 domain masks the binding site on $SH3^N$ and this structural observation is supported by binding studies which show CrkII having a ~6 fold lower binding affinity for PPII-peptide ligands relative to the isolated $SH3^N$ domain. pCrkII in comparison, has a ~16 fold lower binding affinity for PPII-peptide ligands. CrkI does not possess a Tyr221 site and cannot be inhibited. Consistent with the biological data, the structural studies show that pCrkII is unable to participate in protein-protein interactions *via* the SH2 and $SH3^N$ domain. Interestingly, in the unphosphorylated form, the $SH3^N$ domain is also partially inhibited when compared to CrkI (or *v-Crk)* which may explain why the shorter isoform is more transforming.

1.4 The Role of SH3C domain in Crk signaling.

The $SH3^C$ domain of CrkII and CrkL has been a source of intrigue as it has no known binding partner. The $SH3^C$ domain is separated from the other domains by a long 60 amino acid linker. The $SH3^C$ domain of both these proteins is missing an aromatic residue at the binding site partially explaining the lack of evidence for binding partners. The $SH3^C$ domain of CrkII also contains more polar residues that line the binding pocket. [83] However, the $SH3^C$ domain has been shown to play a regulatory role in Crk mediated signaling. Deletion of the $SH3^C$ domain results in increased association of Crk to Abl and elevated Tyr221 phosphorylation. [84]In-fact removal of residues 242-255 from the $SH3^C$ domain relieved the inhibitory effect of the $SH3^{C\,-}$ domain and lead to increased p130Cas and FAK phosphorylation. Structural information about the domains has helped illustrate the inhibitor effect of the $SH3^C$ domain. NMR experiments carried out on the $SH3^C$ domain and the linker revealed an interesting auto-inhibitory role. Mediated by the *cis-trans* isomerization around the Gly237-Pro238 bond in the linker region preceding the $SH3^C$ domain, the domain could adopt two conformations. Fig 1:13

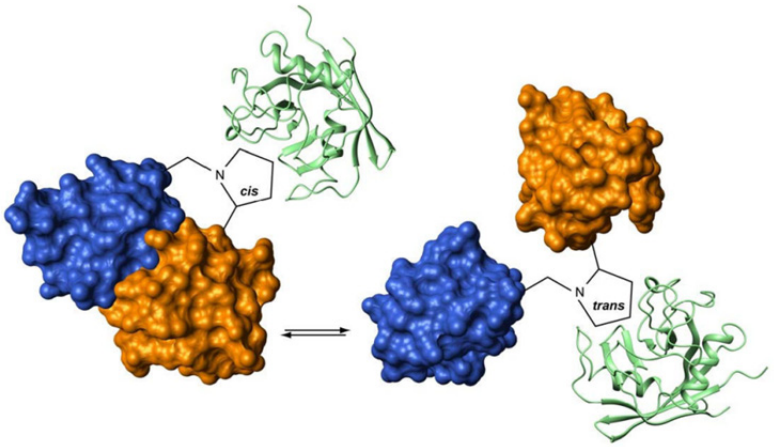


Fig 1:13 Prolyl *cis*-*trans* isomerization mediates the inhibition of the $SH3^N$ domain.

In the *cis* conformation the $SH3^C$ domain folds back over the SH3N domain and inhibit ligand binding. In the *trans* conformation the inhibition is relieved and the $SH3^C$ makes no contact with the $SH3^N$ domain. The *cis* and *trans* conformation exist in an equilibrium with 90% in the *cis* conformation and 10% in the *trans* conformation. Ligand binding to the $SH3^N$ domain shifts the population to the *trans* conformation. The population can also be changed by addition of CypA to catalyze the isomerization rate of the Gly237-Pro238 bond. [85] This observation and the subsequent structure of the $SH3^N$-Linker-$SH3^C$ construct revealed a unique mode of inhibition in CrkII.

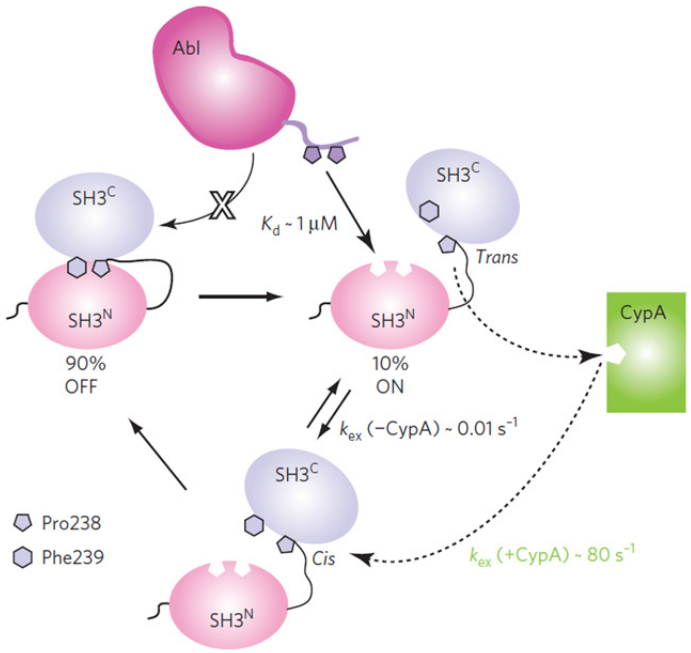


Fig 1:14 Regulation of Crk Activity. Crk exists in an equilibrium (90% inhibited:10% uninhibited) Ligand binds to the uninhibited form and drives the equilibrium towards the uninhibited form.

The existence of an open and closed conformation in CrkII has recently been verified in the full length by employing NMR spectroscopy, thermodynamics and kinetic analysis. A energy difference of ~1.2kcal mol^{-1} separate the open and closed conformations. [86]

The $SH3^C$ domains of CrkII and CrkL contain a buried nuclear export sequence (NES) which consists of the sequence –L-A-L-E-V-G-E-L-V-K-V- [87]. The $SH3^C$ of CrkII has been shown to interact with a nuclear export factor Crm1. Absence of the NES site in CrkII results in the formation of a complex with Wee1, a nuclear tyrosine kinase and promotes apoptosis. [88] Inspection of the structure of CrkII $SH3^C$ shows that this NES sequence is buried in the core of the domain and partial unfolding of the $SH3^C$ domain would be required for Crm1 to access the NES sequence[83]. The $SH3^C$ of CrkL however was shown to form a swap dimer brought about by the local unfolding of the $SH3^C$ domain [89].

The $SH3^C$ domain and linker region have been shown to play a role in the activation of Abl kinase. When residues in the RT-loop of $SH3^C$ with the sequence PNAY (248-251) are deleted along with substitution at P224A and Y221F, the mutant CrkII is unable to activate Abl kinase. [73] In addition recent studies have shown that Abl kinase phosphorylates CrkII at Tyr251 (RT-loop of $SH3^C$). A SH2 domain screen identified the SH2 domain of Abl and Arg kinase as a viable binding partner. Furthermore a phosphopeptide generated from the RT region was able to activate Abl1b. [90]

1.5 Crk proteins and disease.

Crk has been found to be overexpressed in many human cancers including various carcinomas and sarcomas. [91-93] An increase in mRNA levels of Crk is also observed in certain aggressive forms of lung adenocarcinoma.[94]In addition, tumor cells show elevated levels of Tyr221 phosphorylation. Knockdown of Crk expression in human ovarian cancer cell lines show a reduction in formation of focal adhesions, actin disorganization and reduced lamellipodia formation. Replacement of Crk reverts cells back to originial form. (Fig 1:15) Similarly, CrkL has been found to be amplified in tumor specimens including adenocarcinomas and squamous cell carcinomas. siRNA-mediated knockdown of CrkL can rescue the tumor cells from the tumorigenic phenotype

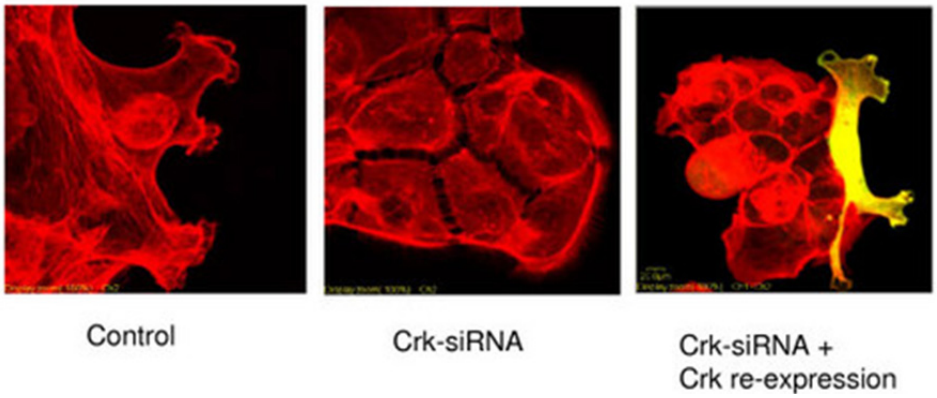


Fig 1:15 Effect of Crk knockdown by siRNA in human ovarian cancer cell line MCAS

As mentioned earlier Crk family adaptors also play a role in mediating the action of the Bcr-Abl, a major player in CML[95, 96] There are multiple reports implicating CrkL as a substrate Bcr-Abl [97, 98]and recently assays have been developed to monitor tyrosine phosphorylation level of CrkL as a diagnostic tool for Philadelphia positive leukemia. [99]

Crk adaptor proteins also play a role in bacterial infection. The *Shigella flexneri* infection was found to require Crk and Abl kinase. [100] Tyr221 phosphorylation is necessary for *Shigella flexneri* infectivity. Unc119, an adaptor protein co-localizes with Crk in *Shigella* infected cells and has been shown to inhibit infection by preventing Abl kinase mediation phosphorylation of Crk. [101] Similar roles are played by Crk in infections caused by *Pseudomonas aeruginosa* [102]and *Helicobacter pylori*[103]. The NS1 protein of the influenza A virus can bind to the $SH3^N$ domain of Crk and CrkL and this association enhances the JNK activation and suppresses apoptosis. [104, 105]

2 Nuclear Magnetic Resonance (NMR) spectroscopy.

The study of biomolecular structure, dynamics and function has been the focus of intense research for the last several decades. Structural information of biomolecules is essential for a complete understanding of how the molecules work. X-ray crystallography and NMR are two complementary methods that are routinely used to elucidate the structure of biomolecules. In fact, there is an advantage to using both these methods to elucidate the structure. NMR spectroscopy has an advantage over X-ray in that it provides dynamic information about the protein at atomic resolution. This includes obtaining information from a wide range of biologically relevant timescales (pico seconds to hours). NMR spectroscopy is also a useful tool to study interactions of biomolecules. Relevant information can be easily obtained by observing changes in chemical shifts from titration experiments and is now used in screening for small molecule inhibitors for protein targets. The process of structure calculation by NMR involves a series of steps that are briefly described below.

2.1.1 Assignment of chemical shifts of protein spectra

A combination of various factors including high field NMR with cryogenic probes, sophisticated NMR experiments, robust software and labeling schemes have led to a revolution in the approach to NMR based protein structure determination. Given the state of the current technology, solution structures of proteins upwards of 50kDa can be solved *ab initio*. The availability and easy of producing ^{15}N labeled protein samples (where every Nitrogen atom in the protein is replaced by the ^{15}N isotope) allows for a fingerprint to be obtained with ease. ^{1}H-^{15}N HSQC experiment is typically the first experiment run on protein samples. This is a two dimensional experiment that allows for the resonance of the amide protein (H) to be correlated with the attached nitrogen (N). A typical protein with 100 residues displays 100 resonances, each resonance from the corresponding amide. (proline residues do now show up as they do not contain an amide bond). The resonance of the amide bonds do not appear in any particular order and are in-fact very sensitive to the protein environment. The initial process for protein structure determination is to assign each resonance to a particular amide of the protein. This is done with the help of triple resonance three dimensional NMR experiments carried out on protein samples labeled with ^{13}C and ^{15}N (naturally occurring ^{14}N and ^{12}C isotopes are not NMR active).

The standard approach for protein backbone assignment includes acquiring triple resonance experiments such as HNCA, HN(CO)CA, HNCO, HN(CA)CO, HBHA(CACB)NH. These experiments rely on the transfer of magnetization transfer between the ^{13}C and the ^{15}N nuclei through bonds, made possible by the large J-couplings Fig 2:1. The scalar couplings between the nuclei make the sequential assignment of the polypeptide chain possible. The experiments provide intra and inter residue correlations to enable the assignment process.

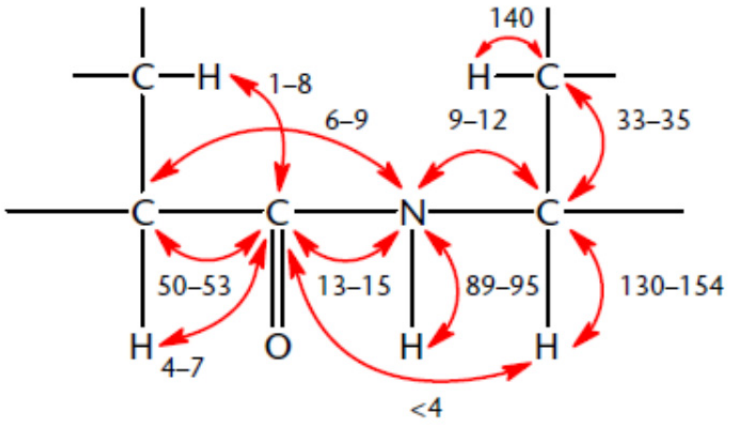


Fig 2:1 ^{1}J and ^{2}J coupling constants for magnetization transfer in ^{13}C,^{15}N-labeled proteins.

An HNCA experiment correlates the ^{13}Cα chemical shift of a residue with the amide proton and the ^{13}Cα chemical shift of the preceding residue (*i-1)* [106, 107]. An HN(CO)CA experiment provides the ^{13}Cα chemical shift of just the preceding residue. Comparing these two spectra allows one to rapidly correlate the chemical shifts of residues that comprise the polypeptide. As displayed in Fig 2:2, in and HNCA experiment the transfer of magnetization takes place from ^{1}H to ^{15}N to the ^{13}Cα via the N-^{13}Cα J- coupling and then back to the ^{1}H via ^{15}N. Because the amide ^{15}N is coupled to its own ^{13}Cα and the ^{13}Cα of the preceding residue, both resonances are observed in the spectra. The strong coupling of the Cα atoms allows for greater intensity of peaks in the spectra and is therefore a common choice for assignment purposes. Similarly, the HN(CO)CA experiment transfers the magnetization in the following manner: the ^{1}H -^{15}N-^{13}CO-^{13}Cα. However in this experiment the resonance is allowed to evolve for the ^{13}Cα, ^{1}H and ^{15}N nuclei only. The ^{13}Cα chemical shift of the preceding residue is observed in this spectrum.

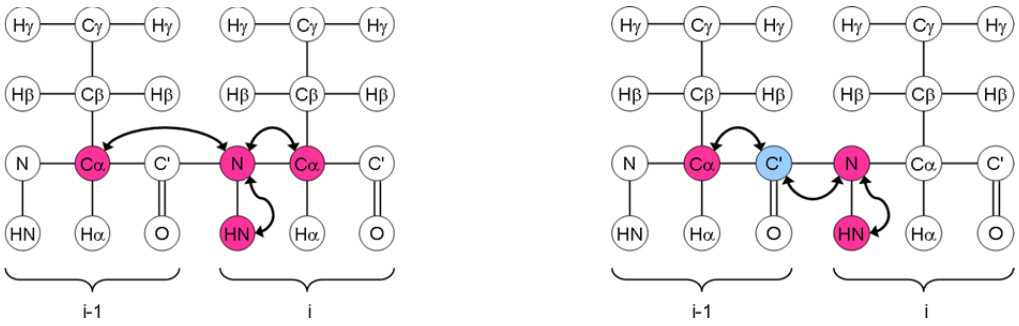


Fig 2:2 Magnetization transfer in an HNCA(left) and an HN(CO)CA(right) experiment.

In a similar fashion, an HNCACB can be used where each NH group is correlated with both the Cα and the Cβ values of the own and previous amino acid. For each amino acid, the Cα and the Cβ values are characteristic. For e.g. Alanine residues usually show a Cβ chemical shift at ~ 20 ppm. Serine and threonine residues have a Cβ value that is always higher than the Cα value. A proline residue is also easily identified because it does not possess a true amide proton and thus in these types of experiments the resonances of a residue preceding proline cannot be observed.

2.1.2 Assignment of side-chains

The assignment of the back-bone allows for the side-chain resonances to be assigned. There are various methods to achieve this. The typical experiments used are HBHA(CO))NH, HC(CO)NH and CC(CO)NH. [108] The quality of these spectra greatly depends on the size of the protein and allowed time. For certain proteins a HCCH-TOCSY along with HCCH-COSY spectra can be used.

The CC(CO)NH experiment requires ^{13}C,^{15}N-labeled samples and magnetization is transferred from the side-chain hydrogen nuclei to the attached ^{13}C nuclei. Using isotrophic ^{13}C mixing, magnetization is transferred between the carbon nuclei to the carbonyl carbon then the amide

nitrogen and then detected at the amide proton. [109, 110](Fig 2:3) This experiment provides a certain convenience during assignment as the side chain resonances appear in a similar manner to which the back-bone assignment was performed. This can also be used to reinforce the initial back-bone assignments performed.

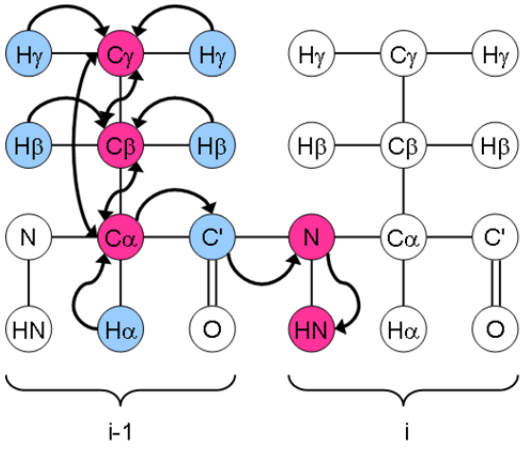


Fig 2:3 Magnetization transfer in a C(CO)NH experiment.

In a similar manner in the H(CCO)NH isotropic ^{13}C mixing allows for the magnetization to transfer between the carbon nuclei then transferred to the amide proton via the carbonyl carbon and the amide nitrogen. The chemical shift is allowed to evolve simultaneously in hall side chain protons. [111]. This allows for all the side-chain protons to be assigned.

2.1.3 Nuclear Overhauser Effect Spectroscopy (NOESY)

The nuclear Overhauser effect (NOE) is an important element in the structure determination process of proteins. A NOESY experiment as it is commonly referred to provide a cross-peak between two protons that are no more than 5Å apart. The NOE is dependent on the distance between the two protons in space and therefore provides structural information. In contrast to

other types of NMR experiments, NOE experiments can be directly related to the protein conformation. THE NOE originates from dipolar interactions between different nuclei. The magnitude of the NOE is inversely dependent on the sixth power of the internuclear distance ($1/r^6$)[112] Two types of NOESY experiments are usually performed. ^{15}N-NOESY-HSQC allow for the magnetization of to evolve in all the protons in the protein and then then transferred to the amide proton *via* the amide nitrogen. The mixing time is an important component of the experiment and usually 80ms is used. The spectra provide chemical shift information for all protons within a 5Å distance. Fig 2:4

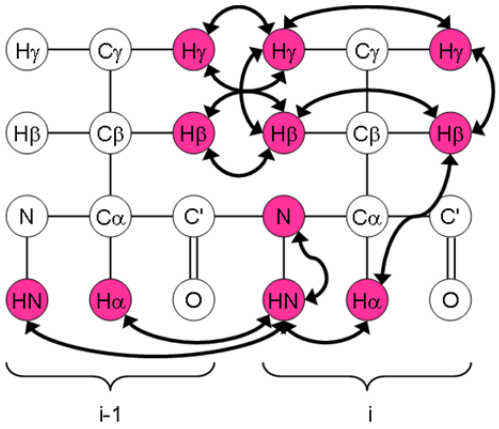


Fig 2:4 Magnetization transfer in a ^{15}N-NOESY-HSQC experiment.

A similar experiment that is an absolute necessity for structural calculations is the ^{13}C-NOESY-HSQC experiment. The magnetization is exchanged between all protons using NOE and then transferred to the neighboring ^{13}C and then back to the proton for detection. The transfer occurs from the aliphatic ^{13}C or the aromatic ^{13}C nuclei. As the Hα chemical shifts of amino acids usually fall in the 4.7ppm range where the water signal resides, the spectral quality improves greatly when acquired in 100% D_2O buffer. The amide protons are usually exchanged and cannot

be observed. This experiment is also useful to assign the protons in the aromatic amino acids. [113] [114, 115]

2.1.4 Protein deuteration

The NMR spectra of proteins tend to deteriorate as the protein size increases (<20kDA) as they tumble slower in solution leading to intrinsically higher $R_{2.}$ A useful technique to circumvent this problem is to record spectra at higher temperatures if the sample allows it. Another approach that was successfully used involved encapsulating the protein in a low viscous solvent [116]. The source is this enhanced relaxation is the high density of protons present in the protein so the replacement of protons with chemically identical isotope Deuteron has proved useful. [117, 118]. This leads to an improvement in sensitivity and resolution. [119] The replacement of deuteron for proton can be achieved easily for recombinant proteins expressed in *E. Coli.* The media for cell cultures is made it D_20 instead of H_20. The amide protons are usually replaced during protein purification as they exchange rapidly with the buffers. For many large systems, this significantly expensive approach of deuteration, has led to the ability to study larger protein systems that were otherwise impenetrable to NMR techniques. The underlying cause of the spectra improvement is based on several factors: i) deuteration decreases the homonuclear proton/proton relaxation pathways as there are less number of ^1H in the sample. The reduced loss of signal due to spin diffusion results in a narrow line width of the observed resonances. ii) The replacement of ^2H for ^1H ($\gamma_D/\ \gamma_H \sim 1/6.5$) causes significant reduction of the heteronuclear dipolar relaxation of the ^{13}C spin. This leads to a signification improvement in triple resonance experiments that involve the ^{13}C spin systems. (e.g 3D-HNCA).

2.1.5 Paramagnetic Relaxation Enhancement (PRE)

To supplement NOE data, paramagnetic relaxation enhancement (PRE) experiments can be employed to provide additional distance restraints needed for protein structure determination. PRE is a very powerful tool for the structural and large-scale dynamic studies of biomolecules in solution. The PRE effect is based on the magnetic dipolar interactions between a given nucleus (e.g., ^{1}H) and the unpaired electron(s) of a nearby paramagnetic center. PRE increases the relaxation rates of the nuclear magnetization. Similar to NOE, the magnitude of PRE is inversely proportional to the electron-nucleus distance (1/r6). Due to the large magnetic moment of the unpaired electron, the observed PRE effects are large and, thus, can provide long-range distance restraints (up to 34 Å). In most cases, nitroxide spin labels (e.g., MTSL) are well suited for PRE measurements. [120, 121] Another advantage of using spin labels or Mn^{2+} measuring PRE is that the g-tensor of the paramagnetic center is isotropic and therefore no pseudo-contact shifts (PCS) are observed, as opposed to the use of lanthanides [122]

The nitroxide spin label, MTSL, is covalently attacked to the side chain of a surface exposed Cysteine. As with any labeling strategy, this process involves the removal of other surface exposed cysteine residues so that each sample can bind one MTSL group. This simplifies data interpretation. [108]

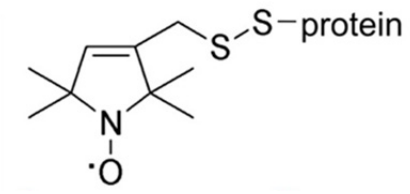


Fig 2:5 Chemical structure of spin label reagent MTSL attached to a protein *via* disulphide bond.

The analysis of the PRE effect is carried out on the basis of a single time-point measurement. ^{1}H-^{15}N HSQC experiments are recorded one with the paramagnetic state and the other with the diamagnetic state, in identical conditions. The relaxation rates can then be extracted from the HSQC spectra using the ratio of cross peak intensities in the paramagnetic and diamagnetic states (I_{para}/I_{dia}) using the following equation

$$\frac{I_{para}}{I_{Dia}} = \frac{R_{2,dia}\exp(-tR_{2,para})}{R_{2,dia}+R_{2,para}}$$

Equation 2-1

In the above equation *t* is the total INEPT evolution type, $R_{2,dia}$ and $R_{2,para}$ are the intrinsic and spin relaxation rates. The distance is calculated using the following equation:

$$r = \left[\frac{K}{R_{2,para}}\left(4\tau_c + \frac{3\tau_c}{1+\tau_h^2\omega_h^2}\right)\right]^{1/6}$$

Equation 2-2

where *r* is the distance between the electron and the nuclear spins τ_c is the correlation time for this interaction, ω_h is the Larmor frequency of the nuclear spin and K is the a constant with a value of 1.23 x 10^{-23} cm^6 s^{-2} . [123]

PRE methods have been applied in various studies which include protein folding studies [124], analysis of protein-protein and protein-nucleic acid complexes [125, 126], structural analysis of membrane pore forming proteins [127] and quantifying low-population states in proteins. [120].

2.1.6 TALOS

TALOS is a database system used to empirically predict *phi* and *psi* backbone torsion angles for a given protein sequence using as a source a set of several chemical shifts (HA, CA, CB, CO, N) Fig 2:6. [128]

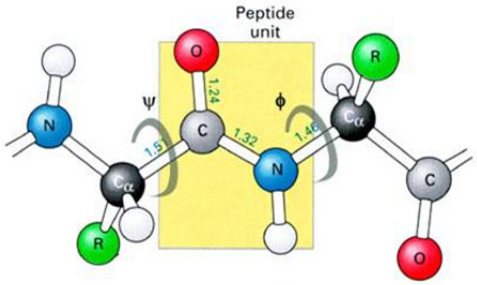


Fig 2:6 Location of the dihedral angles in the polypeptide.

TALOS searches the database for triplets of adjacent residues with secondary chemical shifts and sequence similarity in order to provide the best match to the query 40 triplet of interest. The database contains $^{13}C\alpha$, $^{13}C\ \beta$, ^{13}C ', $^{1}H\alpha$ and ^{15}N chemical shifts for 20 proteins with available high-resolution X-ray structure. TALOS makes quantitative predictions for the protein backbone angles (*phi* and *psi*) estimates the uncertainty for these predictions. The output of TALOS is 10 triplets that have the closest similarity in the secondary chemical shift and amino acid sequence to those of the query sequence. If the central residues in these 10 triplets show similar backbone angles, their averages can be reliably used as angular restraints for the protein being studied. An enhanced version, TALOS+ is now also available. [129]

2.1.7 Residual dipolar coupling (RDC)

Residual dipolar couplings (RDCs) allow for the relative orientation of interacting proteins to be determined by measuring the RDC between nuclei in a partially oriented media. Briefly, the dipolar couplings of the nuclei in solution averages to zero due to the effect of Brownian motion. In order to tease out the coupling, anisotropic conditions are required. The dipolar coupling in solution between two nuclei, A and B can be estimated by the following equation:

$$D_{AB}(\theta,\phi) = A_a^{AB}\left\{(3\cos^2\theta - 1) + \frac{3}{2}R(\sin^2\theta\cos 2\phi)\right\}$$

$$A_a^{AB} = -\left(\frac{\mu_0 h}{16\pi^3}\right) S\gamma_A\gamma_B\langle r_{AB}^{-3}\rangle A_a$$

Equation 2-3

A_a^{AB} and R are the axial and rhombic components. Θ is the angle between the inter-nuclear bond vector and the *z*-axis of the alignment tensor, φ is the angle between the projection of the inter-nuclear bond vector onto the *x*-*y* place and the *x*-axis; μ_o the permeability in vacuum, *h* the Plank's constant, S the generalized order parameter, γ_A and γ_B the gyromagnetic ratio of the two nuclei, r the time averaged inter-nuclear distance and A_a the axial component of the molecular alignment tensor. [130] Dipolar couplings in solids are typically in tens of kHz but is orders of magnitude lower in liquid-crystalline media (<100Hz). In theory any media that generates a weak alignment for the protein sample can be used. Bicelles like DMPC and DHPC[131], filamentous phages[132], polyacrylamide gel and a poly(ethyleneglycol)/hexanol [133]mixture are common solvent choices. Charged molecules can also be used as an alignment media. [134].

N-HN dipolar couplings can be measured using IPAP ^{15}N-HSQC experiment where the components of the multiplets are separated while the 3D (HA)CA(CO)HN and HNCO can be used to detect Cα-Hα, and Cα-C‟ couplings. These measurements are more accurate because the ^{1}H gyromagnetic ratio is large and the bond length is short. The IPAP experiment acquires two spectra, one where the coupling evolves in-phase (IP) and one in anti-phase (AP). The addition and subtraction of the IP and AP spectra give two simplified spectra where the ^{15}N chemical shifts are displaced by either $+J_{NH}/2$ or $-J_{NH}/2$. The coupling is measured easily from the difference in frequency between the peaks in the two specta. [135].

2.1.8 Protein structure calculation

Solution structure from obtained from NMR involves a process of using the experimental information obtained (NOE, PRE and RDC) and converting them into a visible structure. A randomly folded starting structure is generally used to begin the calculations. The structure calculation software produces a set (ensemble) of structures consistent with the experimental NMR data. The goal is to produce a structure that satisfies the experimental and empirical data with minimal overall energy. A successful structure calculation is determined by the level of constraint violations in describing the molecular geometry of the protein. NMR structures are usually represented as a set of 20 minimal energy structures and the success stems from the overall consistency in the set of structures. A rote-mean-square (rmsd) for the atomic coodinates between the structures is usually used as an indicator of the level of confidence in the solved structure – the smaller the rmsd, the more accurate the structure. Additional

verification programs are applied to structure to ensure there is no unusual geometry adopted by the atoms. These include the Ramachandran plot to determine whether the backbone ϕ and ψ angles fall into the favored regions. [136]. There are several programs that can be used for calculating solution structures based of NMR DATA. We have employed XPLOR-NIH for our calculations. [137]

2.1.9 Protein dynamics

NMR is a unique tool for characterizing the dynamics of a biomolecule. The scope of molecular motions with respect to the NMR time scales is summarized in the figure below.

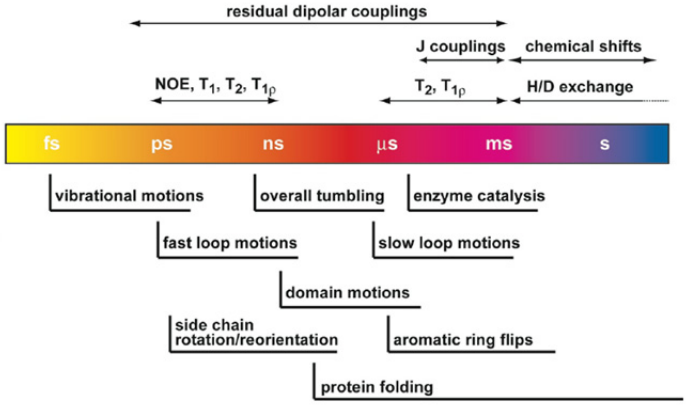


Fig 2:7 Time scales for protein dynamics and NMR techniques.

Dynamics play an important role in ligand binding, molecular recognition, allostery and catalysis. [138]. Powerful techniques exist in solution NMR to investigate these dynamic properties at the various timescales. The dynamics (or protein motion) generally affect the relaxation rates of the spin systems and is related to the timescale of the motions. Overall rotation of protein molecules occurs in the nano second (ns) time scale, fast loop motions in the pico second (ps) time scale and domain motions and catalysis in the micro-milli (µs-ms) time scale. The ps-ns

time scale are usually referred to as the fast time scale motions and the µs-ms time scale motions as slow time scale motions. The order parameter, S^2 is a measure of the angular fluctuation of a bond vector and reports on the flexibility of the protein at that site in the pico-to-nanosecond (ps-ns) timescale. [139]. The value can vary from S^2=1, for completely rigid, to S^2=0, for freely moving. [138]. The determination of order parameters require three measurements, the longitudinal relaxation rate (R_1), the transverse relaxation rate (R_2), and the heteronuclear NOE. Conformation change in proteins can lead to broadening of NMR signal. This is usually an indication that the protein is undergoing chemical exchange in the µs-ms timescale. The presence of the other conformations can be detected from measuring the broadening of the linewidth. Carr-Purcell-Meiboom-Gill(CPMG) experiments have been designed to study these systems. The exchange process usually leads to a dephasing of the coherent magnetization. This can be refocused by applying 180° pulses separated by a time τ_{cp}. The relation between R_2 and the strength of the refocusing field τ_{cp} can be used to obtain the difference in chemical shift of the conformers, their exchange rates and relative populations. [140].

2.1.10 Ligand binding studies by NMR.

Interactions of proteins with other macromolecules or small molecules play important roles in all biological processes. NMR spectroscopy is a powerful technique that can be used to characterize protein complexes under physiological conditions at atomic level even if the interactions are weak and transient[141, 142]. NMR is an excellent tool for characterizing protein interactions e.g. protein/protein, protein/nucleic acid, protein/ligand or nucleic

acid/ligand interactions without the need for high-resolution structure determination. The principle behind this approach, known as NMR mapping, is that titration of a ligand that interacts with a molecule induces changes in its NMR spectrum for signals of atoms near the binding site. A variety of experiments have been developed to study biomolecular interactions such as chemical shift perturbation mapping, saturation transfer difference spectroscopy (STD), differential line broadening, the transferred cross-saturation experiment (TCS) [143], isotope labeling methods [177, 178] (isotope filtered NMR experiments) and [144]residual dipolar coupling (RDCs) [108]. Chemical shift perturbation is the most widely used NMR method to map protein interfaces [141]. The technique is based mainly on analyzing ^{1}H-^{15}N HSQC or the ^{1}H-^{13}C HMQC spectra. The chemical shift in both nitrogen and proton dimension is sensitive to the chemical environment of the two nuclei. The 15N-1H HSQC spectrum of one protein is monitored when the unlabeled interaction partner is titrated in, and the perturbations of the chemical shifts are recorded [145]. Although, the chemical shifts provide structural information the evidence of the interaction can be determined. Chemical shift perturbation measurements solely provide information about the interacting surfaces of the binding partners.

2.2 Fluorescence Resonance Energy Transfer (FRET)

Fluorescence Resonance Energy Transfer (FRET) is a physical process where energy is transferred from donor fluorophore to an acceptor fluorophore. The efficiency of this transfer is depended on several factors: (i) the distance between the two fluorophores, (ii) the extent of

the overlap between the donor emission and acceptor excitation spectra, (iii) the quantum yield of the donor, and (iv) the fluorescence lifetime of the donor. [146] Along with FRAP and FlIM, this technique has become the standard for live cell imaging and is used to monitor the spatial and temporal information about biomolecules in cells. FRET can be used to follow proteins inside cells in real-time and follow biochemical reactions. [147]. A diagram of how the FRET energy transfer occurs during protein-protein interaction is shown in Fig 2:8.

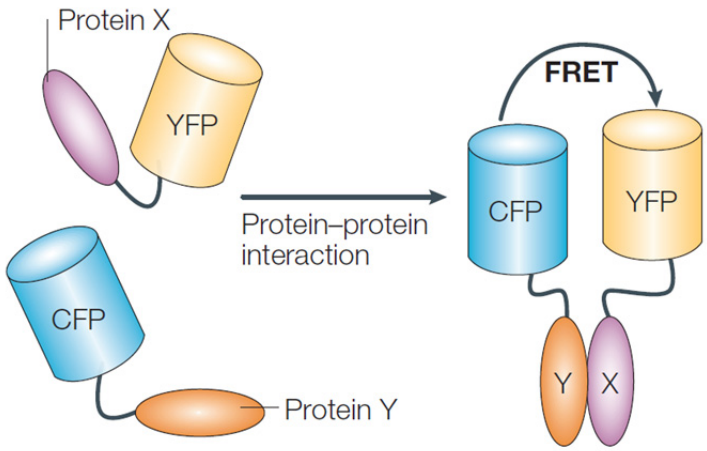


Fig 2:8 Intermolecular FRET: Protein X is labeled with yellow fluorescent protein (YFP) and protein Y is labeled with cyan fluorescent protein (YFP).

FRET transfer occurs when the donor and acceptor are within 10nm of each other and usually in cells this distance indicates protein-protein interaction. The FRET efficiency, E is the quantum yield and depends on R_0, the Försters distance. ($E \sim 1/R^6$) The success of the FRET experiment is contingent upon the choice of fluorophores. Sufficient overlap between the emission of the donor and the excitation of the acceptor must be present. The excitation and emission profile of YFP and CFP are shown below.

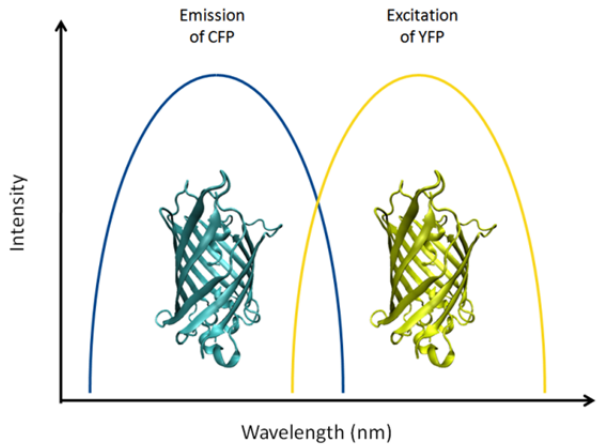


Fig 2:9 Excitation and emission profiles of donor CFP and acceptor YFP.

CFP is excited at 436nm and emits at 480nm. YFP is excited at 514nm and emits at 538nm. There are several methods to measure FRET in cells. Various software packages are now available to detect and calculate E. An important point to note here is that the loss of excited-state energy via FRET means that the quantum yield of the donor decreases and the acceptor molecule is excited and emits photons. Therefore, a decrease on the donor spectrum and the increase in the acceptor spectrum reveals the occurrence of FRET.

The simplest method is ratiometric imaging which involves exciting the donor fluorophore, and measuring the emission of the donor and the acceptor. A plot of the ratio of the two for each pixel of the image reveals the occurrence of FRET. [148] The ratiometric method has its limitations as the data is not reliable if the concentrations of the two fluorophores are not the same. The sensitized emission method can correct for this discrepancy by using a specific set of measurements of the donor and acceptor fluorescence. [146, 149]. Another technique called 3-cube FRET which takes into account the donor spectral bleed into the acceptor channel and the

direct acceptor excitation. Three sets of measurements are necessary for this technique as the baseline values are subtracted from the measured intensity to obtain the real FRET intensity of the experiment. [150]. Irrespective of the method, the popularity of FRET is based on the relatively simple protocol with minimal instrumentation and the availability of a wide range of FRET probes.

2.3 Isothermal Titration Calorimetry (ITC)

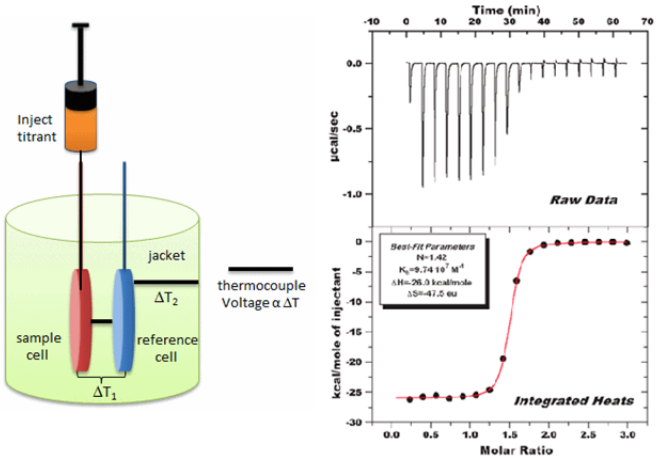


Fig 2:10 Layout of an ITC machine (left). Sample titration plot from an ITC experiment (right).

Isothermal titration calorimetry (ITC) is widely used for calorimetric measurements in biomolecular studies. ITC measures the heat evolution (or absorption) during the course of molecular association. The appeal of this technique lies in the fact that several physical parameters of a system can be obtained from one study. No other biophysical technique can claim this. These parameters include the standard free energy change (ΔG), the enthalpy change (ΔH), the entropy change (ΔS), and the stoichiometry (n) of the association event. Furthermore,

the heat capacity change (Δ*Cp*) of the binding reaction can be obtained from the variable temperature experiments. [151] The Gibbs free energy (Δ*G*) is the most important thermodynamic parameter of the binding event, because it determines the stability of any given biological complex. It can be calculated from the following equation:

$$\Delta G = -RTlnK_B \qquad \text{Equation 2-4}$$

where R is the universal gas constant, T is the temperature and K_B is the equilibrium binding constant.

The enthalpy change of binding (Δ*H*) is a measure of the disruption of protein-solvent hydrogen bonds and van der Waals interactions, formation of protein-ligand bonds, salt bridges and van der Waals contacts, and solvent reorganization near protein surfaces.

While these values are measured, the binding entropy (Δ*S*) can be calculated by the equation:

$$\Delta G = \Delta H - T\Delta S \qquad \text{Equation 2-5}$$

Hydration is the main driving force of Δ*S* and a positive Δ*S* is a strong indication that the water network at the interface has been disrupted due to protein-ligand binding. [152] The heat capacity change (Δ*Cp*) is almost always negative when the complexed state of the biomolecule is treated as a reference state. Δ*Cp* becomes substantial when water is released from the surface and is proportional to the surface area involved. Therefore, the heat capacity can translate thermodynamic data into the structural information. [153]. The stoichiometry (*n*) is usually derived from the molar ratio of the interacting species at the equivalence point.

3 Research Outline

The study of Crk family of adaptor proteins has provided the basis for understanding the role played by adaptor proteins in cell signaling. Given the multitude of proteins the adaptor proteins can interact with, predicting an outcome for a specific signaling output remains challenging. The difficulty is magnified by the fact that signaling output is context dependent where the observed phenotype differs in different cell types for a given input. The work presented here attempts to take a structural approach to address some important questions in the area of Crk adaptor proteins. The goal of the current research is as follows:

1) **Biophysical characterization of adaptor protein CrkL**

The absence of quantitative data on the binding properties of the modular adaptor protein CrkL makes comparative studies with CrkII challenging. We use ITC and NMR to explore the underlying basis of substrate selection by Crk and CrkL.

2) **Study the structural and dynamic properties of CrkL and pCrkL**

It is postulated, based on the high sequence homology, that CrkII and CrkL have similar structural and dynamic properties. We use NMR to determine the structure of CrkL and phosphorylated CrkL(pCrkL). Solution structures provide information in atomic resolution and can provide valuable insight into the mechanistic basis of CrkL signaling.

3) **Investigate the basis of proline mediated heterogeneity in CrkII.**

The linker region in CrkII has been shown to undergo *cis-trans* isomerization. The effect of this isomerization in the full-length protein has not been studied. We use NMR to investigate this phenomena in an effort to gain functional insight into the role of isomerization in CrkII signaling

4 Structural determination of CrkL

4.1 Introduction

The high sequence similarity and domain architecture of CrkII and CrkL alludes to similarity of function and structure of the two proteins. CrkII and CrkL are of comparable length (CrkII-304 aa and CrkL-303 aa) and share a 50% sequence identify overall and a 76% identity between the domains. It is not surprising that biochemical and cellular assays identify proteins that are able to interact with both CrkII and CrkL. Both proteins require the SH2 domain and the SH3 domains for biological activity. However, several lines of evidence suggest that these two proteins are not functionally redundant. CrkII and CrkL knock-out mutants are perinatal lethal but display different developmental defects. CrkL knock-out mice show a phenotype that is characteristic of DIGeorge/velocardiofacial syndrome with cranial and cardiac developmental defects. [154, 155]. CrkI and CrkII null mutants display defects in cardiac and skeletal development. [156] As Crk and CrkL knockouts are both embryonic lethal, it suggests that the proteins have non-redundant function. Studies into the Reelin pathway suggest that both CrkII and CrkL are essential for the propagation of the signal. Absence of CrkII and CrkL causes rapid Dab1 degradation and a termination of Reelin based signaling.[157, 158]. A modeling study based on the CrkII structure determined that the CrkL SH2 domain has a ~30 fold higher affinity for pTyr 463 of FGFR1 than the CrkII SH2 domain. [159] The study presented below challenges the hypothesis that CrkII and CrkL are structurally similar. We use solution NMR to obtain a structure of CrkL and phosphorylated CrkL (pCrkL). The study reveals several interesting differences between CrkII and CrkL.

4.2 Results

4.2.1 Analysis of CrkL using NMR

The full length CrkL protein (1-303) is ~35kDa protein and traditional NMR methods are not applicable. The main reason for this is that at large molecular weights the magnetization relaxes rapidly and as a result the signal is broadened. An approach that has been used quite successfully is 'divide and conquer'. This involves studying smaller fragments of the protein initially and then transferring the information to the full-length protein. Initially, we assigned the NMR resonance of individual domains: SH2 (1-105), $SH3^{N}$ (127-180), L-$SH3^{C}$ (188-303) and the $SH3^{C}$. Each of the domains was expressed and a ^{1}H-^{15}N HSQC experiment was acquired. (**Fig** 4:1)This initial analysis revealed some interesting results. The $SH3^{C}$ of CrkII does not bind the canonical polyproline II (PPII)-type sequence as it lacks an aromatic residue (Trp) at the binding site. [83] The $SH3^{C}$ domain of CrkL does not have a known binding partner. A crystal structure of the $SH3^{C}$ domain reported that the function of the domain was to promote dimerization of the CrkL molecule[89]. Our size-exclusion chromatography data and our NMR data did not confirm these findings. We find the $SH3^{C}$ domain to be exclusively monomeric at high concentrations (1mM) and to be well folded as judged from the quality of the spectra.

Our lab had published the structure of the SH3-L-SH3 (135-297) domain and the fl-SH3 (190-297) domain of chicken CrkII. This study revealed a novel regulatory role for the $SH3^{C}$ domain in that it functions as an inhibitor for the $SH3^{N}$ domain. This inhibition was shown to be mediated by the *cis-trans* isomerization of the Gly-237-Pro238 bond. (Fig 1:13) Although CrkL has high sequence similarity to CrkII, it showed no evidence of heterogeneity in the linker region thus negating the possibility of a similar type on inhibition scheme as observed in CrkII **Fig** 4:1.

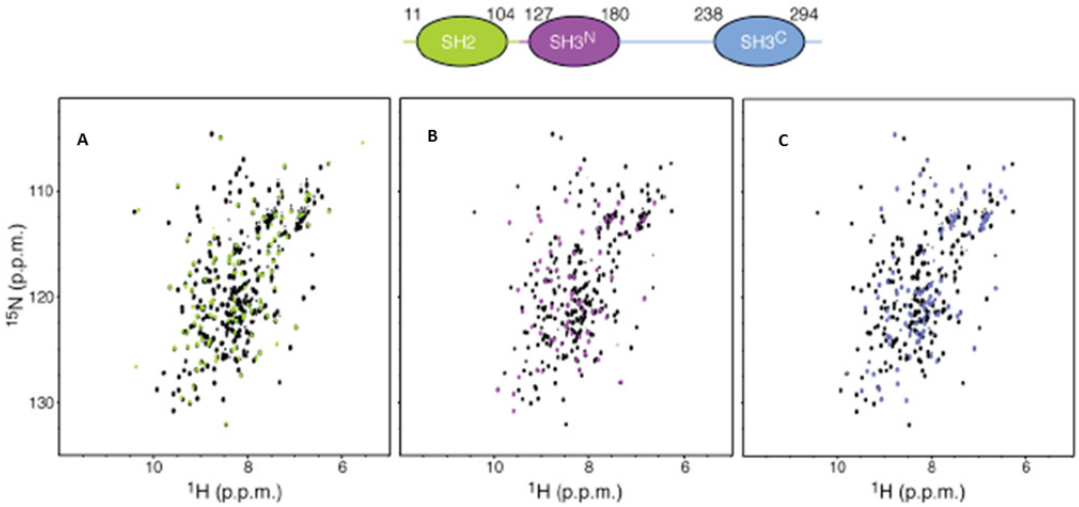


Fig 4:1 1H-15N HSQC spectra of CrkL domains overlayed on full-length CrkL. The cross-peaks of each domain follows the color code of the schematic.

A standard set of multi-dimensional through-bond heteronuclear scalar correlation experiments were carried out to achieve the complete backbone and side-chain chemical shifts. ^{1}H, U-^{13}C/ ^{15}N samples for the domains were prepared and a set of 3D-HNCO, 3D-HNCA, 3D-HNCACB experiments yielded the necessary information. A set of 3D-HN(CCO)NH and 3D-C(CO)NH, experiments enabled the assignment of the side-chain chemical shifts. Although initial experiments were carried out at a temperature of 295K, we soon discovered that the quality of the 3D experiments improved significantly at 303K. The protein samples proved to be stable at this temperature and this was used for future experiments.

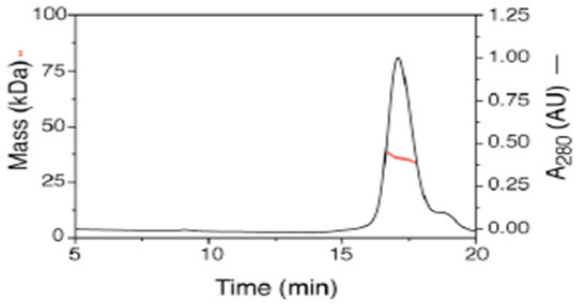


Fig 4:2 Multi-angle laser light scattering (MALLS) shows that CrkL is monomeric in solution

The full-length CrkL (1-303) recombinant protein was monomeric even at concentrations of ~1mM Fig **4:**2. There was significant overlap in the chemical shifts of full length CrkL and the individual domains. A majority of the chemical shifts were transferred to the full length spectra and later confirmed by preparing a deuterated sample of the full length protein. Even at 35kDA, deuteration significantly improved the spectral quality and provided excellent spectra. To further confirm the assignment, we selectively labeled the backbone ^{15}N amides of full-length CrkL with ^{15}N –labeled Lys, Ala, Tyr, Val, Iso, Leu, Phe using an auxotrophic strain (CT-19). Arginine was observed using reverse labeling.

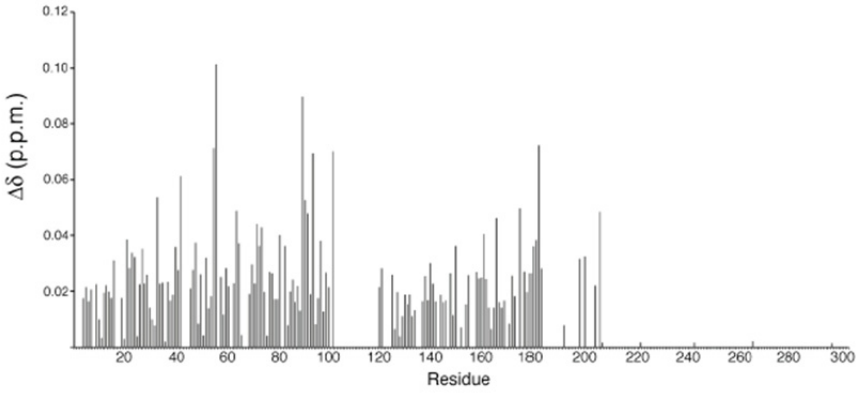


Fig 4:3 Chemical shift differences (Δδ) plotted as a function of residue number between isolated domains and full length CrkL.

The assignment of the respective domains and the transfer of assignments to the full-length CrkL (1-303) protein allowed us to compare the chemical shifts of each residue in the context of the isolated domain and the full-length protein. Several residues from the SH2 and $SH3^{N}$ domain are affected while the residues from the $SH3^{C}$ domain experience no changes. This suggests in CrkL(1-303) the $SH3^{C}$ domain does not interact with either of the two domains (**Fig 4:3**.)

4.2.2 Structure of CrkL and pCrkL

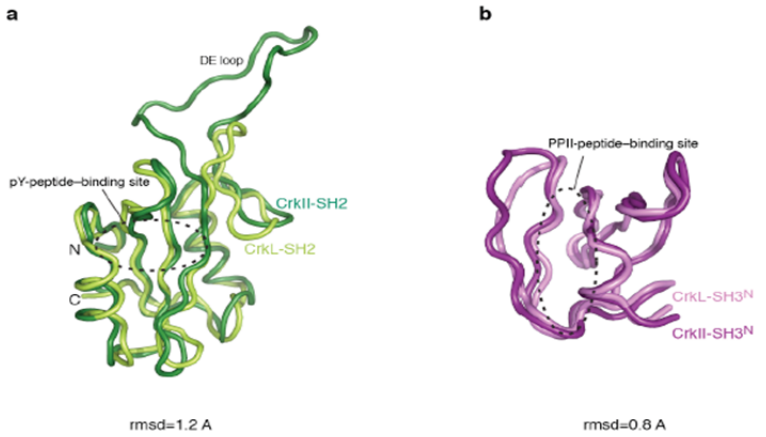


Fig 4:4 (a) Overlay of the structures of CrkL SH2 (this work) and CrkII SH2 domains. The DE loop, which was shown to bind to the SH3 domain of Abl, is present only in CrkII. **(b)** Overlay of the structures of CrkL $SH3^N$ (this work) and CrkII $SH3^N$ domains. Overall the structures are very similar other than some structural heterogeneity in the loops.

The analysis provided us with evidence that the structure of CrkL was indeed different from that of CrkII. To solve the solution structure of CrkL (1-303) we decided to follow a strategy that was successfully used for CrkII. [160] NOE based distance restraints obtained by using 3D ^{15}N-edited NOESY-HSQC and 3D ^{13}C-edited NOESY-HSQC experiments allowed us to solve the structure of the individual domains. Superposition of the individual domains to the respective domains of CrkII revealed little difference. The binding pockets of the SH2 and $SH3^N$ domains are nearly identical which is not surprising, given that they share high sequence similarity and show similar affinity for cellular ligands.(Fig 4:4) Both the SH2 and $SH3^N$ domains have the canonical fold. The SH2 domain of CrkII contains a seventeen residue long proline rich linker between the β-strands D and E (DE loop) which forms a binding site for the SH3 domain of Abl kinase[161]. The linker is absent in the SH2 of CrkL explaining its inability to bind to Abl SH3. (**Fig 4:4a**)

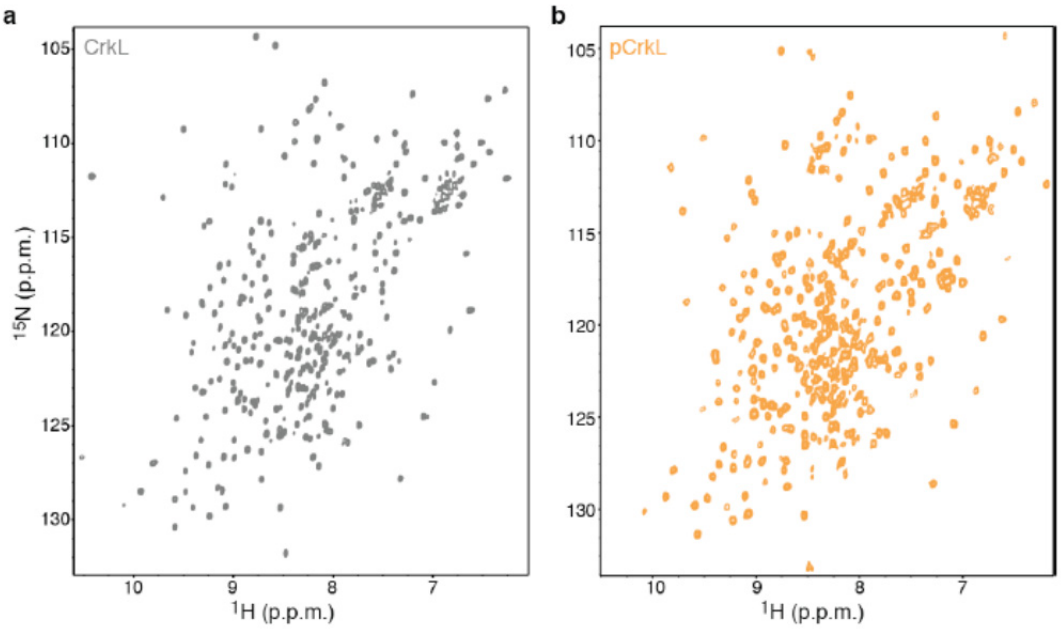


Fig 4:5 ^{1}H-^{15}N HSQC spectra of full length CrkL(1-303) and pCrkL

The solution structure of CrkL (1-303) was solved with the help of NOE, paramagnetic resonance and residual dipolar coupling. Paramagnetic resonance enhancement (PRE) measurements were made using site directed spin labeling. These experiments provide long range distances restraints information (<~30Å). To achieve these labels, sites were generated on CrkL by a single amino acid to Cysteine to which a nitroxide spine label (MTSL) can be attached. The PRE effect can be converted to distance by measuring the intensity ratio of paramagnetic over diamagnetic samples. [123] CrkL contains two Cys residues (C44, C249) which were mutated to serine, an isosteric mutation. These changes did not alter the CrkL structure as observed by ^{1}H-^{15}N HSQC. Cysteine residues were then introduced by point mutation at the following sites: Ser20, Ile90, Asp133, Gly147 and Val175. These sites were selected based on the chemical shift analysis **(Fig 4:3**) as they were expected not to perturb the structure. Residual dipolar coupling (RDCs) were used to better define the relative orientation of the domains. RDC measurements were obtained for the ^{1}H-^{15}N and Cα-C' dipolar coupling.

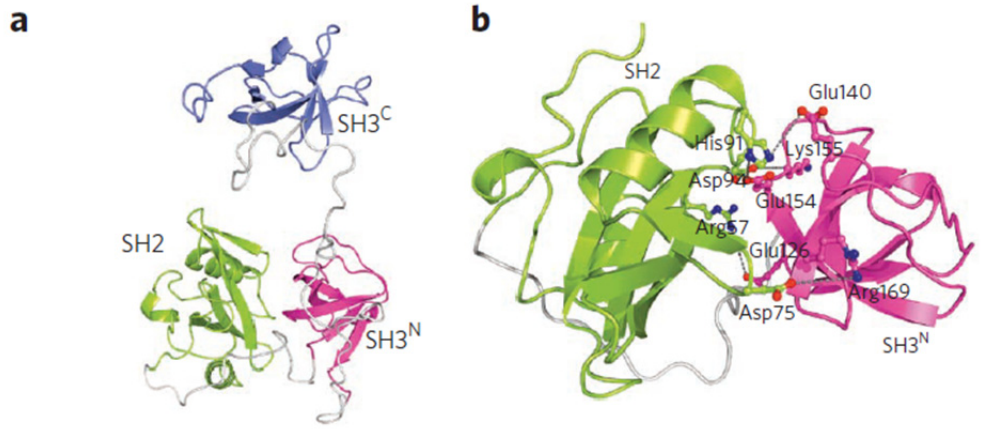


Fig 4:6 (a) Structure of CrkL. The SH2, SH3N and SH3C domains are colored green, magenta and blue, respectively. The linker regions are colored gray. The SH3C domain does not interact with the other domains. **(b)** Close-up view of the SH2-SH3N interface in CrkL. Only polar or charged residues mediate the interaction between the two domains**.**

The solution structure of CrkL reveals a rather unique structure. The SH2 domain and the $SH3^N$ domain interact directly with each other. The interaction is mediated by a surface of ~600$Å^2$. This is in good agreement with the chemical shift analysis. **Fig 4:3** Unlike the SH2-SH3 interface of CrkII, in CrkL SH2-SH3 interface is made up of polar residues.

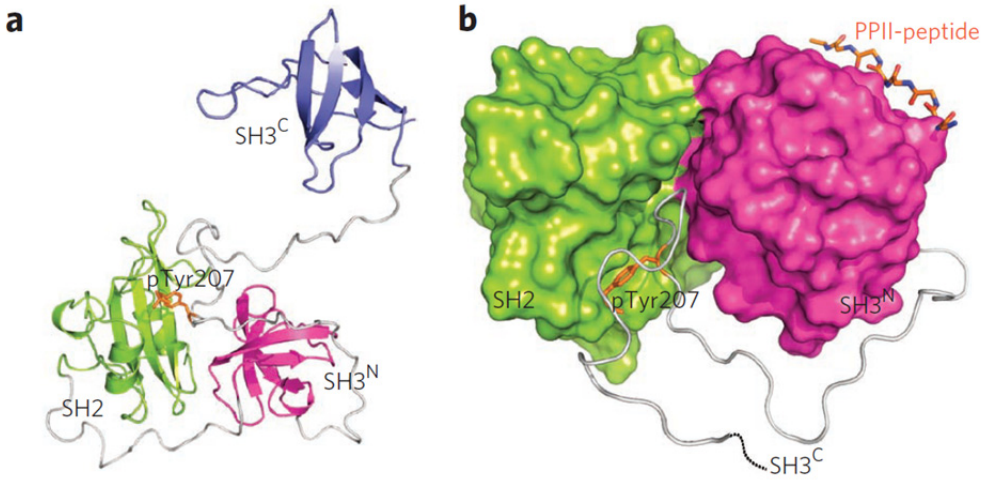


Fig 4:7 Structure of pCrkL. **(a)** Structure of pCrkL. pTyr207 is shown as orange sticks. **(b)** Close-up view of the pTyr-binding site. The SH2-$SH3^N$ interface undergoes slight adjustment to accommodate the binding of the linker to SH2.

Phosphorylated CrkL (pCrkL) was prepared by treating CrkL with catalytic amount of Abl Kinase. A majority of the assignment could be transferred for the pCrkL spectrum from the CrkL assignment. (Fig 4:5) A 3D-HNCA experiment was run to confirm the assignment. Using a similar strategy the structure of pCrkL was solved.

4.2.3 Dynamic properties of CrkL

Dynamic motions in proteins are important for biological function proteins. We used NMR relaxation methodologies to explore the dynamic properties of CrkL and pCrkL. ^{1}H-^{15}N NOE, longitudinal relaxation rate R_1 and traverse relaxation rate R_2 were measured for the full length protein.

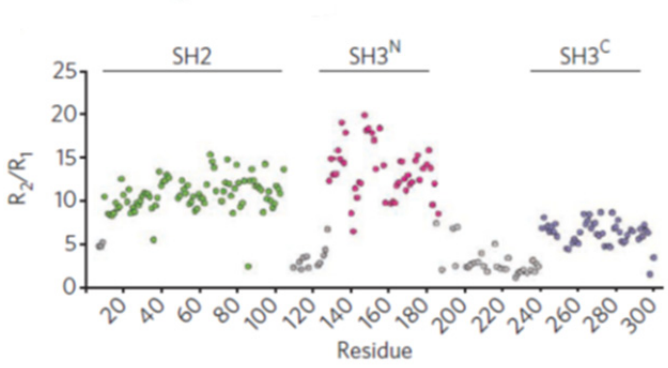


Fig 4:8 Plot of the R2/R1 ratio. 15N relaxation rates of the CrkL backbone as a function of residue number. The R2/R1 ratio provides information about the tumbling of the molecule, with higher values indicating slower tumbling

Both relaxation rates are sensitive to the overall motional properties of protein. The R_2/R_1 ratio provides a good estimation of the correlation time τ_c if residues that belong to the rigid part of

the protein are considered. [162] The data shows that the SH2 and the $SH3^N$ domain tumble with a τ_c of 11ns and the $SH3^C$ tumbles with a τ_c of 7.8ns.(**Fig 4:9**a) The $SH3^C$ domain tumbles much faster than the other two domains supporting the observation from our structural studies and chemical shift analysis that the SH2 and the $SH3^N$domain act as a rigid body and tumble as one unit. The organization of the domains in CrkL is different from that observed in CrkII. The SH2 domain of CrkII interacts the linker region (226-237) and the $SH3^C$ forming a repressed structure whereas in CrkL the $SH3^C$ domain is independent and forms no contact with the SH2 of $SH3^N$ domains. [85, 160].

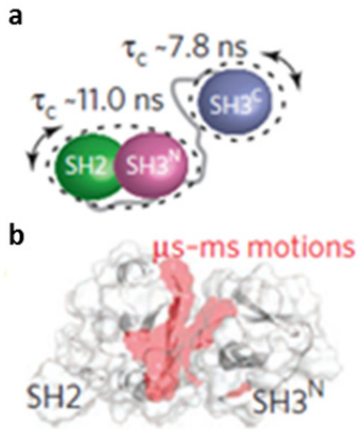


Fig 4:9 (a) Correlation times (τc) for the tumbling of CrkL. The SH2-SH3N module tumbles as a rigid unit, whereas the SH3C domain tumbles much faster and independently of the other domains. **(b)** Residues undergoing substantial μs–ms motions, as denoted by enhanced contribution to R_2 (*R*ex) values, are mapped on the structure of CrkL in red. Almost all residues located at the interface between the SH2 and SH3N domains show relatively high *R*ex values, indicating that the binding interface is dynamic.

4.2.4 $SH3^N$ domain in CrkL is not inhibited.

Structural analysis of the CrkL structure showed that the binding site for the PPII-peptide ligand (the $SH3^N$ domain) is completely exposed.(**Fig 4:19**) To test this we compared the binding affinity of a PPII-peptide ligand to isolated $SH3^N$ domain and full length CrkL (1-303). A peptide containing the consensus sequence for the $SH3^N$ domain that contains the first –PXXP- motif of Abl kinase was used in our experiments. ITC data showed that the $SH3^N$ domain has identical affinity for the ligand in both contexts. (**Fig 4:11**,a and b) Next we used NMR to monitor the effect of the PPII-peptide ligand binding to CrkL. The spectra revealed that the ligand binding causes minimal change to the rest of the protein and only the binding surface on $SH3^N$ is affected. **Fig 4:10**

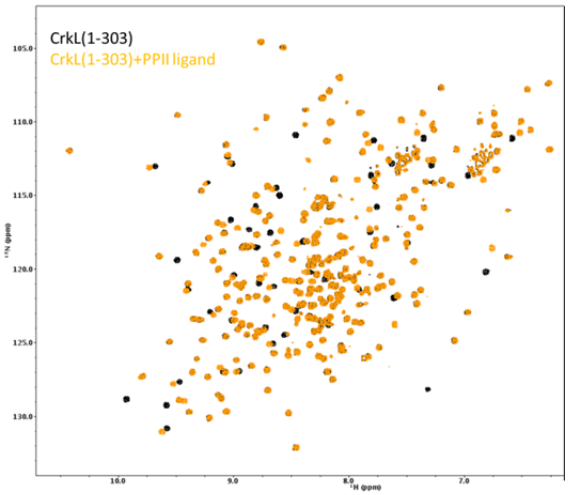


Fig 4:10 Spectra showing interaction of PPII-peptide ligand is limited to the $SH3^N$ domain only.

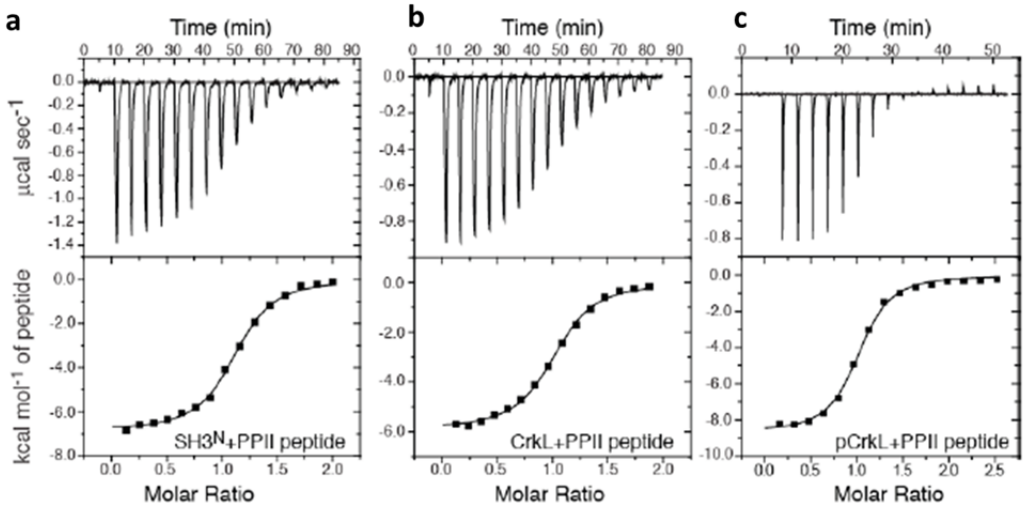


Fig 4:11 ITC traces and binding isotherms of titrations performed on isolated $SH3^N$ (a), full-length CrkL(b) and pCrkL(c).

Two independent studies have observed the inhibition of the $SH3^N$ domain in CrkII. The $SH3^C$ domain along with the linker connecting the two SH3 domains is able to adopt a conformation that blocks the access of PPII ligand. [85, 163]. In the full length CrkII structure, access to the PPII-peptide ligand binding site is blocked by a combination of the linker (226-237) and the $SH3^C$ domain. [160].

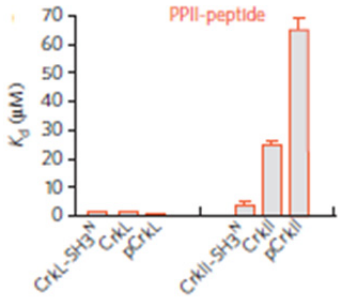


Fig 4:12 K_d values of PPII-peptide complexes with CrkL and CrkII fragments. Standard error was determined from three independent experiments.

Upon comparison, the $SH3^N$ of CrkL has a ~10 fold higher affinity for PPII-peptide ligand than CrkII. This means that a SH3 binding protein will preferably bind to CrkL than CrkII.(**Fig 4:12**)

4.2.5 CrkL inhibition via Y207 phosphorylation

The solution structure of phosphorylated full-length CrkII revealed that the adaptor protein was in the inhibited form.(**Fig 1:12**b) Phosphorylation at Tyr221 creates an internal binding site for the SH2 domain of CrkII and prevents CrkII from interacting with other proteins via the SH2 domain. The structure also shows that the domain reorganization blocks access to the $SH3^N$ binding site. CrkL also has an internal binding site for its SH2 domain that is created upon Tyr207 phosphorylation. The solution structure of pCrkL revealed a similar mode of domain reorganization upon Tyr207 phosphorylation. The pTyr207 motif binds the SH2 domain of CrkL in the canonical fashion where the negative charge of the phosphate group is stabilized by Arg21 and Arg39. To accommodate the interaction the SH2-SH3 unit of CrkL does not disengage, but simply reorient.**Fig 4:7**

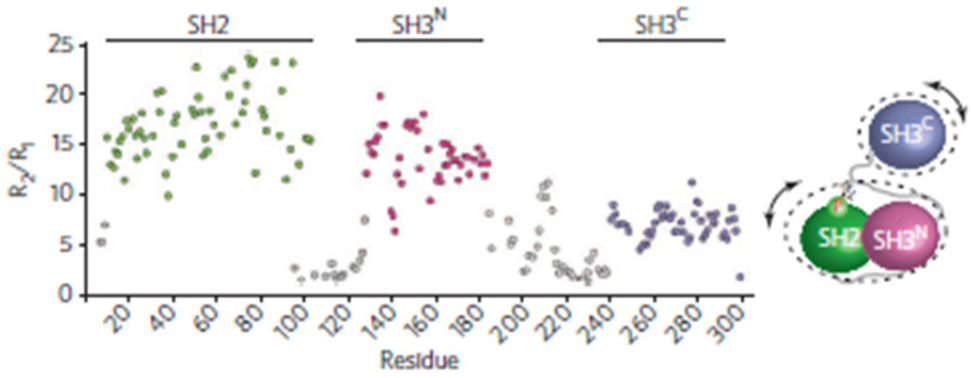


Fig 4:13 Plot of the R2/R1 ratio of pCrkL as a function of residue number. The scheme at right shows that the SH2-SH3N module in pCrkL tumbles as a unit, as in CrkL, whereas the SH3C domain tumbles much faster and independently of the other domains.

R_2/R_1 measurements show that the $SH3^C$ domain remains unperturbed. The phosphorylation status of CrkL appears to have no effect on the dynamic properties of $SH3^C$. **Fig 4:13**

How does an SH2 domain of CrkL already engaged in an interaction detach from the target upon Tyr207 phosphorylation to bind intra-molecularly? To investigate this we devised a simple experiment. We produced ^{15}N labeled linker-CrkL(188-304) which contains pTyr207 and added full-length unlabeled CrkL. ^{1}H-^{15}N HSQC spectra showed chemical shifts consistent with the unlabeled CrkL interacting with pTyr207 of the linker as a complex is formed between the SH2 domain of CrkL and pTyr207 of the linker-CrkL. Upon addition of catalytic amounts of Abl KD along with ATP and $MgCl_2$ we observe the full-length CrkL detaches from the complex. The only way for this to occur is for the SH2 domain of full-length CrkL to preferentially form a complex intra-molecularly with its newly formed binding site at its own pTyr207.(**Fig 4:14 a, b**) This aspect was further tested in cells where we find that phosphorylated CrkL is less likely to form a complex with paxillin via the SH2 domain.**Fig 4:14**c The data suggests that phosphorylated CrkL is inhibited to the extent that it can no longer interact with other proteins *via* the SH2 domain.

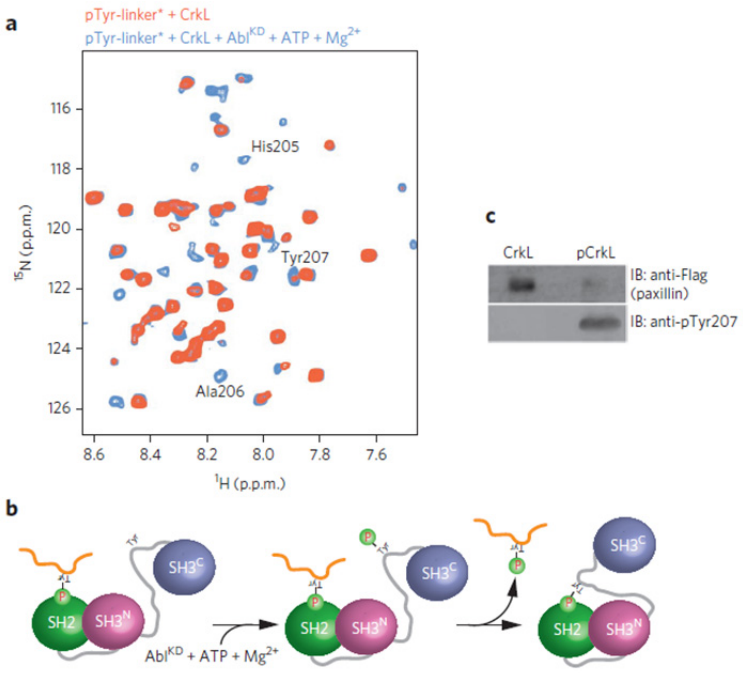


Fig 4:14 Effect of Tyr207 phosphorylation on CrkL folding and its association with Abl kinase. **(a)** ^{1}H-^{15}N HSQC NMR spectra of the linker region of CrkL containing the phosphorylated Tyr207 (pTyr-linker) in the presence of CrkL(orange) and after the addition of catalytic amounts of AblKD and ATP-Mg2+ (blue). The pTyr-linker is 15N-labeled, whereas CrkL and AblKD are unlabeled. Asterisk denotes isotopic labeling. (b) Analysis of the NMR experiments in a shows that the pTyr-linker binds the SH2 domain of CrkL. Phosphorylation of Tyr207 in CrkLinduces the intramolecular association of pTyr207 and SH2. As a result, the pTyr-linker is displaced. (c) Pulldown of CrkLand pCrkLwith paxillin, an SH2-binding physiological partner of CrkL

4.2.6 Unlike CrkII, the SH3^{N} domain of pCrkL is not inhibited

Structural analysis of the pCrkL structure revealed that the binding surface of the SH3^{N} domain is exposed. This implies that CrkL should be able to interact with PPII ligands both in the

phosphorylated and unphosphorylated form. This is different from CrkII where upon phosphoryalation at Tyr221, CrkII is essentially inactive as the SH2 is engaged intra-molecularly, binding the internal site and the domain rearrangements block access to the $SH3^N$ PPII binding site. ITC experiments show that pCrkL has similar binding affinity to PPII-peptide ligand as full-length CrkL and $SH3^N$. **Fig 4:11**, **Fig 4:12**. A pull-down experiment with DOCK180 which binds $SH3^N$ of CrkL showed comparable association with both CrkL and pCrkL.

Fig 4:15 Pull-down experiment of CrkL and pCrkL with DOCK1, an $SH3^N$ binding partner of CrkL. Membrane was blotted for DOCK1 and pTyr207.

The notion that pCrkL can stay bound to a ligand *via* the $SH3^N$ domain was tested *in vitro.* Abl kinase contains three –P-X-X-P motifs that are known to bind the $SH3^N$ of CrkII and CrkL. [71]. We prepared a complex of unlabeled Abl^{PxxP} (Abl kinase domain and the first –P-X-X-P- motif) and ^{15}N labeled CrkL. NMR experiments show that only the $SH3^N$ domain of CrkL is engaged in the interaction with the Abl^{PxxP} construct. Upon addition of ATP and $MgCl_2$, NMR experiments show the expected rearrangement of the SH2 domain of CrkL interacts intra-molecularly. Interestingly, Abl^{PxxP} is still bound to the $SH3^N$ domain. (**Fig 4:16**a, **Fig 4:21**)

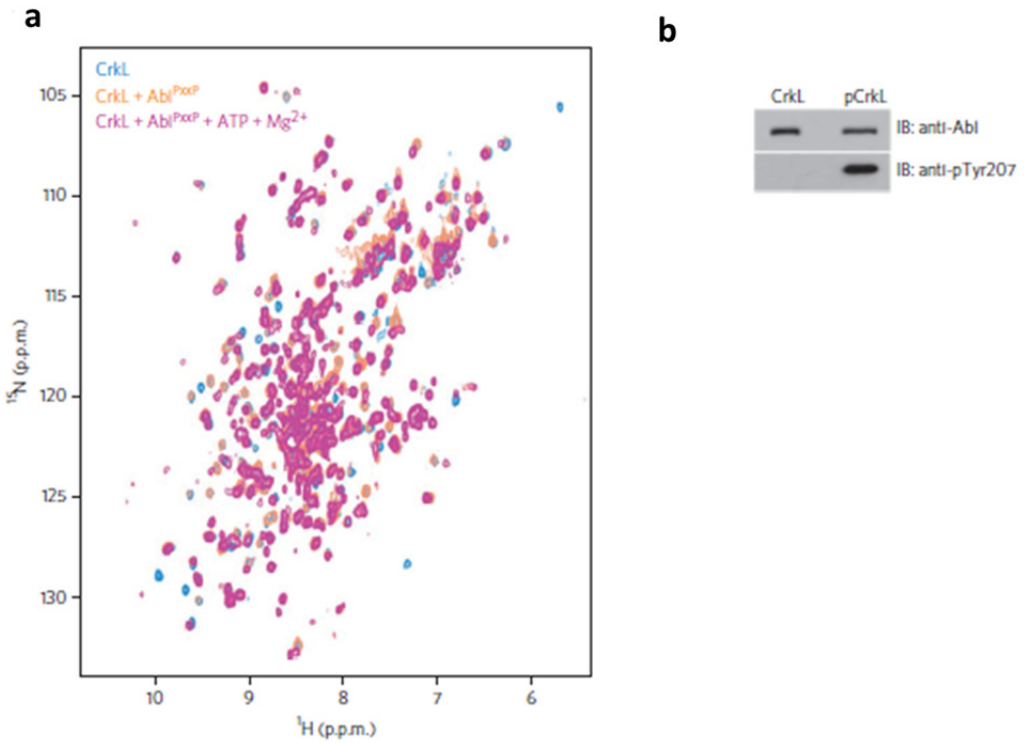


Fig 4:16 (a) 1H-15N HSQC NMR spectra of free CrkL(blue), in complex with AblPxxP (orange) and after adding ATP+Mg2+ (magenta). AblPxxP is a construct that encompasses the kinase domain and the first PxxP motif that binds CrkL. **(b)** Pull down of CrkL and pCrkL with full-length Abl (form 1b) Membrane blotted for Abl and pTyr207.

Pull down experiments with CrkL and pCrkL show that Abl kinase can associate with both forms of the CrkL protein.

4.2.7 Discussion

The high sequence similarity of CrkII and CrkL lead to a presumption that both proteins behave in a similar manner **Fig 4:17**a. Both proteins display similar propensity to bind to other cellular proteins making identification of binding partners that are exclusively for CrkL or CrkII difficult. The solution structure of CrkII and pCrkII highlighted the key features of the adaptor protein and helped explain the observations made in cellular and biochemical studies. CrkII in solution exists as a monomer and exists in a closed conformation where the SH2 domain can access pTyr motifs. The $SH3^N$ domain is not accessible because it is inhibited by a combination of structural restraints including residues from the DE loop of the SH2, the linker region (226-237) and the $SH3^C$ domain. pCrkII is also monomeric and the SH2 domain binds intra-molecularly to Tyr221. This also blocks access to $SH3^N$. As an adaptor protein, CrkII is thus able to bind with pTyr motifs only in the un-phosphorylated form. The $SH3^N$ domain is intrinsically inhibited from interactions in both forms of the protein.

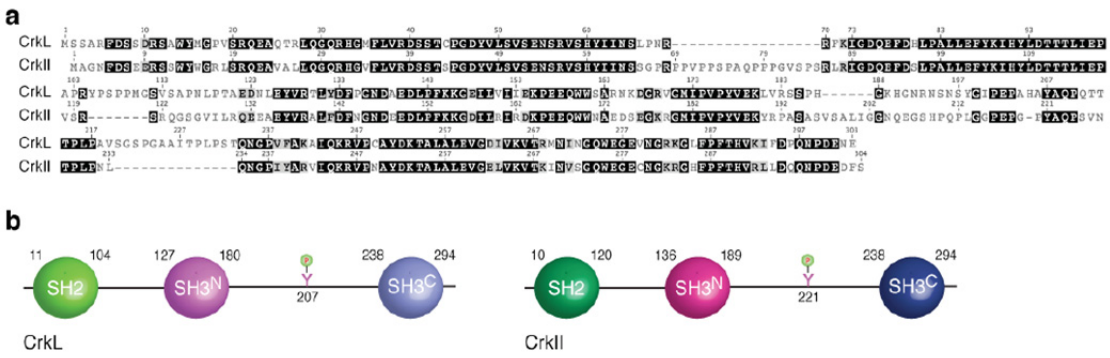


Fig 4:17 Sequence and domain organization of CrkL and CrkII. (a) Sequence alignment of human CrkII and human CrkL. (b) Domain organization of CrkII and CrkL. Tyr207 in CrkL and Tyr221 in CrkII are phosphorylated by Abl kinase.

CrkL, despite the high sequence similarity displays structural properties that are different from that of CrkII. The $SH3^C$ domains of CrkII and CrkL have no known binding partner. The solution structures of domains show that although the $SH3^C$ domain adopts the canonical fold associated with SH3 domains, it is missing a critical aromatic residue at the binding surface and therefore unable to interact with –P-X-X-P motifs. [83, 160]. The $SH3^C$ domain, however can play a role in inhibiting ligand binding to the $SH3^N$ domain. [85]. The $SH3^C$ domain of CrkL was reported to be a dimer and assist in the dimerization of full length CrkL. [89]. This dimerization would require partial or complete unfolding of the $SH3^C$ domain and expose a NES site for Crm1 binding. The $SH3^C$ had been reported to interact with CD34, cell surface transmembrane protein.[164]. Harkiolaki *et.al.* failed to detect binding of the putative binding site on CD34 with $SH3^C$ using ITC and our efforts to detect interaction with NMR also yielded no results. A recent study identified the $SH3^C$ domain of CrkL interacting with the PSI domain of β_1 integrin. Initially identified from a study assaying for tumor homing peptides, the study uncovered a unique interaction between the $SH3^C$ of CrkL and a peptide sequence (YRCTLNSPFFWEDMTHECHA). [165]. The peptide was identified to be part of the extracellular domain however, in our hands neither the peptide nor the purified PSI domain showed any interaction with the $SH3^C$ or full-length CrkL. Considering the structural, biochemical and cellular studies we conclude the $SH3^C$ domain of CrkL has no known function-either as a binding partner or a regulatory role.

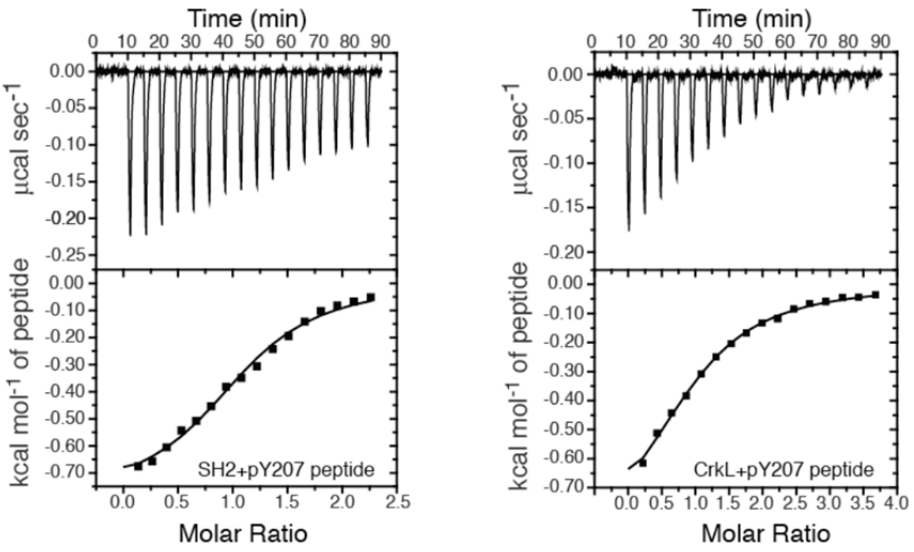


Fig 4:18 ITC traces and binding isotherms of titrations performed on isolated SH2 domain and full-length CrkL with phosphopeptide CrkL(pTyr207). pTyr207 binds isolated SH2 with a K_d ~7µM and full-length SH2 with a K_d~23 µM.

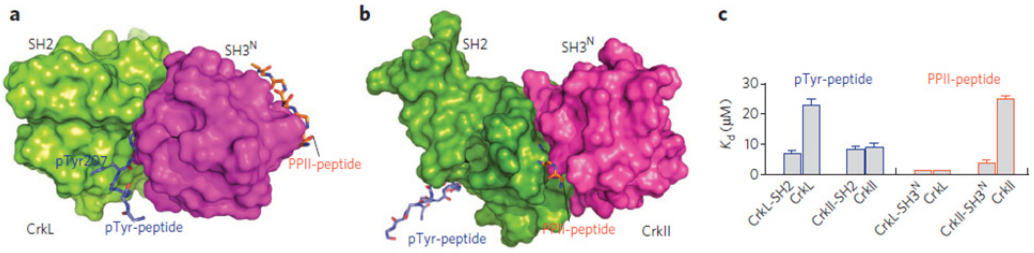


Fig 4:19 Binding of pTyr peptide and PPII-ligand peptide to CrkL and CrkII. (a,b) The Structure of SH2-$SH3^N$ domains in CrkL(a) and CrkII(b)The pTyr peptide binding site in CrkL is partially masked but is completely accessible in CrkII. The PPII binding site in CrkL is exposed whereas in CrkL it is masked. (c) Summary of the Dissociation constants (K_d) of pTyr-peptide and PPII ligand with CrkL and CrkII.

Our structural studies of full-length CrkL also revealed several key differences between CrkII and CrkL. In the unphosphorylated form, the SH2 domain of CrkL is inhibited **Fig 4:19c.** Isolated SH2

domain of CrkL has a ~3 fold higher affinity for pTyr ligands than full length CrkL (**Fig 4:18**). The SH2 domain of CrkII can interact with ligands unimpeded **Fig 4:19**b. SH2 domain of CrkII, isolated or in context of the full-length protein interact with pTyr ligand with equal affinity. Moreover, isolated SH2 domains of CrkL and CrkII have similar affinity for pTyr ligands.(**Fig 4:20**a) This difference becomes relevant when we consider that CrkL and CrkII are both present in the cytoplasm and even though a pTyr ligand may have similar affinity for the SH2 domains of CrkII and CrkL, the ligand will preferably bind to CrkII simply because the domain organization of CrkL results in the ligand binding site at its SH2 domain to be buried **Fig 4:19**a. p130CAS, an important ligand for the Crk family of adaptors contains multiple pTyr sites and is able to simultaneously bind several SH2 domain containing proteins. [55]. Our studies indicate that CrkII would more likely form complexes with p130CAS than CrkL.

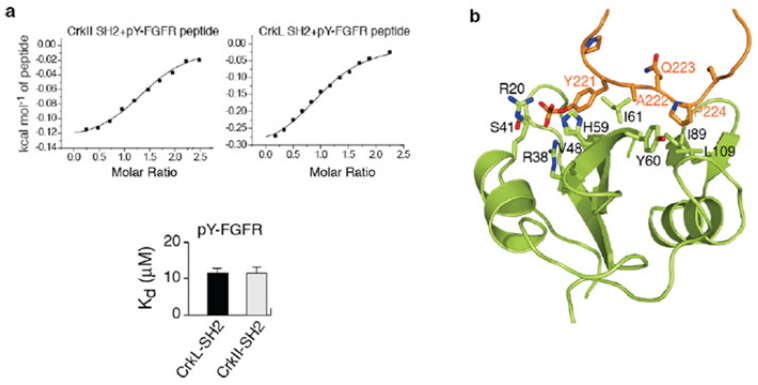


Fig 4:20 (a) Binding isotherm of isolated SH2 domain of CrkII and CrkL with pTyr peptide from FGFR peptide. Isolated SH2 domain of CrkII and CrkL have similar binding affinity. (b) The peptide binding mechanism is identical for both SH2 domains.

In terms of interactions with the $SH3^N$ domain, CrkL is more accessible than CrkII both in the phosphorylated and unphosphorylated forms.(**Fig 4:19**c). Proteins that interact with the $SH3^N$ domain of Crk proteins will more likely interact with CrkL than CrkII. In this scenario, a

CrkL/p130CAS complex is more likely to bind to DOCK180 (or C3G) than a CrkII/p130CAS complex as the PPII-binding site of $SH3^N$ in CrkII remains inhibited. The association of DOCK180 and CrkL increase cell migration and invasion and may explain why CrkL is more transforming than CrkII in fibroblasts [166]. This also explains why CrkL is the preferred substrate for Bcr-Abl as shown in **Fig 4:16** and Fig **4:21**. Abl kinase can remain bound to CrkL even after phosphorylation thereby perpetuation signaling at the complex. The Abl-CrkII complex dissociates upon CrkII Tyr221 phosphorylation.

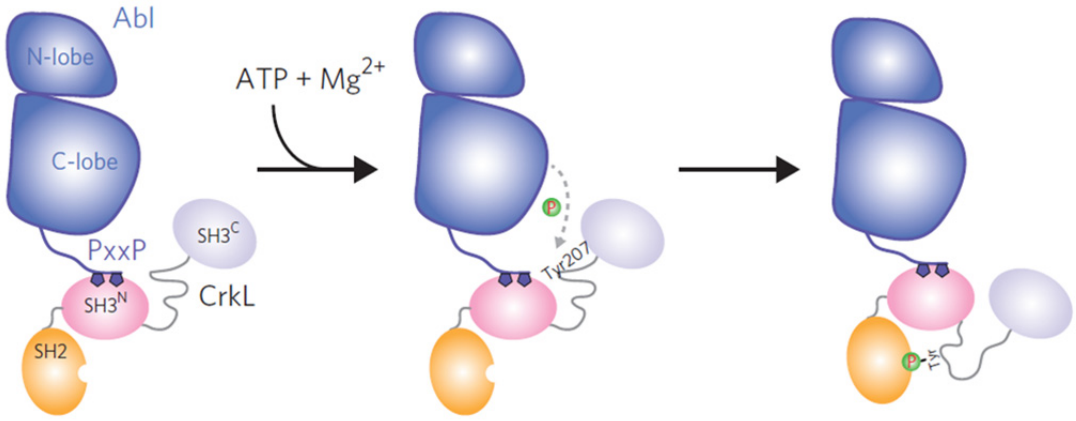


Fig 4:21 Abl Kinase docks on the $SH3^N$ domain of CrkL and phosphorylates Tyr207. Intramolecular binding of SH2 domain to pTyr207 does not disrupt Abl-CrkL complex.

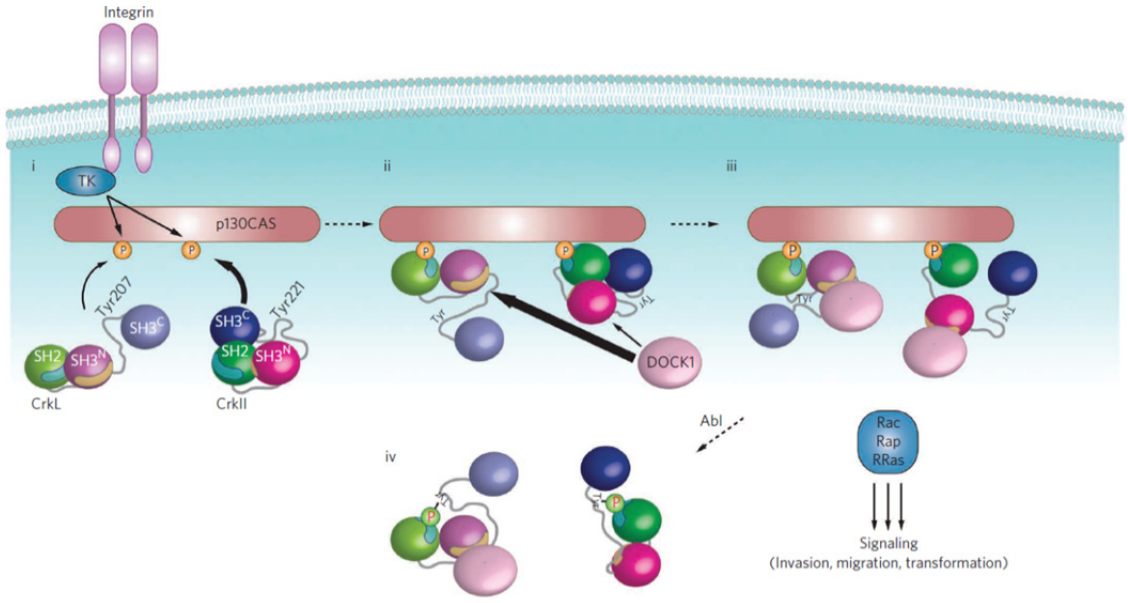


Fig 4:22 CrkL versus CrkII signaling. Integrin activation elicits p130CAS phosphorylation by tyrosine kinases (TK), and, as a result, CrkLand CrkII are recruited. (ii) GEFs (for example, DOCK1 and C3G) associate with CrkLand CrkII via their SH3N domain, giving rise to efficient localized activation (iii) of small GTPases (for example, Rac, Rap and RRas) at the membrane. (iv) Abl-induced phosphorylation of CrkLand CrkII forces their dissociation from p130CAS and thus results in signaling suppression. The distinct structural organization of CrkLand CrkII modulates the interactions with their physiological partners to a different extent. The blue and brown shaded regions in SH2 and SH3N denote the pTyr- and PPII-binding sites, respectively

The SH2 of CrkII contains an extended DE loop (**Fig 4:4**) that can bind Abl SH3 domain. The SH2 domain of CrkL does not contain these extra residues. Although there is no study that links changes in Abl kinase activity solely on the interaction of the DE loop, it is possible that in the CrkII-Abl complex, this interaction stabilizes the complex even after Tyr221 phosphorylation and subsequent $SH3^N$ domain inhibition. The structure of the CrkII SH2 domain and Abl SH3 domain alludes to this possibility. {Donaldson, 2002 #5328} A recent study has identified Tyr251 of CrkII as a site phosphorylated by Abl kinase and EGFR. This phosphorylation creates a binding site for

Abl SH2 binding and in the process activates Abl. In addition, Crk SH2 can also bind this motif as the accompanying sequence includes –^{248}PNAY^{251}- . CrkL contains a similar motif -^{248}PCAY^{251}-. It will be interesting to investigate what the structural outcome is for these modifications and whether these phosphorylations occur independently or in a proccessive manner.

5 CypA acts as a switch for Abl phosphorylation

5.1.1 Introduction

Crk adaptor proteins with their modular nature have necessitated an assumption that all protein interactions with these molecules occur through the binding surfaces of the SH2 and SH3 domains. The existence of *cis-trans* isomerization in the linker region of CrkII presents the possibility of these proteins interacting with a new class of proteins through the linker [85]. Prolyl isomerization is an intrinsically slow process and a number of enzymes have been identified that catalyze this process. This long linker (~50 aa) is conserved and other than the inhibitory Tyr221 (Tyr207 in CrkL), no significant cellular function has been attributed to this region. Cyclophilin A, a prolyl *cis-trans* isomerase (PPIase) can interact with the heterogeneous Gly237-Pro238 bond of CrkII and catalyze the slow isomerization process.{Sarkar, 2011 #11114}

5.1.2 PPIase family

Cyclophilin A belongs to a family of PPIases that catalyze the slow interconversion between the *cis* and trans prolyl conformers. CypA, a member of the prolyl-cis/trans isomerase family (PPIase family) of proteins is involved in numerous cellular processes is the most abundant and widely expressed protein in cells. The PPIase family includes Cyclophilins, FK-506 binding proteins (FKBPs) and parvulins . CypA catalyzes the *cis-trans* isomerization of the peptidyl-prolyl bonds in proteins and therefore the primary function associated with CypA has been as a molecular chaperone[167]. The enzymatic function of CypA also plays a role in intracellular trafficking, signal transduction and regulation of transcription. Analogous to phosphorylation and

acetylation, cis-trans isomerization has been purported to be an additional level of regulation available for proteins.{Andreotti, 2003 #9240}

CypA was initially identified as a target of the immunosuppressive drug Cyclosporin A (CsA) and the subsequent complex of CypA-CsA was shown to inhibit calcineurin. FK-560 was also shown to interact with CsA. Although CsA binds to the catalytic center of the isomerases, they yield their immunosuppressant activity by forming a complex and inhibiting the phosphatase activity of calcineurin. [168-170] The role of these PPIases in cells is implied from the catalytic activity – to accelerate protein folding.

5.1.3 Cyclophilins (Cyps) family

Cyclophilin A(CypA) was the first identified member of the of the family of cyclophilins. Cyps are evolutionarily conserved and are present in prokaryotes and eukaryotes. There are 16 cyclophilins in humans out of which the major ones include, Cyp(A-E) and Cyp40. They differ in structure, size and cellular distribution. CypA is mainly cytosolic while CypB and CypC is associated with the ER. CypD localizes in the mitrochondia and CypE is a nuclear protein. The size of Cyps vary from 18kDa (CypA) up to 40kDa (Cyp40). [171]

5.1.4 CypA

CypA is the mostly widely studied member of the PPIase family and folds as a β-barrel with eight β-beta strands connected by loops and three α-helices. The active site of the protein is hydrophobic in nature and there is very little structure change during the catalytic process. However extensive interactions are observed between the residues of the active site of CypA and the binding protein.(**Fig 5:1**a) [172]

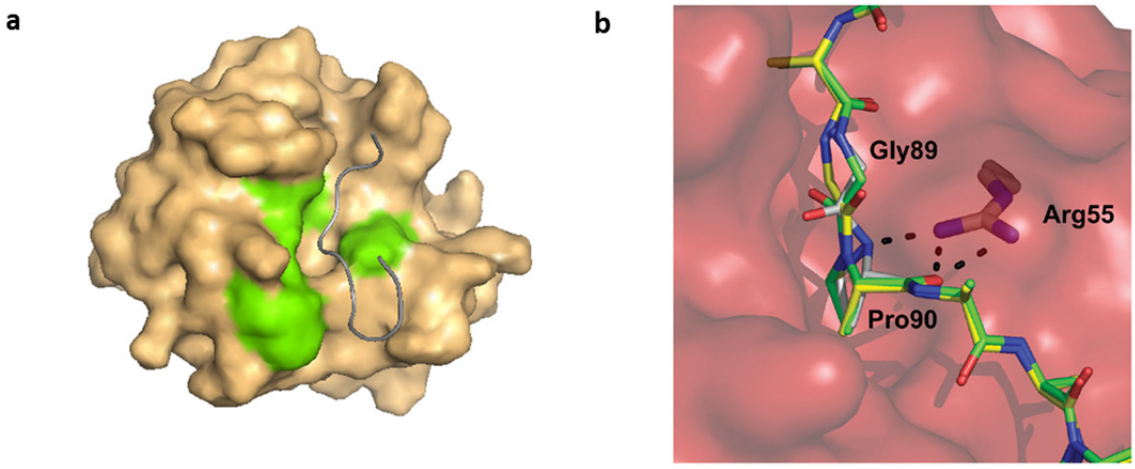


Fig 5:1 (a) Surface representation of CypA with ligand. The most effected residues are shown in green. **(b)**A zoomed view of the active-site of CypA bound to ligand. Arg55 forms the essential contacts with the isomeric Proline.

Residues that are involved in the interactions with the substrate include Arg55, Ile57, Phe60, Met61, Gln63, Ala101, Asn102, Gln111, Phe112, Trp 121, Leu122 and His126. [173] Arg55 is directly involved in the catalytic process and a substitution of Arg55 to an Ala substantially reduces the catalytic efficiency of CypA. Several crystal and solution structures of CypA and putative substrates have revealed that CypA is capable of binding proline in both the imide bond conformations. [174] In the crystal structures of HIV-1 CA^N peptide with CypA, the heterogeneous proline (Pro90) is found in both conformations with little change to the overall structure. The residues following Pro90 are identical in both forms but there are some differences for the residues N-terminal to Pro90. Residues 88 and 89 of the HIV-1CA^N display changes in the backbone conformation.

5.1.5 Cyclophilin and disease

Cyclophilins are involved in various pathological conditions including HIV, hepatitis B and C viral infections,[175, 176] and various cardiovascular diseases including atherosclerosis,[177, 178] While these aspects have been studied in some detail, recently CypA has been shown to be overexpressed in certain forms of cancers including small cell lung cancer[179-181] pancreatic,[182-184] breast [185, 186]and colorectal cancers[187-189]. Tumor cells show elevated levels of CypA mRNA. {Obchoei, 2009 #9196} Proteomic studies on cells treated with anticancer drugs revealed a decrease in levels of CypA.[187, 189]

Table 1. CypA mRNA expression in different tissue types.

Tissue type (total libraries)	Tag counts per 200,000 (Mean ±SEM)	
	Normal	Tumor
Brain (77)	181±28	270±14
Breast (49)	229±16	319±22
Colon (4)	87±23	234±2
Kidney (3)	195±98	540
Liver (2)	156	943
Lymph node (4)	316±156	612±51
Muscle (3)	23±5	332
Pancreas (3)	92	140±4
Peritoneum (2)	134	258
Retina (10)	188±34	452
Skin (18)	236±26	423±121
Stomach (11)	181±61	353±25
Thyroid (2)	351	401
Prostate (5)	352±110	410±69
Lung (4)	426	364±60
Bone marrow (6)	314±137	85±38

Fig 5:2 Expression profile comparing CypA expression levels in tumor tissue and the normal counterpart. Profile created from SAGE map. http://www.ncbi.nlm.nih. {Obchoei, 2009 #9196}

Although the mechanistic basis of the CypA's involvement in each of these cancer types is not clear recent studies provide some intriguing observations. CD147, a cell surface receptor

can interact with CypA using a proline-containing peptide in its transmembrane domain and stimulate cell proliferation in human pancreatic cancer cells. [182] Knockdown of CypA in a osteosarcoma cell line, U2OS decreases cell migration by altering the F-actin structure. {Calhoun, 2009 #13190} CsA treatment of ER-MDA-231 breast cancer cells inhibit cell motility and cell invasiveness in a dose dependent manner. {Zheng, 2008 #13191}. Hela cells display increased resistance to hypoxia and cisplatin-induced cell death following overexpression of CypA[190]. Taken together these studies points to an important role of CypA in carcinogenesis.

5.2 Results

5.2.1 Proline induced conformational heterogeneity in CrkII

Fig 5:3 Sequence alignment of CrkII(216-231) shows absolute conservation of amino acids around Tyr221.

The conservation of a Proline residue around the regulatory phosphorylation site on CrkII (Y221) is intriguing as it suggests the possibility of conformational heterogeneity and an additional level of regulation for CrkII. NMR analysis of the region flanking Pro220 (Glu217-Qln223) shows a duplication of peaks indicating the presence of two conformations in slow exchange (**Fig 5:4**,b).

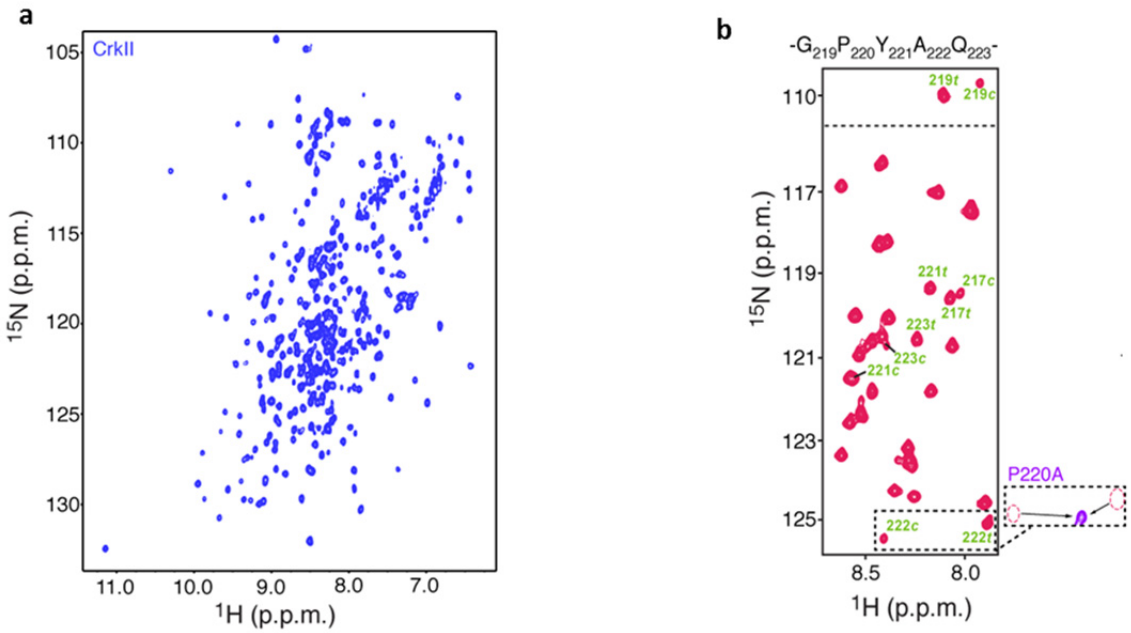


Fig 5:4. **a**. ^{1}H-15N HSQC of CrkII (1-304) **b.** ^{1}H-15N HSQC of hl-SH3^{c}-P238A.

Characteristic NOE correlation between Gly219-Pro220 as well as $^{13}C_{\beta}$ and $^{13}C_{\gamma}$ chemical shift analysis of Pro220 indicate that the Gly219-Pro220 bond undergoes cis-trans isomerization (**Fig 5:5**a). A Pro220Ala mutation of the site restores the homogeneity of the region validating that the G-P bond is the source of the conformational heterogeneity (**Fig 5:5**b). Because peptidyl-prolyl cis-trans isomerization is slow with respect to the NMR chemical shift scale [192]two resonance frequencies are observed for each nuclear spin affected by the isomerization process. We conclude that the source of the conformational heterogeneity in the phosphorylation site of CrkII is the presence of cis-trans isomerization about the Gly219-Pro220 prolyl bond.

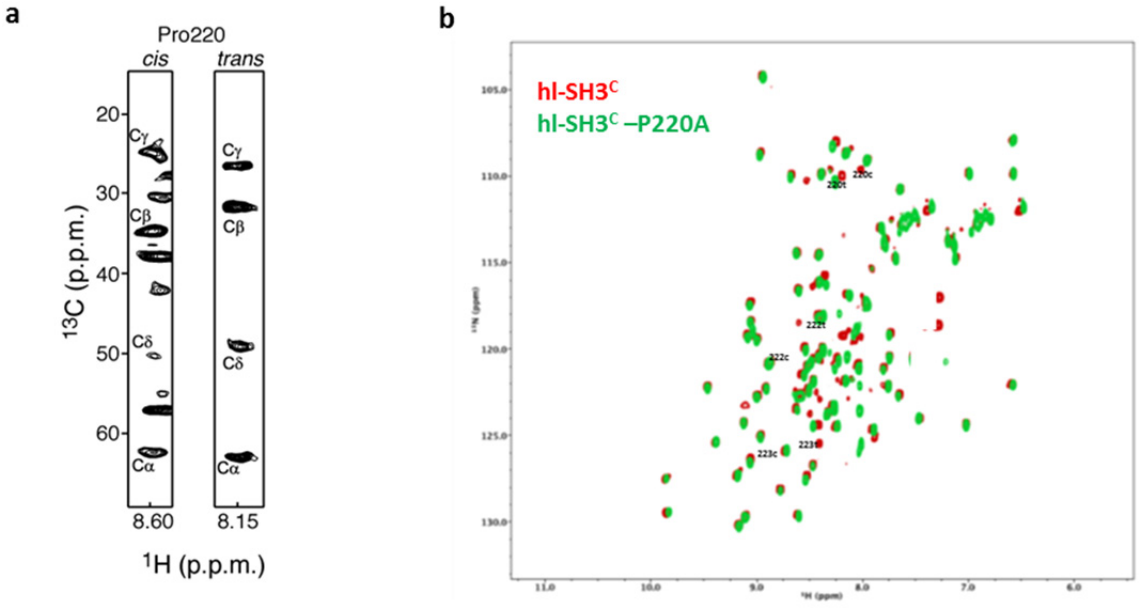
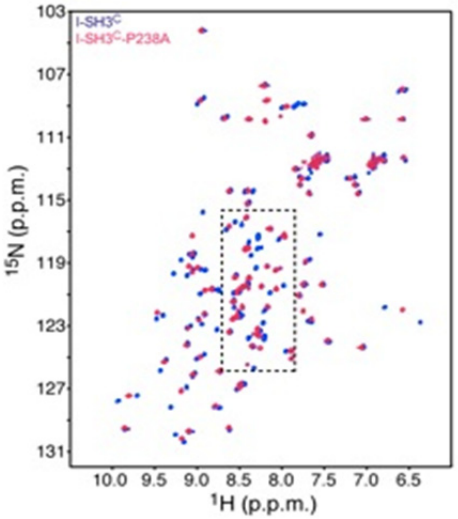
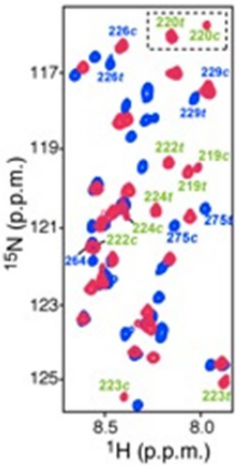

a
Pro220
cis
trans
20
30
40
50
60
13C (p.p.m.)
Cγ
Cβ
Cδ
Cα
8.60
8.15
1H (p.p.m.)
b
hl-SH3C
hl-SH3C –P220A

Fig 5:5 (a) Strips from CCO-NH spectra showing the ^{13}C chemical shift of Pro220 carbon skeleton. The chemical shift difference between the $^{13}C_{\beta}$ and $^{13}C_{\gamma}$ nuclei for the two conformations of residue Pro220 are 9.2ppm and 4.6ppm which, on the basis statistical analysis of ^{13}C chemical shifts[191] of proline residues



in proteins corroborate that the two forms are in the *cis* and *trans* conformations.(b) Overlay of ^{1}H-^{15}N



HSQC of hl-SH3^{C} (red) and hl-SH3^{C} -Pro220 (green)

Fig 5:6 ^{1}H-15N HSQC of l-SH3^{C} P238A CrkII. Zoom (right)

Volume integration of isolated cross-peaks corresponding to the two conformations indicates that the trans:cis population ratio is 3:1 at room temperature. The effect of Pro220 cis-trans

isomerization is local, with only 9 residues showing duplicate peaks. In fact the full length spectra of CrkII shows the expected number of peaks indicating that the intact protein is essentially homogeneous(**Fig 5:4 A**).[160, 163] It should be noted that the isomerization process at Pro220 is distinct from, and independent of, the isomerization process at Pro238 previously identified in CrkII from chicken[85, 163]. (Substitution of Pro238 by Ala abolishes the Pro238 isomerization but has no effect on the Pro220 isomerization (**Fig 5:6**). In contrast to the isomerization process at Pro220, which is evolutionary conserved among CrkII proteins, Pro238 isomerization appears to be species specific and while is present in chicken it appears not to be present in human CrkII (**Fig 5:6**). In addition, a synthetic peptide encompassing the Pro216-Ser225 region shows *cis-trans* isomerization strongly supporting the notion that this process is intrinsic to the sequence (**Fig 5:7**). Given that this sequence is highly conserved around Pro220, it provides strong evidence that cis-trans isomerization about Pro220 is conserved in mammalian CrkII.(**Fig 5:3**)

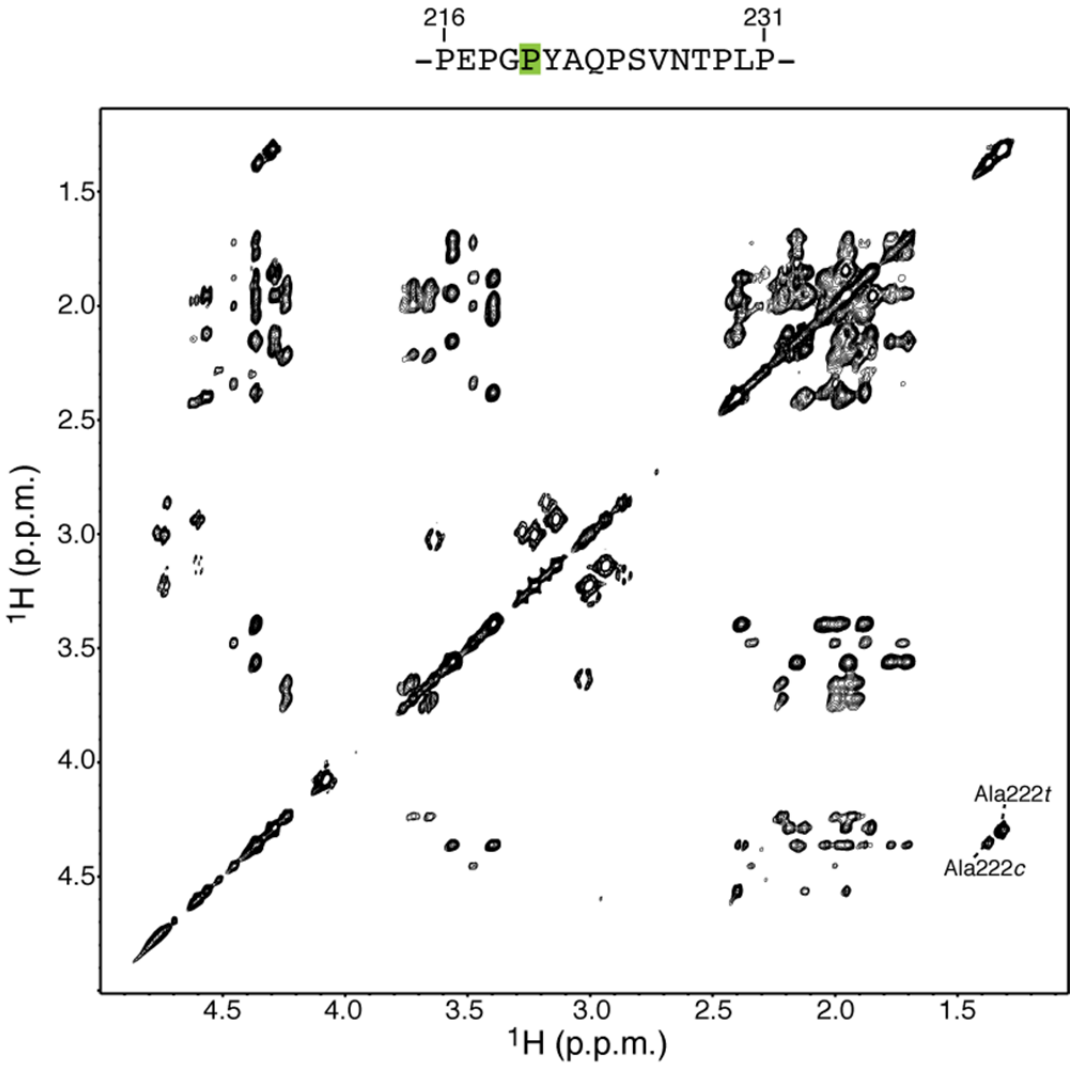


Fig 5:7 ^{1}H-^{1}H NOESY experiment of the synthetic peptide displays *cis-trans* isomerization about Pro220. The two sets of peaks, one for each conformation is easily identified.

5.2.2 CypA binds to CrkII around G219-P220 motif

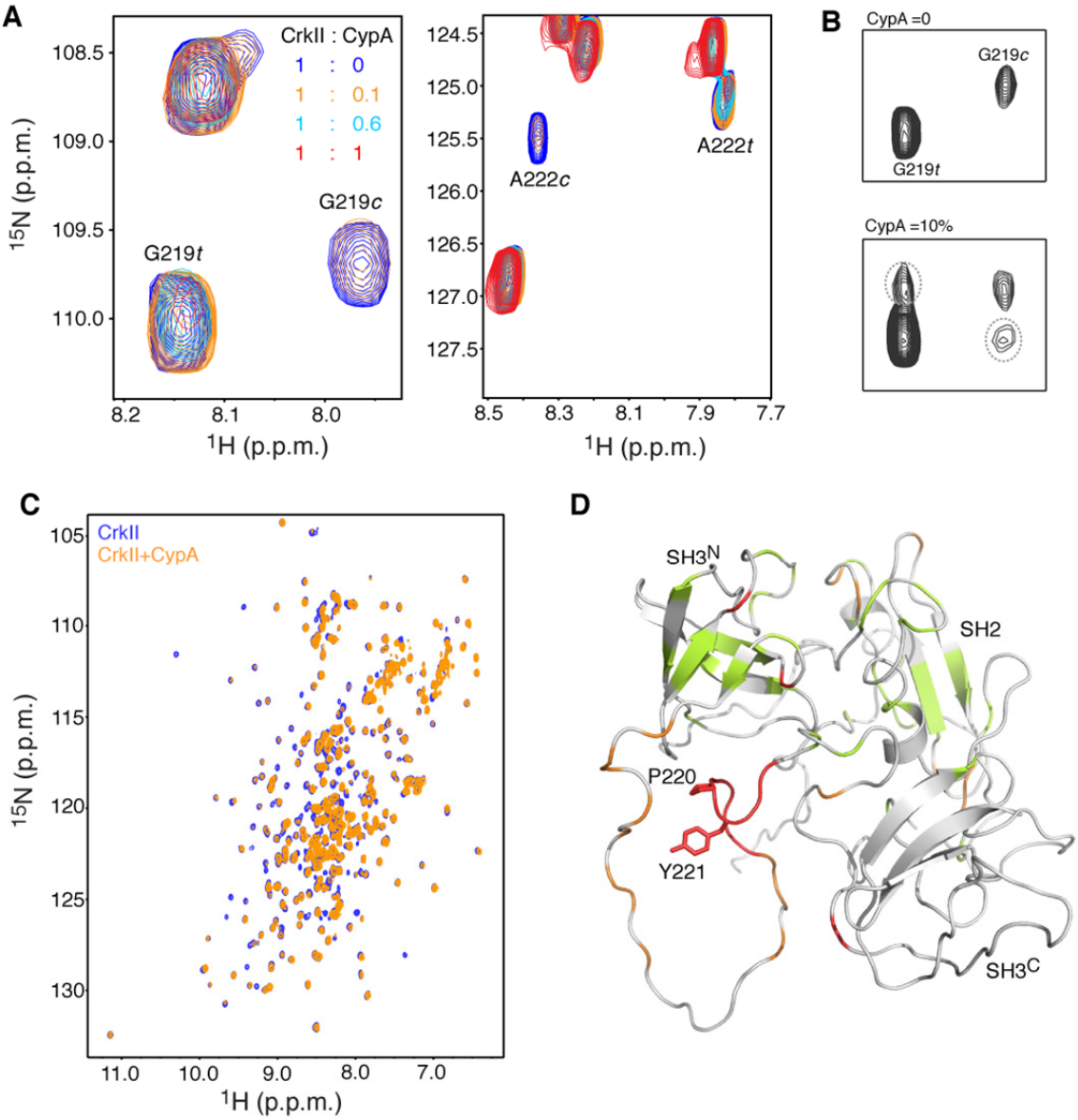


Fig 5:8 (A) A titration experiment where unlabeled CypA was added to ^{15}N-labeled CrkII. Binding of CypA leads to peak broadening for the residues near the binding site. **(B)** ZZ-Exchange spectroscopy. Catalytic amounts of CypA lead to the appearance of exchange peaks shown in dotted lines. Exchange peaks appear when the rate is quite fast. **(C)** CypA interaction with CrkII. Peak broadening observed for residues that come in contact with CypA. (D) Chemical shift mapping of residues perturbed upon CypA binding. Scale Red –most peak broadening, green –moderate, orange-slight, white-unaffected.

The presence of heterogeneous proline residue undergoing *cis-trans* isomerization provides a potential binding site for PPIases. In fact the Gly219-Pro220 motif presents a favorable binding

site for CypA. [193]We used NMR spectroscopy to probe the interaction of CypA and CrkII. NMR spectroscopy is one of the few techniques that is able to detect motion at the slow time scale of proline isomerization(k_{ex}~$0.01s^{-1}$). Unlabeled CypA was titrated to ^{15}N-labeled CrkII. The protein backbone signals are sensitive to local changes in the environment and binding events are indicated as shift in the peaks or a broadening of the signals in a NH cross correlation spectroscopy. [145] Catalytic amounts of CypA resulted in significant line broadening of residues near Pro220 (**Fig 5:8A,C**).

^{1}H-^{15}N NMR exchange spectroscopy shows that the addition of catalytic amounts of CypA results in the appearance of exchange peaks in the 2D exchange spectrum, providing direct evidence that CypA catalyses the Gly219-Pro220 cis-trans isomerization process[194].**Fig 5:8**B Increasing molar amount of CypA leads to specific chemical shift perturbations effecting residues mostly near the binding site. (**Fig 5:8C,D**) The reverse titration of unlabeled CrkII to ^{15}N-labeled CypA further corroborates the specific interaction between the two proteins.(**Fig 5:9**A) Chemical shift mapping reveals the binding of CrkII to be localized to the catalytic site of CypA. In fact the residues that are perturbed by the binding are identical to interaction sites of known CypA-ligand structures. Addition of Cyclosporin A, an inhibitor that binds the catalytic site of CypA with high affinity is able to disrupt the CrkII-CypA interaction.(**Fig.5:10**) **Error! Reference source not found.** [168, 195] NMR analysis of the binding revealed a dissociation constant (K_d) of 18uM, which is lower than typical CypA-ligand interactions indicating that the binding is stronger in this case. [193] Pro220 to Ala mutation abrogates the binding of CypA to CrkII providing further evidence of the specificity of the complex (**Fig 5:9**B).

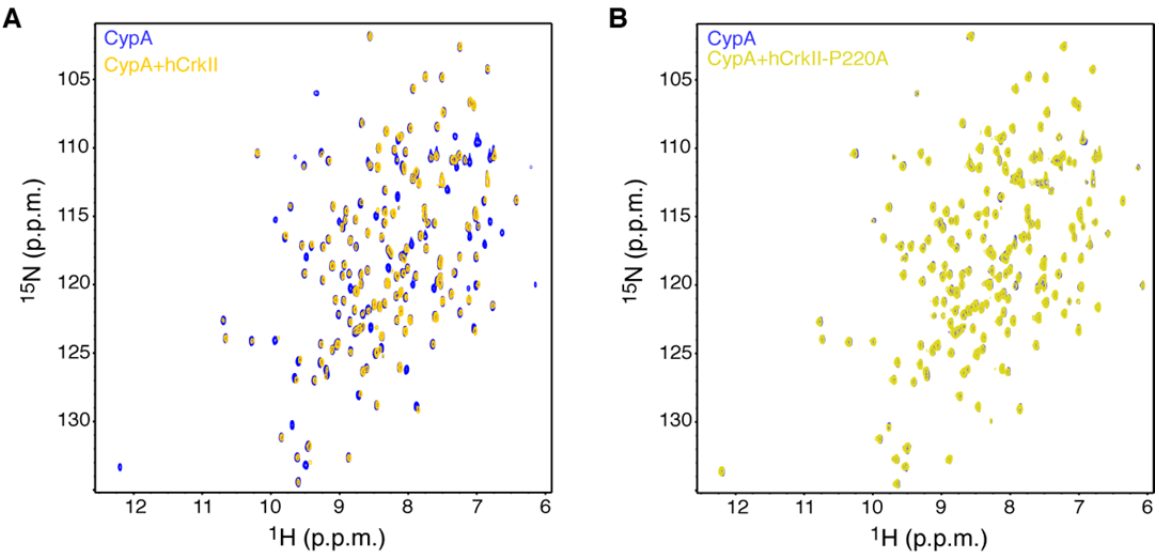
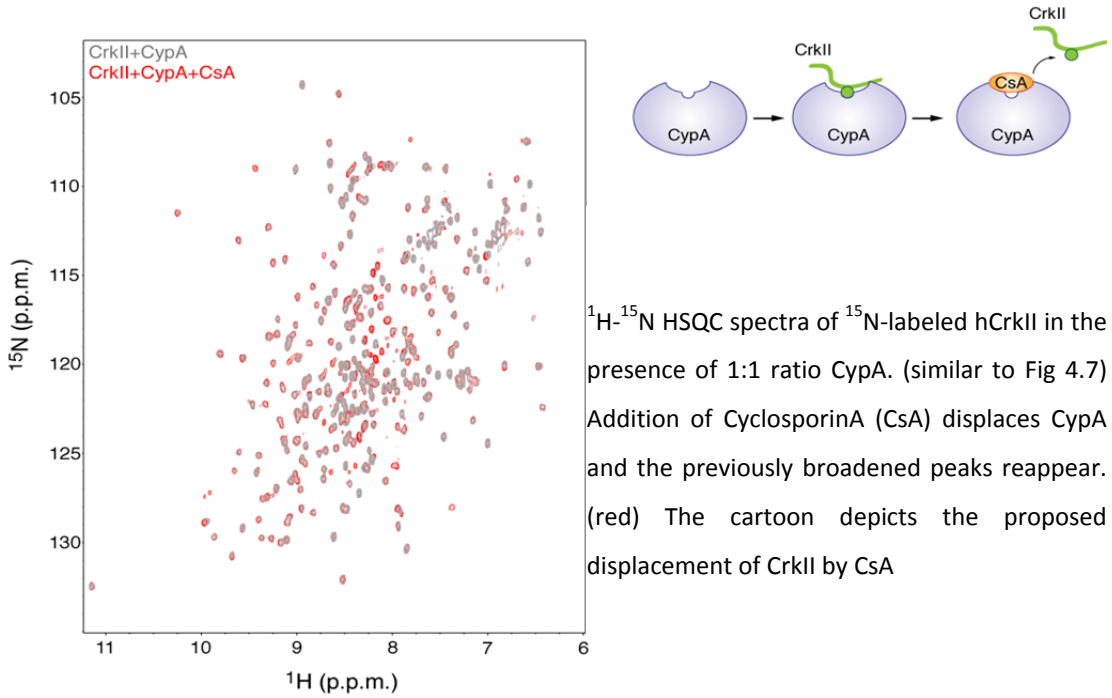


Fig 5:9 (A) ^{1}H-^{15}N HSQC of ^{15}N-labeled CypA and unlabeled hCrkII at 1:1 molar ratio. Chemical shift mapping of the affected region reveals that CypA uses the active site to interact with CrkII. **(B)** hCrkII-P220A mutation does not interact with CypA confirming that the Gly219-Pro220 site is indeed the binding region for CypA.



^{1}H-^{15}N HSQC spectra of ^{15}N-labeled hCrkII in the presence of 1:1 ratio CypA. (similar to Fig 4.7) Addition of CyclosporinA (CsA) displaces CypA and the previously broadened peaks reappear. (red) The cartoon depicts the proposed displacement of CrkII by CsA

Fig 5:10 Spectra of CrkII-CypA Interaction

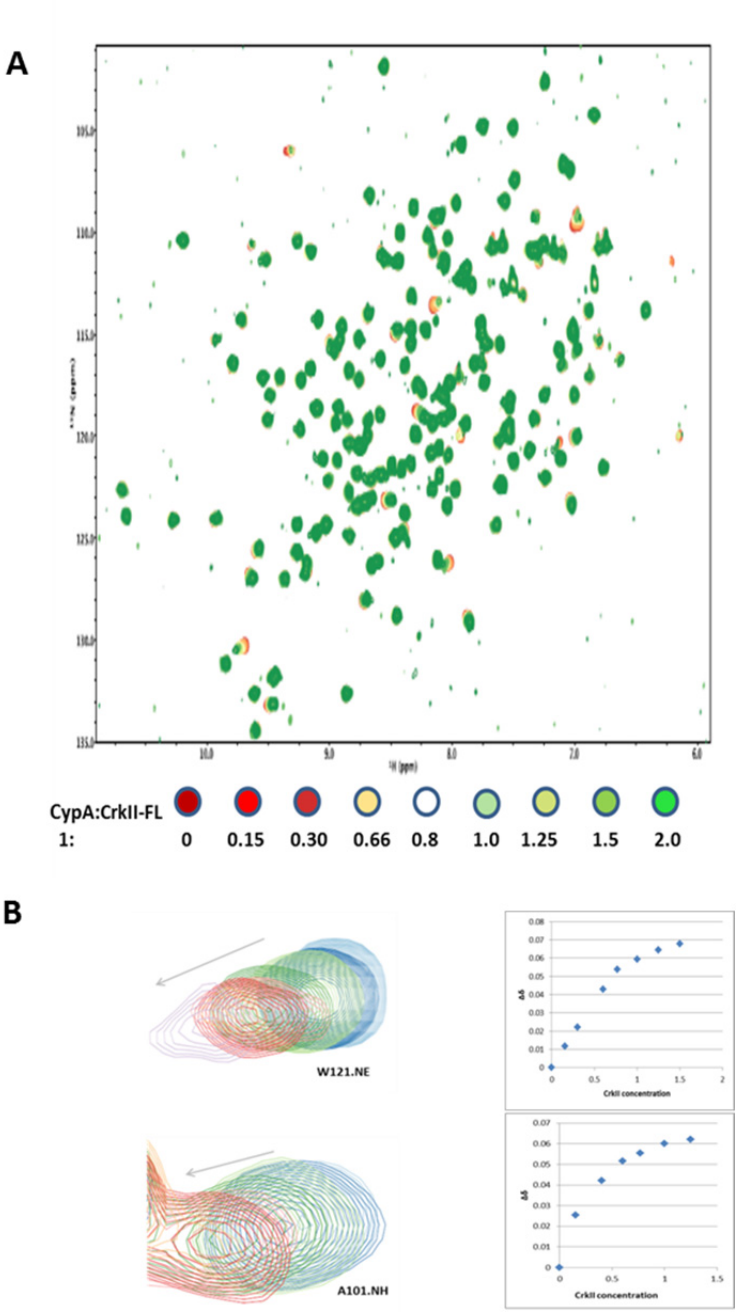


Fig 5:11 (A) ^{1}H-^{15}N HSQC spectra of ^{15}N –labeled CypA acquired with varying concentration of CrkII. The titration is color coded. Due to the relatively weak binding and the interaction displaying properties of fast-exchange in the NMR time scale, the dissociation constant could be extracted by fitting a plot of chemical shift change as a function of concentration of CrkII.

5.2.3 Structural basis for the interaction of CypA with CrkII

To understand the basis of this interaction we solved the solution structure of CypA in complex with a peptide encompassing the linker region of CrkII (Pro216-Ser225). NMR analysis of the CypA-peptide complex is similar to the binding of full length CrkII and CypA suggesting that the contacts between the two are identical. This is in agreement with the structural analysis of the interaction of CypA with other substrates. [196] A set of 3D ^{15}N-edited NOESY-HSQC and 3D ^{13}C-edited NOESY-HSQC provided sufficient information for the structure calculation. The structure reveals that Gly219-P220 region of CrkII binds deep into the active site of CypA. (**Fig 5:12**C,D) with an overall arrangement that is nearly identical to the CypA-HIV-CA structure. Pro218 is also buried into the CypA cleft, fitting snugly into a pocket formed primarily by CypA residues Thr73,Lys82, Ala101, Ala103, Thr107, and Gln111 (**Fig 5:12C,D)** The backbone carbonyl of Tyr221, the residue that becomes phosphorylated by the Abl kinase, forms a hydrogen bond with Trp121. In contrast, the side chain of Tyr221 is mostly exposed to the solvent and appears to form no contacts with CypA. The last C-terminal residue of the Crk peptide that interacts with CypA is Ala222, whose methyl group forms Van der Waals contacts with CypA residues Ile57 and Phe60. Thus, CypA uses its catalytic site to interact specifically with the CrkII linker region that encompasses the tyrosine-phosphorylation site (**Fig 5:12**A,B).

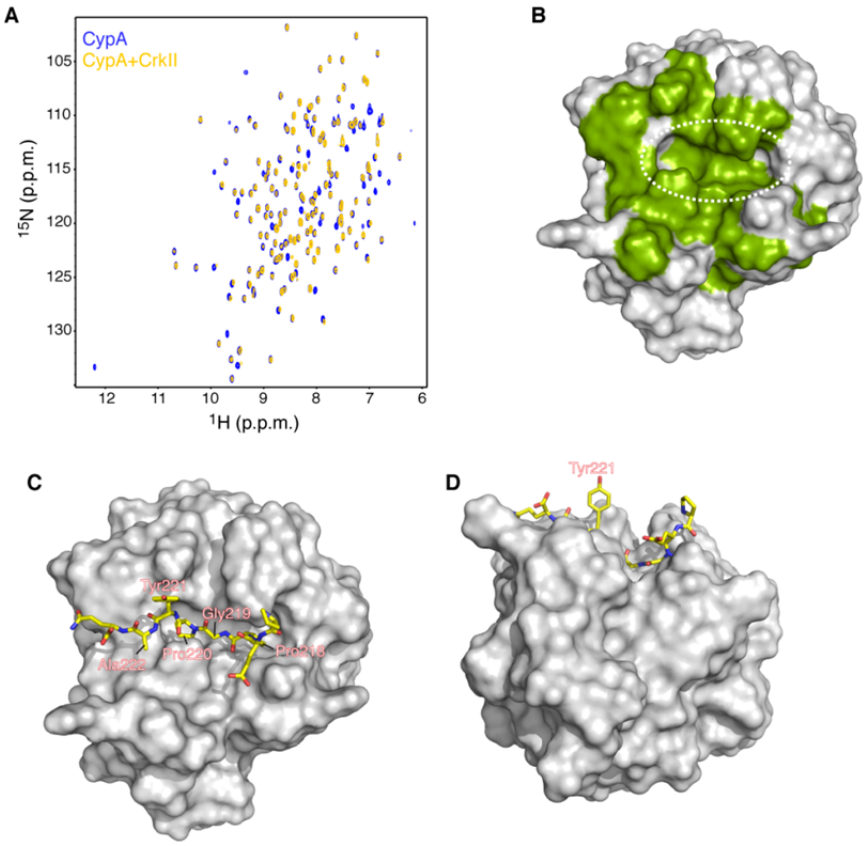


Fig 5:12 (A-B) CypA uses the active site to interact with CrkII. (**C-D)**The peptide binds deep inside the CypA active site.

The structural analysis suggested the possibility of CypA interfering with the access of Tyr221 by Abl kinase. Structural studies indicate that atleast four amino acids on either side of the phosphorylation site must make contact with the kinase for efficient catalysis. [197] When CypA is bound to CrkII, these residues are buried in the catalytic cleft and therefore the interaction of CypA with CrkII should inhibit the phosphorylation of Tyr221 by Abl. (**Fig 5:12**C,D)To test this hypothesis we carried out a kinase assay to measure the efficiency of CrkII Tyr221 phosphorylation in the presence and absence of Abl kinase. The data show that CrkII Tyr221 phosphorylation is significantly lower in the presence of CypA (up to a factor of ~12). As expected, phosphorylation increases with time since the dissociation rate of CypA-CrkII complex is relatively fast. **Fig 5:13**

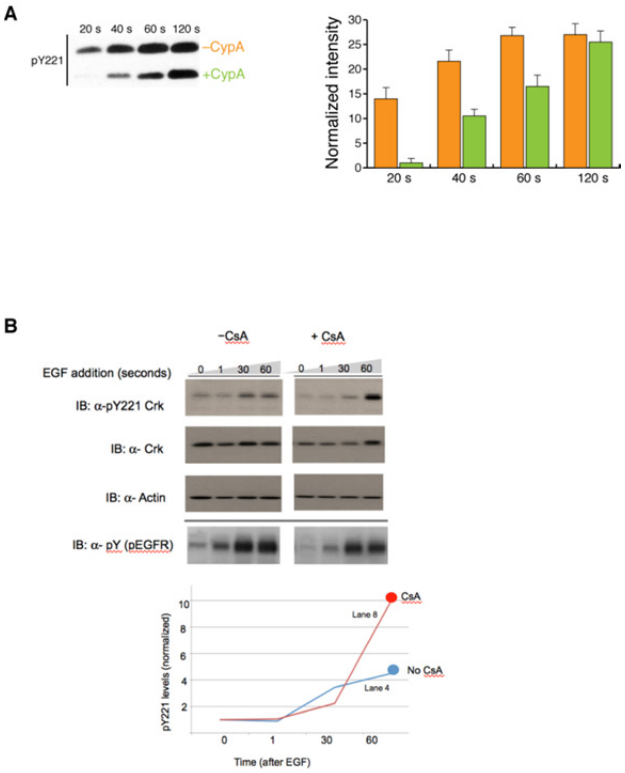


Fig 5:13 (A) Western blot for Kinase assay performed with Abl kinase and CrkII as substrate in the presence and absence of CypA. The membrane was probed with anti-CrkII pY221 antibody to monitor phosphorylation. Quantification performed in ImageJ. In the presence of CypA there is a ~12 fold inhibition of Tyr221 phosphorylation. **(B)** Western Blot probing for pY221 phosphorylation in EGF-stimulated MDA-MB-468 cells in the presence and absence of CsA.

To examine this effect of CypA on CrkII phosphorylation *in vivo* EGF-stimulated CrkII phosphorylation in breast cancer line MDA-MB-468. CsA treatment increases Tyr221 phosphorylation by a factor of ~3. (Fig 5:13**B**) CypA is recruited by CrkII in cells and can inhibit phosphorylation.

5.2.4 *In-vivo* localization of CypA and CrkII

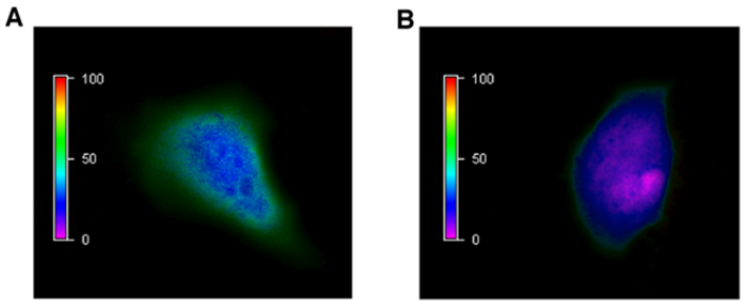


Fig 5:14(A) FRET analysis performed on HELA cells transfected with CrkII-YFP and CypA-CFP. The cell shows a maximum FRET efficiency of 70% near the membrane. (B) Treatment of cell with CsA leads to a decrease in FRET efficiency ~15%.

To investigate the frequency with which this interaction occurs in cells we used in cell FRET analysis. YFP and CFP constructs of CrKII and CypA were prepared and expressed in HELA cells. FRET analysis was performed on cells that were confirmed to be expressing both proteins. Both CypA and CrkII were observed to be in the cytosol. Maximum FRET was observed near the membrane (~70%) and the addition of CsA lead to a dramatic loss of FRET. (~15%.) Transfection of CrkII-P220A mutant also showed no interaction with CypA. This is consistent with our NMR studies which showed that the interaction is mediated by the catalytic region of CypA and the G219-P220 bond of CrkII.

5.3 Discussion.

The role of CypA as a PPIase has been the highlight of numerous studies. Its primary role in catalyzing the *cis-trans* prolyl isomerization has been studied in detail. Pin1 has been shown to interact with proteins in a phosphorylation-dependent manner. It plays an important role in the packaging of HIV-1 virion which requires the interaction of CypA and the HIV-CA protein.[198, 199] In this case CypA plays the role of a binding partner and not as a catalyst.[200] Recently there have been reports of CypA acting as a binding partner for extracellular receptor CD147.

CypA interacts with the receptor CD147 using an alternative binding site from that of the catalytic center that induces chemotaxis upon binding to CD147.[201] Also, the role of CypA as a catalyst has not been demonstrated in cell signaling.

We have shown that CypA interacts with the adaptor protein CrkII using a proline containing motif in the linker region. The structural data reported here demonstrate that the interaction of CrkII and CypA resembles previously published data of CypA and HIV-1 CA peptide structure. Although CypA interacts with CrkII using the catalytic center our data show that CrkII interacts with CypA as a binding partner and not a ligand. We show that this interaction inhibits the phosphorylation of CrkII by Abl and delays the inactivation of CrkII. Abl phosphorylation of CrkII at Y221 creates a binding motif for the SH2 domain and the intramolecular interaction prevents the SH2 domain from interacting with other binding partners. The reorganization also prevents access to the interacting surface of the $SH3^N$ domain thus rendering CrkII inactive. In light of the fact that the only known function of CrkII is protein binding and that the SH2 and SH3 domains are essential for CrkII function, delay in the inhibitory Y221 phosphorylation would have a direct impact on its ability to interact with cellular proteins.

It should be noted that in our assays, even in the presence of CypA, CrkII phosphorylation levels reach a plateau but it is the delay in phosphorylation that is most interesting.**Fig 5:13** CrkII forms dynamic signaling complexes in cells that are in constant turnover and a delay of a few seconds can have profound impact on the outcome of a particular signaling cascade[76].

Interestingly, CrkL which shares a 76% sequence identity with CrkII does not contain a heterogeneous proline in the linker region. The residue Tyr207 that is phosphorylated by Abl kinase is preceded by Ala (Fig 5:3) CrkL is thus not expected to interact with CypA. A detailed NMR analysis of CrkL and CypA demonstrate that this is indeed the case. (**Fig 5:15**)

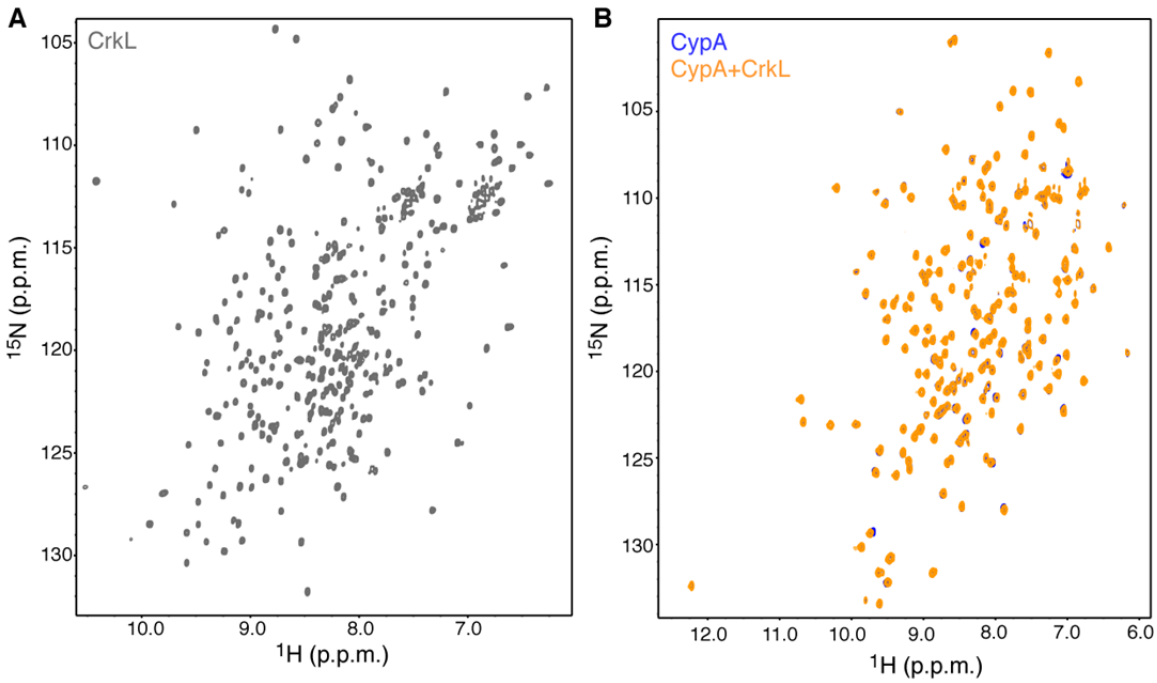


Fig 5:15 (A) ^{1}H-^{15}N HSQC of ^{15}N-labeled CrkL. (B) ^{1}H-^{15}N HSQC of ^{15}N-labeled CrkL and 2 fold molar excess of unlabeled CypA shows no interaction between these two proteins.

Kinase assay performed with Abl kinase and CrkL as a substrate also show inhibition of phosphorylation by CypA. (**Fig 5:16**)

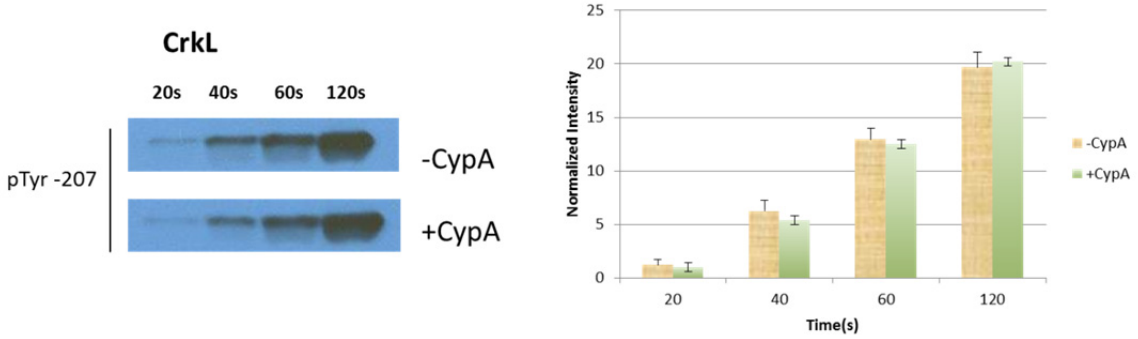


Fig 5:16 Western blot for kinase assay performed with Abl Kinase and CrkL as a substrate. The membrane was probed with anti-CrkL-Y207 antibody. Quantification performed in ImageJ.

This suggests that CrkII may be subject to an additional level of regulation that is not available for CrkL. This can also provide an explanation for the observation that CrkL is usually found to be the preferred substrate for Bcr-Abl. It is possible that in tumor cells where Bcr-Abl is active, CypA

levels are also elevated. Since both CrkII and CrkL are equally good substrates for Bcr-Abl, the presence of CypA will inhibit CrkII Tyr221 phosphorylation compared to CrkL.

Because of the relatively weak interaction between CypA and CrkII, this interaction may have failed identification from screenings run in search of binding partners for Crk. In certain cellular context, especially in tumor cells where the levels of CrkII and CypA are elevated, the CrkII-CypA complex can have a significant effect in cell phenotype.

Along with NMR, we use a variety of biochemical techniques to identify a previously unreported interaction between CypA and CrkII.

6 Materials and Methods.

6.1 Protein Preperation of Crk fragments and Abl kinase.

The various fragments of Crk protein were cloned from a template plasmid that encodes full length CrkII or CrkL, The constructs were cloned into the pet42a vector using the NcoI and Xho1 restriction sites. A Tev protease cleavage site was introduced between the histidine tag and the protein The cloned vectors were sequenced to verify the insert. Vector were then transformed into BL21(DE3) cells followed by plating on Lysogeny broth (LB) plate. A single colony was picked from the plate and was transferred into 5 ml of autoclaved LB media (for the preparation of deuterated samples, LB media was dissolved in 75 % D2O). After ~ 8 hour growing, the cells were transferred in 100 mL of LB media. Before transferring, the cells was centrifuged and washed with fresh LB media and kept overnight for growing at 37 °C. Then the cells were transferred into 1L of LB media. The cells were grown at 37 °C in the presence of ampicilin and chloramphenicol. Protein synthesis was induced by the addition of 0.5 mM of IPTG at A600 ~ 0.4 and cells were harvested at A600 ~ 1.0 and centrifuged at 5000 *x g* in a 6-liter JLA-8.100 rotor (Beckman). The lysate was loaded on Ni-NTA agarose resin (GE) preequilibrated with Tris buffer and 1 M NaCl. For the final purification step, the sample was concentrated and applied to a Superdex 75 size-exclusion column (GE). For NMR studies, the samples were dialyzed in NMR buffer (50 mM KPi (pH 6.8), 140 mM NaCl and 1 mM DTT) and concentrated using Amicon cell units (Millipore). All fragments are monomeric in solution at concentrations used for the NMR studies (typically 0.6-1.0 mM), as indicated by gel filtration and light scattering.

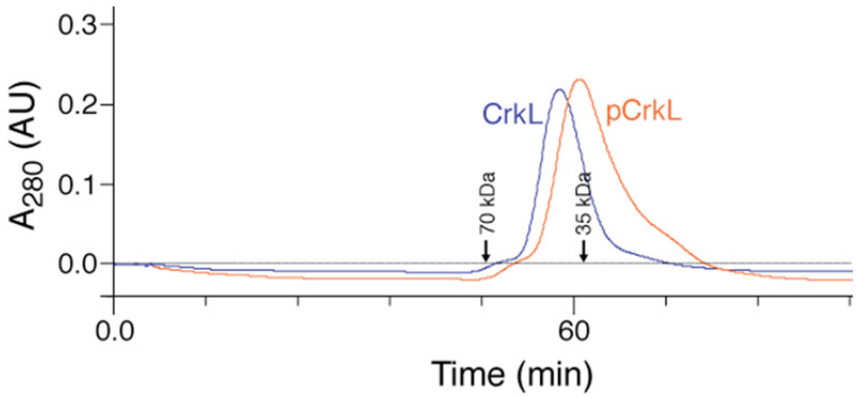


Fig 6:1 Gel filtration profile of full-length CrkL and pCrkL.

Isotopically labeled samples for NMR studies were prepared by growing the cells in M9 minimal media containing 1 gL-1 of 15NH4Cl and 2 gL-1 13C6-glucose, 2ml of 1M MgSO4L-1, 100µl of 1M CaCl2 L-1.

AblKD and AblPxxP were expressed and purified as described before[202] with minor modifications, including extension of the incubation period to 24 h. Protein concentration was determined spectrophotometrically at 280 nm using an extinction coefficient

6.2 Purification of CypA

His6-tagged CypA was transformed into BL21(DE3) cells and plated on LB plate. A single colony was picked from the plate and was transferred into 5 ml of autoclaved LB media. After ~ 8 hour growing, the cells were transferred in 100 mL of LB media. Before transferring, the cells was centrifuged and washed with fresh LB media and kept overnight for growing at 37 °C. Then the cells were transferred into 1L of LB media. The cells were grown at 18 °C in the presence of ampicilin. Protein synthesis was induced by the addition of 0.5 mM of IPTG at A600 ~ 0.4 and

cells were harvested at A600 ~ 1.0. The cells were re-suspended into Tris, 400mMNaCl, 30mM immidazole, 10mM BME, pH8.0.

Cell were sonicated and centrifuged at 50,000×g using a JA-25.50 rotor (Beckman). The supernatant was loaded on a Ni-column pre-equilibrated with buffer containing 50mM Tris, 400mMNaCl, 30mM immidazole, 10mM BME with a flow rate of 0.7/ml/min, followed by two-column volume wash with 50mM Tris, 1MNaCl, 30mM immidazole, 10mM BME, pH8.0 (high-salt buffer). In next step column was washed with two column volume of low salt buffer, i.e. 50mM Tris, 400mMNaCl, 30mM immidazole, 10mM BME, pH8.0. Finally the protein was eluted with buffer containing 50mM Tris, 140mM NaCl, 400mM immidazole, 10mM BME with a pH8.0.

The eluted protein was then exchange into buffer containig 140mM NaCl, 50mM KPi, pH6.5, 1mM BME, 5% glycerol using an amicon stirred cell. Later on, the sample was concentrated and applied to the Superdex-75 size exclusion column, previously equilibrated with the same buffer. Finally the protein was exchanged into desired salt buffer. Protein concentration was determined spectrophotometrically at 280 nm using an extinction coefficient of 8470 M-1 cm-1.

6.3 NMR Spectroscopy

All NMR experiments were performed on Varian 800- and 600-MHz and Bruker 700- and 600-MHz spectrometer. Sequential assignment of the 1H, 13C, and 15N protein backbone

chemical shifts was achieved by means of through-bond heteronuclear scalar correlations using the following 3D pulse sequences: 3D HNCO, 3D HN(CA)CO, 3D HNCA, 3D HN(CO)CA, 3D HNCACB, and 3D HN(CO)CACB. Side-chain assignment was performed using 3D C(CO)NH and 3D H(CCO)NH spectra. NOEs were assigned and collected on the basis of 3D 15N-NOESY-HSQC and 13C-NOESYHSQC spectra using mixing time of 100 and 80ms respectively. All NMR samples were prepared in 50 mM KPi (pH 6.8), 140 mM potassium phosphate, 1 mM BME and 1g L-1 NaN3.

Concentration of all protein NMR sample was 0.6-0.8 mM. Spectra were recorded at 22 °C, 25 °C and 32 °C. 2D ^{1}H-^{15}N HSQC spectra were recorded for all constructs in different concentrations, pH, and temperatures too. All spectra were processed using the NMRPipe software package and analyzed with NMRView.

For all the experiment using CypA and c-Abl we use similar buffer condition with pH6.5 and temperature 20°C or 25°C.

6.3.1 2D ^{1}H-^{15}N Heteronuclear (ZZ NMR experiments)

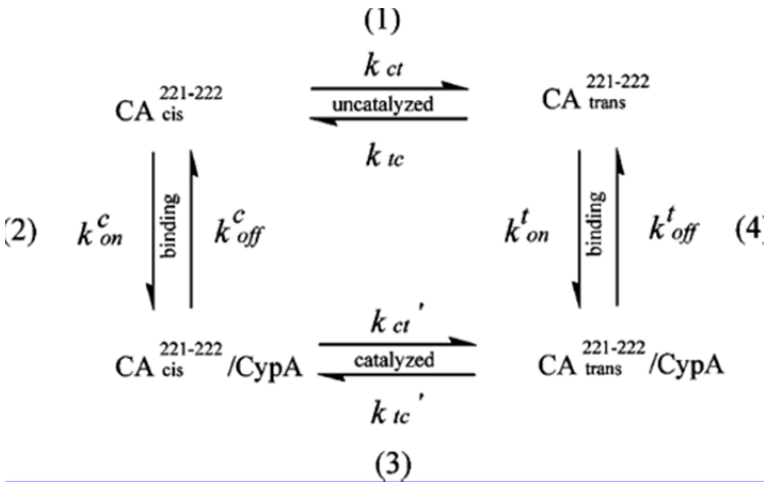


Fig 6:2 Schematic representation of CypA catalyzed *cis-trans* isomerization.

The 2D 1H-15N heteronuclear (ZZ) NMR exchange experiments [203]were used to measure the rate of CypA-catalyzed *cis-trans* isomerization (kex) of Gly237-Pro238. This pulse sequence allows observation of the transfer of 15N-labeled heteronuclear zz-magnetization between the two conformations. The rate constant of *cis-trans* interconversion in the presence of CypA, can be determined by measuring the intensity of the exchange and auto peaks as a function of the

mixing time and then to fit to the equations described below[204] The equations and methods are stated below.

$$I_{tt}(\tau_m) = I_t(0)[(-(\lambda_2 - a_{11})\, e(-\lambda_1\tau_m) + (\lambda_1 - a_{11})\, e(-\lambda_2\tau_m)/(\lambda_1 - \lambda_2)] \quad (1a)$$

$$I_{tc}(\tau_m) = I_t(0)[(a_{21}\, e(-\lambda_1\tau_m) - a_{21}\, e(-\lambda_2\tau_m)/(\lambda_1 - \lambda_2)] \quad (1b)$$

In eq a and b, I_{tt} is the intensity of the X trans of X-P auto peak (in this case Gly219) and I_{tc} is the intensity of the *trans* exchange peak at mixing time τ_m. $I_t(0)$ is the intensity of the trans peak at τ_m = 0.$\lambda_{1,2}$= ½[(a_{11}+a_{22}) ± [(a_{11}-a_{22})^2 + 4$k_{tc}k_{ct}$]$^{1/2}$], a_{11} = R_{1t} + k_{tc}, a21 = -k_{tc} and a_{22} = R_{1c} + k_{ct}. R_{1c} and R_{1t} are the longitudinal relaxation rates.

The rate of CypA-catalyzed chemical exchange (k_{ex}) in presence of sub stoichiometric amount of CypA is the sum of the individual rate constants in steps 2-4 in Scheme 1 as shown in eq 2. p_c and p_t denotes populations of X in the *cis* and *trans* conformation, respectively, and were determined from the relative intensities of the *trans* and *cis* resonances of X in an HSQC spectrum. The equilibrium constant (K_{eq}) for the *cis* and *trans* conformations can be determined by eq 1. From eq 1, the individual rate constants k_{ct} and k_{tc} are simplified to p_tk_{ex} and p_ck_{ex}, respectively, in eqs 1a and b. Fitting the exchange curves with eqs a and b results in the rate of CypA-catalyzed X-P *cis-trans* isomerization (k_{ex}).

6.3.2 Residual dipolar Coupling.

Alignment of the proteins for RDC measurements was achieved using poly(ethylene glycol)/alcohol mixtures[133]. A 5% C12E5/hexanol (molar ratio=0.96) mixture was prepared.

C12E5 was used to a final concentration of 5% (w/w) in 90% H2O:10% D2O solution. The pH was adjusted using sodium hydroxide. The amount of hexanol was added dropwise, while vigorously shaking, to a final molar ratio C12E5:hexanol of 0.96. Air bubbles were removed by centrifugation at 5,000 × g for few minutes. The HDO quadrupolar deuterium splitting was checked to confirm the presence of the crystalline phase (a splitting of ~20 Hz was observed). For the measurement of RDCs in the protein, 250 µl of the C12E5: hexanol stock solution was added into 50 µl of protein in buffer. 15N-HSQC (IPAP) and HNCO based experiments were used to measure one-bond N-H and CαC' RDCs[205]. the accuracy of experimentally determined RDC values, the back calculated values were plotted versus the measured ones. The program MODULE was used to properly define the aligment tensor parameters from residual dipolar couplings. The perfect agreement between experimental and calculated RDC values signifies the accuracy and overall quality of our final structure

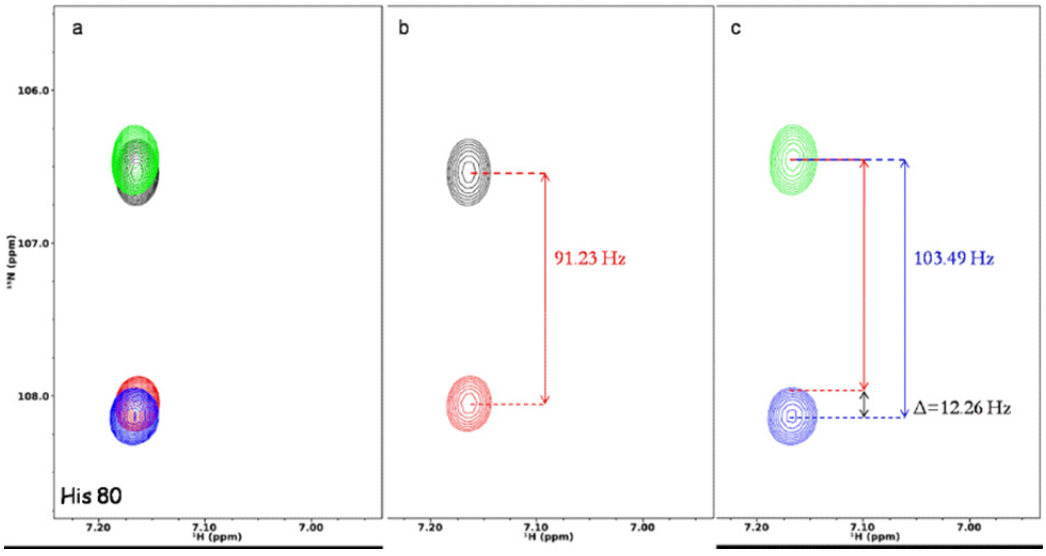


Fig 6:3 A sample of the splitting observed in RDC measurements.

6.3.3 Relaxation measurements and analysis

Three relaxation parameters were measured for all backbone amides of CrkL: the ^{1}H-^{15}N NOE, the longitudinal relaxation rate R1 and the transverse relaxation rate R_2. ^{15}N R_1 values were measured from 2D spectra recorded with relaxation delays 10, 60, 100, 200, 300, 450, 600, 750, 900, 1200 and 1400 ms; 15N R2 values were measured from 2D spectra recorded with relaxation delays 6.4, 19.2, 32.0, 44.8, 57.6, 70.4, 83.2, 96.0 ms. Data sets were acquired as 152 x 1,024 complex points in the t1 x t2 time-domain dimensions. Data points were fitted as a function of the length of the parametric relaxation delay to two-parameter decay curves of the form $I(t)=I_0 e^{-Rt}$, where I is the intensity of the magnetization. ^{1}H-^{15}N NOE data were obtained by recording, in an interleaved manner, one spectrum with a 3^{-s} recycle delay followed by a 3^{-s} saturation and another spectrum with no saturation and a 6-s recycle delay. Correlation times and Rex values were determined by using the model free approach[206].

6.4 Isothermal titration Calorimetry

All calorimetric titrations were performed on a iTC200 microcalorimeter (GE). Protein samples were extensively dialyzed against the ITC buffer containing 50 mM KPi (pH 6.8), 150 mM NaCl, and 1 mM TCEP. All solutions were filtered using membrane filters (pore size, 0.45mm) and thoroughly degassed for 20 min by gentle stirring under vacuum. The sample cell was usually filled with a 40 µM and the injection syringe with 400 µM solutions respectively. Ligand solutions were prepared by dissolving the peptide ligand in the flow through of the last buffer exchange. Each titration typically consisted of a preliminary injection followed by ~16 subsequent injections. Data for the preliminary injection, which are affected by diffusion of the solution from and into the injection syringe during the initial equilibration period, were discarded. The data were fitted with Origin 7.0.

The sequences of the CrkL-pTyr207-, CrkII-pTyr221-, PPII- and FGFR-pTyr-peptides used are EPAHApYAQPTT, PEPGPpYAQPSV, YLQAPELPTKTRTS and AGVSEpYELPEDPR, respectively

6.5 Structure Calculation

Structure calculations were performed with Xplor-NIH. The $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{13}C'$, H_{α}, ^{15}N and NH chemical shifts served as input for the TALOS+ program[129] to extract dihedral (ϕ and ψ) angles. The initial structure was calculated using NOEs, PREs, dihedral angles and hydrogen bonds. RDC restraints were included in the final stages of the calculation. Ramachandran statistics are as follows: most favored regions, 90%; allowed regions, 8%; disallowed regions, 2%. The summary of NMR restraints and structure refinement statistics are provided below.

Supplementary Table 1 NMR and refinement statistics for CrkL structures

	CrkL	pCrkL
NMR distance and dihedral constraints		
Distance constraints		
Total NOE	3964	3814
Intra-residue	763	733
Inter-residue		
Sequential ($\|i-j\|=1$)	1010	1001
Medium-range ($\|i-j\|<4$)	522	483
Long-range ($\|i-j\|>5$)	1669	1598
Inter-domain (SH2-SH3^N)	8	8
Hydrogen bonds	63	61
Paramagnetic relaxation enhancement rates (PREs)	37	35
Total dihedral angle restraints		
ϕ	90	86
ψ	93	90
Total RDCs		
$^1D_{HN}$	87	82
Q (%)	21	23
Structure statistics		
Violations (mean and s.d.)		
Distance constraints (Å)	0.11±0.017	0.12±0.019
Dihedral angle constraints (°)	0.68±0.13	0.79±0.14
Max. dihedral angle violation (°)	8.6	6.3
Max. distance constraint violation (Å)	0.91	0.87
Deviations from idealized geometry		
Bond lengths (Å)	0.006	0.008
Bond angles (°)	0.8	1.0
Impropers (°)	1.65	1.89
Average pairwise r.m.s. deviation** (Å)		
Heavy	1.0	1.1
Backbone	0.7	0.7

** Pairwise r.m.s. deviation was calculated among 20 refined structures for the SH2-SH3^N region

Fig 6:4 Summary table of statistics for solution structure of CrkL and pCrkL.

6.6 PRE measurements

Nitroxide spin labels (MTSL; Toronto Research Chemicals Inc.) were introduced via cysteine-specific modification of engineered CrkL derivatives containing single-solvent-accessible cysteine residues. The wild-type Cys44 and Cys249 residues were mutated to Ser to provide the protein scaffold for introducing the cysteine-specific modifications at the following sites: Ser20, Ile90, Asp133, Gly147, and Val175. These mutants and their MTSL derivatives were determined not to perturb the CrkL structure, as assessed by ^{1}H-^{15}N HSQC spectra, and were used for measuring PRE rates. After purification, proteins were exchanged into phosphate buffer (50 mM KPi (pH 6.8), 150 mM NaCl, and 1 mM β-mercaptoethanol). β- mercaptoethanol was removed by Zeba spin desalting column (Thermo Scientific) according to the manufacturer's protocol. MTSL was added from a concentrated stock in acetonitrile at a 10-fold excess, and the reaction was allowed to proceed at 4 °C for ~12h. The completion of the reaction was confirmed by mass spectrometry. Excess MTSL was removed by extensive dialysis using an Amicon stirred cell, and the pH was corrected to 6.8. PRE-derived distances were determined from 1H-15N-HSQC spectra of CrkL by measuring peak intensities before (paramagnetic) and after (diamagnetic) reduction of the nitroxide spin label with ascorbic acid. PRE values then were converted to distances by using a modified Solomon-Bloembergen equation for transverse relaxation, as described previously.[123]. Ensemble simulated annealing refinement was used as described. [207]. Two sets of restraints were incorporated into subsequent structure calculations. Amide resonances strongly affected by the presence of the spin label in the peptide (Ipara/Idia < 0.15) and whose resonances broaden beyond detection in the paramagnetic spectrum were restrained with only an upper-bound distance estimated from the noise of the spectrum plus 4 Å. Amides whose resonances appear in the paramagnetic spectra (Ipara/Idia < 0.85) were restrained as the calculated distance with (+/-) 4-Å upper/lower bounds

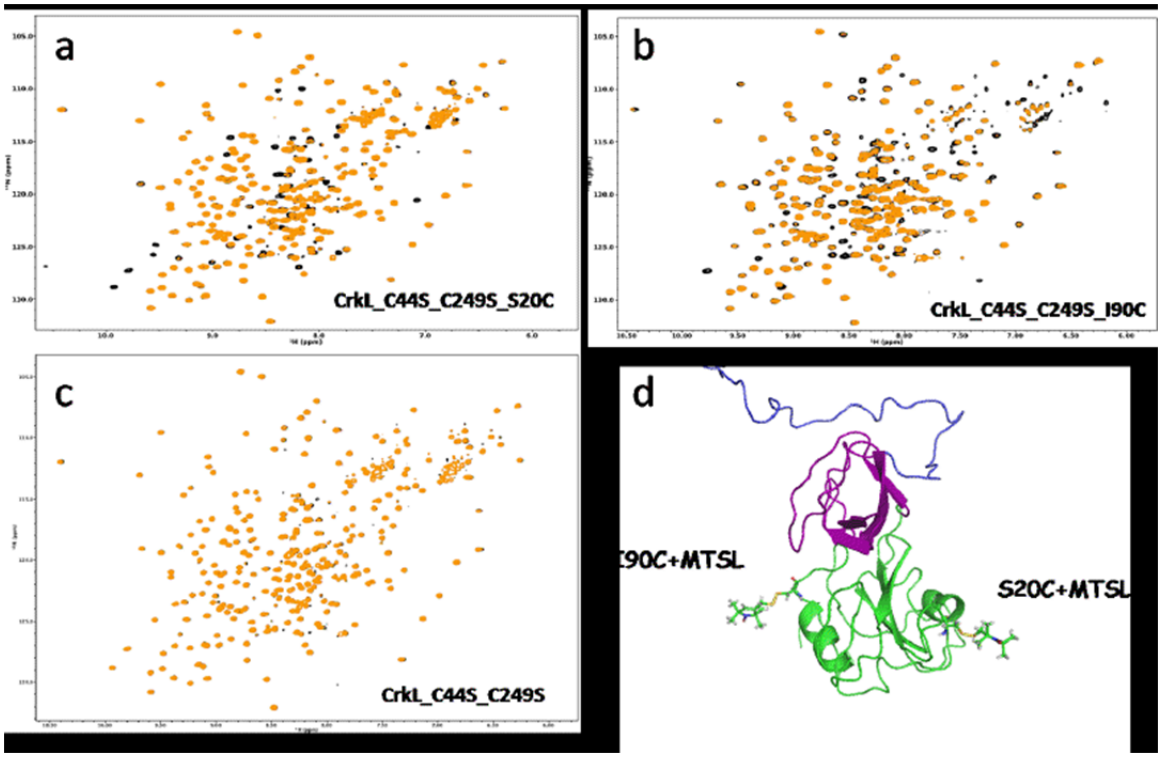


Fig 6:5 (a) Spectra of PRE effects of S20C mutation. (b) Spectra of PRE effects of 190C.

(c) Spectra of non-cysteine mutant CrkL(control) . (d) Ribbon Diagram depicting location of S20C in the context of the SH2-SH3 fragment of CrkL.

6.7 Mass Spectrometry

CrkL and pCrkL samples (50-100 µM in NMR buffer) were diluted to 0.1 X in molecular biology grade water. Each sample was then mixed in a 1:1 ratio with sinapinic acid and spotted in 0.5 µl volumes on a stainless steel MALDI-TOF sample plate and allowed to air dry. All samples were then analyzed in a Voyager DE MALDI-TOF instrument from Applied Biosystems in linear positive mode using the following settings: 25,000 V accelerating voltage, 93% grid voltage, 0.3% guide wire voltage, and 750 nsec extraction delay time. All spectra were analyzed using Data Explorer software v 5.10.2 from Applied Biosystems.

6.8 Kinase assay and Western Blots

6.10. Western Blot

We performed western blot by using purified protein. For detection of phophorylated Tyr221 phospho-CrkII (Tyr221) antibody was purchased from Cell Signaling. All reactions are carried out at room temperature and all proteins are exchanged into the same buffer condition (50mM KPi, 140mM NaCl, 1mM BME, pH6.5). First the substrate CrkII or CrkL was aliquated and then an excess (~6-7 times) of CypA was added to achieve saturation. Abl kinase was then added in catalytic amounts and then the mixture was supplemented with 1mM $MgCl_2$. The reaction was started with the addition of ATP and stopped by adding SDS sample buffer. All the samples were run on a 10% SDS gel. In the next step the protein was transfered from SDS-gel to nitro-cellulose membrane using gel-electrophoresis [transfer buffer: 700ml H2O, 100ml 10X transfer buffer (30.4g Tris, 144.1g glycine),200ml MeOH]. Nitrocellulose membrane was blocked in milk (5% non-fat dry milk in 1X TBST buffer and 0.2% tween). 1X TBST buffer is made from 10X TBST buffer i.e. tris-base 24.2g/L, NaCl 80g/L and pH7.5. After washing 2-3 times with 1X TBST buffer membrane was incubated overnight in primar antibody solution (substrate to antibody ratio; 1:1000). Next morning we wash out exces of primary antibody by using 1X TBST buffer for 3-4 times. We then incubate it with secoundry antibody in blocking solution for firther one more hour. Secoundry antibody was use for detection purpose. In our case we use secondary antibody from goat with horseradish peroxidase label. We use Pierce ECL Western Blotting Substrate for the detection of horseradish peroxidase (HRP) on immunoblots. Finally, the blots were exposed to X-ray film to obtain the results. Blots were analyzed using Image J software.

6.9 FRET analysis

Generation of Constructs

The vectors containing YFP and CFP were obtained from Addgene. Full-length CrkII and CypA were closed into the vectors. For CrkII we subcloned the region encoding the CrkII sequence into

pcDNA3-YFP vector from Addgene using the Not1 and BamH1 sites for CRKII-YFP. CypA was subclosed into pcDNA3-CFP vector using the Not1 and EcoRI sites for CypA-CFP. Both constructs were inserted at the N-terminal of the fluorescent tag. Clones were confirmed by sequencing

Cell Culture

HELA cells were cultered in DMEM supplemented with 10% Fetal Bovine Serum, 1% Penicillin/Streptomycin and 1% Glutamax in a humidified incubator at 37C 5% CO_2.The cells were grown to 70% confluency and then transfected with the appropriate plasmids at 0.4ug using X-tremeGENE Transfection Reagent (Roche) as per their instructions. Cells were washed with PBS 4-6 hours after delivery. Images were acquired 24hrs after transfection. The cells were imaged live as this gave the most consistent results. A Nikon Ti Eclipse Inverted Microscope was used to capture images.

FRET measurements

Observations were made using 433 nm excitation and emission window collected between 507-547nm using a Nikon Ti Eclipse Inverted Microscope. The Sensitized emission method was used to image the FRET effect. The built-in Nikon software (NIS-Element) was used to measure FRET efficiency.

The FRET efficiency is calculated using the following formula

$$FRET_{CORRECTED} = Da_{FRET} - Dd_{FRET} \cdot \frac{Da_{DONOR}}{Dd_{DONOR}} - Aa_{FRET} \cdot \frac{Da_{ACCEPTOR}}{Aa_{ACCEPTOR}}$$

Sentisized Emission

$$\text{EFFICIENCY [\%]} = (FRET_{CORRECTED} / Dd_{FRET}) * 100$$

Fig 6:6 Equation to calculate FRET efficiency.

7 Bibliography

1. Birge, R.B., et al., *Crk and CrkL adaptor proteins: networks for physiological and pathological signaling.* Cell Commun Signal, 2009. **7**: p. 13.
2. Mayer, B.J., M. Hamaguchi, and H. Hanafusa, *A novel viral oncogene with structural similarity to phospholipase C.* Nature, 1988. **332**(6161): p. 272-5.
3. Mayer, B.J. and H. Hanafusa, *Mutagenic analysis of the v-crk oncogene: requirement for SH2 and SH3 domains and correlation between increased cellular phosphotyrosine and transformation.* J Virol, 1990. **64**(8): p. 3581-9.
4. Matsuda, M., et al., *Two species of human CRK cDNA encode proteins with distinct biological activities.* Mol Cell Biol, 1992. **12**(8): p. 3482-9.
5. Pawson, T., *Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems.* Cell, 2004. **116**(2): p. 191-203.
6. Ren, R., et al., *Identification of a ten-amino acid proline-rich SH3 binding site.* Science, 1993. **259**(5098): p. 1157-61.
7. Feller, S.M., *Crk family adaptors-signalling complex formation and biological roles.* Oncogene, 2001. **20**(44): p. 6348-71.
8. Schlessinger, J. and M.A. Lemmon, *SH2 and PTB domains in tyrosine kinase signaling.* Sci STKE, 2003. **2003**(191): p. RE12.
9. Lowenstein, E.J., et al., *The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling.* Cell, 1992. **70**(3): p. 431-42.
10. Schlessinger, J. and A. Ullrich, *Growth factor signaling by receptor tyrosine kinases.* Neuron, 1992. **9**(3): p. 383-91.
11. Bradshaw, J.M., V. Mitaxov, and G. Waksman, *Investigation of phosphotyrosine recognition by the SH2 domain of the Src kinase.* J Mol Biol, 1999. **293**(4): p. 971-85.
12. Songyang, Z., et al., *A single point mutation switches the specificity of group III Src homology (SH) 2 domains to that of group I SH2 domains.* J Biol Chem, 1995. **270**(44): p. 26029-32.
13. Songyang, Z. and L.C. Cantley, *SH2 domain specificity determination using oriented phosphopeptide library.* Methods Enzymol, 1995. **254**: p. 523-35.
14. Pascal, S.M., et al., *Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide.* Cell, 1994. **77**(3): p. 461-72.
15. Rahuel, J., et al., *Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode.* Nat Struct Biol, 1996. **3**(7): p. 586-9.
16. Kuriyan, J. and D. Cowburn, *Modular peptide recognition domains in eukaryotic signaling.* Annu Rev Biophys Biomol Struct, 1997. **26**: p. 259-88.
17. Birge, R.B., et al., *Identification and characterization of a high-affinity interaction between v-Crk and tyrosine-phosphorylated paxillin in CT10-transformed fibroblasts.* Mol Cell Biol, 1993. **13**(8): p. 4648-56.

18. Sakai, R., et al., *A novel signaling molecule, p130, forms stable complexes in vivo with v-Crk and v-Src in a tyrosine phosphorylation-dependent manner.* EMBO J, 1994. **13**(16): p. 3748-56.
19. Garcia-Guzman, M., et al., *Met-induced JNK activation is mediated by the adapter protein Crk and correlates with the Gab1 - Crk signaling complex formation.* Oncogene, 1999. **18**(54): p. 7775-86.
20. Smit, L. and J. Borst, *The Cbl family of signal transduction molecules.* Crit Rev Oncog, 1997. **8**(4): p. 359-79.
21. Chen, C.H., et al., *Two closely spaced tyrosines regulate NFAT signaling in B cells via Syk association with Vav.* Mol Cell Biol, 2011. **31**(14): p. 2984-96.
22. Zhang, Y.Y., et al., *Simultaneous Binding of Two Peptidyl Ligands by a Src Homology 2 Domain.* Biochemistry, 2011. **50**(35): p. 7637-7646.
23. Nguyen, J.T., et al., *Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors.* Science, 1998. **282**(5396): p. 2088-92.
24. Lim, W.A., F.M. Richards, and R.O. Fox, *Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains.* Nature, 1994. **372**(6504): p. 375-9.
25. Feng, S., et al., *Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands.* Proc Natl Acad Sci U S A, 1995. **92**(26): p. 12408-15.
26. Knudsen, B.S., et al., *Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins.* EMBO J, 1995. **14**(10): p. 2191-8.
27. Wu, X., et al., *Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk.* Structure, 1995. **3**(2): p. 215-26.
28. Glenney, J.R., Jr. and L. Zokas, *Novel tyrosine kinase substrates from Rous sarcoma virus-transformed cells are present in the membrane skeleton.* J Cell Biol, 1989. **108**(6): p. 2401-8.
29. West, K.A., et al., *The LD4 motif of paxillin regulates cell spreading and motility through an interaction with paxillin kinase linker (PKL).* J Cell Biol, 2001. **154**(1): p. 161-76.
30. Petit, V., et al., *Phosphorylation of tyrosine residues 31 and 118 on paxillin regulates cell migration through an association with CRK in NBT-II cells.* J Cell Biol, 2000. **148**(5): p. 957-70.
31. Cote, J.F., C.E. Turner, and M.L. Tremblay, *Intact LIM 3 and LIM 4 domains of paxillin are required for the association to a novel polyproline region (Pro 2) of protein-tyrosine phosphatase-PEST.* J Biol Chem, 1999. **274**(29): p. 20550-60.
32. Ito, A., et al., *A truncated isoform of the PP2A B56 subunit promotes cell motility through paxillin phosphorylation.* EMBO J, 2000. **19**(4): p. 562-71.
33. Manes, S., et al., *Concerted activity of tyrosine phosphatase SHP-2 and focal adhesion kinase in regulation of cell motility.* Mol Cell Biol, 1999. **19**(4): p. 3125-35.
34. Young, M.R., S.W. Liu, and J. Meisinger, *Protein phosphatase-2A restricts migration of Lewis lung carcinoma cells by modulating the phosphorylation of focal adhesion proteins.* Int J Cancer, 2003. **103**(1): p. 38-44.
35. Hirata, T., et al., *Amplification, up-regulation and over-expression of C3G (CRK SH3 domain-binding guanine nucleotide-releasing factor) in non-small cell lung cancers.* J Hum Genet, 2004. **49**(6): p. 290-5.
36. Radha, V., et al., *The guanine nucleotide exchange factor, C3G regulates differentiation and survival of human neuroblastoma cells.* J Neurochem, 2008. **107**(5): p. 1424-35.

37. Gotoh, T., et al., *Identification of Rap1 as a Target for the Crk Sh3 Domain-Binding Guanine Nucleotide-Releasing Factor C3g.* Molecular and Cellular Biology, 1995. **15**(12): p. 6746-6753.
38. Ichiba, T., et al., *Enhancement of guanine-nucleotide exchange activity of C3G for Rap1 by the expression of Crk, CrkL, and Grb2.* J Biol Chem, 1997. **272**(35): p. 22215-20.
39. Radha, V., A. Rajanna, and G. Swarup, *Phosphorylated guanine nucleotide exchange factor C3G, induced by pervanadate and Src family kinases localizes to the Golgi and subcortical actin cytoskeleton.* BMC Cell Biol, 2004. **5**: p. 31.
40. Mitra, A. and V. Radha, *F-actin-binding domain of c-Abl regulates localized phosphorylation of C3G: role of C3G in c-Abl-mediated cell death.* Oncogene, 2010. **29**(32): p. 4528-42.
41. Sakakibara, A., et al., *Novel function of Chat in controlling cell adhesion via Cas-Crk-C3G-pathway-mediated Rap1 activation.* J Cell Sci, 2002. **115**(Pt 24): p. 4915-24.
42. Vuori, K., et al., *Introduction of p130cas signaling complex formation upon integrin-mediated cell adhesion: a role for Src family kinases.* Mol Cell Biol, 1996. **16**(6): p. 2606-13.
43. Cote, J.F. and K. Vuori, *Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity.* J Cell Sci, 2002. **115**(Pt 24): p. 4901-13.
44. Brugnera, E., et al., *Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex.* Nat Cell Biol, 2002. **4**(8): p. 574-82.
45. Grimsley, C.M., et al., *Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration.* J Biol Chem, 2004. **279**(7): p. 6087-97.
46. Posern, G., et al., *Development of highly selective SH3 binding peptides for Crk and CRKL which disrupt Crk-complexes with DOCK180, SoS and C3G.* Oncogene, 1998. **16**(15): p. 1903-12.
47. Valles, A.M., M. Beuvin, and B. Boyer, *Activation of Rac1 by paxillin-Crk-DOCK180 signaling complex is antagonized by Rap1 in migrating NBT-II cells.* J Biol Chem, 2004. **279**(43): p. 44490-6.
48. Kiyokawa, E., et al., *Activation of Rac1 by a Crk SH3-binding protein, DOCK180.* Genes Dev, 1998. **12**(21): p. 3331-6.
49. Akakura, S., et al., *C-terminal SH3 domain of CrkII regulates the assembly and function of the DOCK180/ELMO Rac-GEF.* J Cell Physiol, 2005. **204**(1): p. 344-51.
50. Feng, H., et al., *Phosphorylation of dedicator of cytokinesis 1 (Dock180) at tyrosine residue Y722 by Src family kinases mediates EGFRvIII-driven glioblastoma tumorigenesis.* Proc Natl Acad Sci U S A, 2012. **109**(8): p. 3018-23.
51. Reynolds, A.B., et al., *Stable association of activated pp60src with two tyrosine-phosphorylated cellular proteins.* Mol Cell Biol, 1989. **9**(9): p. 3951-8.
52. Kanner, S.B., et al., *The SH2 and SH3 domains of pp60src direct stable association with tyrosine phosphorylated proteins p130 and p110.* EMBO J, 1991. **10**(7): p. 1689-98.
53. Songyang, Z. and L.C. Cantley, *The use of peptide library for the determination of kinase peptide substrates.* Methods Mol Biol, 1998. **87**: p. 87-98.
54. Bouton, A.H., R.B. Riggins, and P.J. Bruce-Staskal, *Functions of the adapter protein Cas: signal convergence and the determination of cellular responses.* Oncogene, 2001. **20**(44): p. 6448-58.
55. Sawada, Y., et al., *Force sensing by mechanical extension of the Src family kinase substrate p130Cas.* Cell, 2006. **127**(5): p. 1015-26.

56. Pellicena, P. and W.T. Miller, *Processive phosphorylation of p130Cas by Src depends on SH3-polyproline interactions.* J Biol Chem, 2001. **276**(30): p. 28190-6.
57. Klemke, R.L., et al., *CAS/Crk coupling serves as a "molecular switch" for induction of cell migration.* J Cell Biol, 1998. **140**(4): p. 961-72.
58. Auvinen, M., et al., *Ornithine decarboxylase- and ras-induced cell transformations: reversal by protein tyrosine kinase inhibitors and role of pp130CAS.* Mol Cell Biol, 1995. **15**(12): p. 6513-25.
59. Schrecengost, R.S., et al., *Breast cancer antiestrogen resistance-3 expression regulates breast cancer cell migration through promotion of p130Cas membrane localization and membrane ruffling.* Cancer Res, 2007. **67**(13): p. 6174-82.
60. Wong, S. and O.N. Witte, *The BCR-ABL story: bench to bedside and back.* Annu Rev Immunol, 2004. **22**: p. 247-306.
61. Rosenberg, N. and D. Baltimore, *The effect of helper virus on Abelson virus-induced transformation of lymphoid cells.* J Exp Med, 1978. **147**(4): p. 1126-41.
62. Schwartzberg, P.L., et al., *Mice homozygous for the ablm1 mutation show poor viability and depletion of selected B and T cell populations.* Cell, 1991. **65**(7): p. 1165-75.
63. Advani, A.S. and A.M. Pendergast, *Bcr-Abl variants: biological and clinical aspects.* Leuk Res, 2002. **26**(8): p. 713-20.
64. Lugo, T.G. and O.N. Witte, *The BCR-ABL oncogene transforms Rat-1 cells and cooperates with v-myc.* Mol Cell Biol, 1989. **9**(3): p. 1263-70.
65. McLaughlin, J., E. Chianese, and O.N. Witte, *In vitro transformation of immature hematopoietic cells by the P210 BCR/ABL oncogene product of the Philadelphia chromosome.* Proc Natl Acad Sci U S A, 1987. **84**(18): p. 6558-62.
66. Hantschel, O. and G. Superti-Furga, *Regulation of the C-Abl and Bcr-Abl tyrosine kinases.* Nature Reviews Molecular Cell Biology, 2004. **5**(1): p. 33-44.
67. Nagar, B., et al., *Structural basis for the autoinhibition of c-Abl tyrosine kinase.* Cell, 2003. **112**(6): p. 859-871.
68. Choi, Y., et al., *N-myristoylated c-Abl tyrosine kinase localizes to the endoplasmic reticulum upon binding to an allosteric inhibitor.* J Biol Chem, 2009. **284**(42): p. 29005-14.
69. Franz, W.M., P. Berger, and J.Y. Wang, *Deletion of an N-terminal regulatory domain of the c-abl tyrosine kinase activates its oncogenic potential.* EMBO J, 1989. **8**(1): p. 137-47.
70. Hantschel, O., et al., *A myristoyl/phosphotyrosine switch regulates c-Abl.* Cell, 2003. **112**(6): p. 845-57.
71. Ren, R., Z.S. Ye, and D. Baltimore, *Abl protein-tyrosine kinase selects the Crk adapter as a substrate using SH3-binding sites.* Genes Dev, 1994. **8**(7): p. 783-95.
72. Shishido, T., et al., *Crk family adaptor proteins trans-activate c-Abl kinase.* Genes Cells, 2001. **6**(5): p. 431-40.
73. Reichman, C., et al., *Transactivation of Abl by the Crk II adapter protein requires a PNAY sequence in the Crk C-terminal SH3 domain.* Oncogene, 2005. **24**(55): p. 8187-99.
74. Chodniewicz, D. and R.L. Klemke, *Regulation of integrin-mediated cellular responses through assembly of a CAS/Crk scaffold.* Biochimica Et Biophysica Acta-Molecular Cell Research, 2004. **1692**(2-3): p. 63-76.
75. Huang, X., et al., *Induction of cell retraction by the combined actions of Abl-CrkII and Rho-ROCK1 signaling.* J Cell Biol, 2008. **183**(4): p. 711-23.
76. Escalante, M., et al., *Phosphorylation of c-Crk II on the negative regulatory Tyr222 mediates nerve growth factor-induced cell spreading and morphogenesis.* J Biol Chem, 2000. **275**(32): p. 24787-97.

77. Kain, K.H., S. Gooch, and R.L. Klemke, *Cytoplasmic c-Abl provides a molecular 'Rheostat' controlling carcinoma cell survival and invasion.* Oncogene, 2003. **22**(38): p. 6071-80.
78. Cho, S.Y. and R.L. Klemke, *Extracellular-regulated kinase activation and CAS/Crk coupling regulate cell migration and suppress apoptosis during invasion of the extracellular matrix.* J Cell Biol, 2000. **149**(1): p. 223-36.
79. Noren, N.K., et al., *The EphB4 receptor suppresses breast cancer cell tumorigenicity through an Abl-Crk pathway.* Nat Cell Biol, 2006. **8**(8): p. 815-25.
80. Senechal, K., et al., *Structural requirements for function of the Crkl adapter protein in fibroblasts and hematopoietic cells.* Mol Cell Biol, 1998. **18**(9): p. 5082-90.
81. Feller, S.M., B. Knudsen, and H. Hanafusa, *C-Abl Kinase Regulates the Protein-Binding Activity of C-Crk.* Embo Journal, 1994. **13**(10): p. 2341-2351.
82. Rosen, M.K., et al., *Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein.* Nature, 1995. **374**(6521): p. 477-9.
83. Muralidharan, V., et al., *Solution structure and folding characteristics of the C-terminal SH3 domain of c-Crk-II.* Biochemistry, 2006. **45**(29): p. 8874-84.
84. Zvara, A., et al., *Activation of the focal adhesion kinase signaling pathway by structural alterations in the carboxyl-terminal region of c-Crk II.* Oncogene, 2001. **20**(8): p. 951-61.
85. Sarkar, P., et al., *Proline cis-trans isomerization controls autoinhibition of a signaling protein.* Mol Cell, 2007. **25**(3): p. 413-26.
86. Cho, J.H., et al., *Tuning protein autoinhibition by domain destabilization.* Nat Struct Mol Biol, 2011. **18**(5): p. 550-5.
87. la Cour, T., et al., *Analysis and prediction of leucine-rich nuclear export signals.* Protein Eng Des Sel, 2004. **17**(6): p. 527-36.
88. Smith, J.J., et al., *Apoptotic regulation by the Crk adapter protein mediated by interactions with Wee1 and Crm1/exportin.* Mol Cell Biol, 2002. **22**(5): p. 1412-23.
89. Harkiolaki, M., et al., *The C-terminal SH3 domain of CRKL as a dynamic dimerization module transiently exposing a nuclear export signal.* Structure, 2006. **14**(12): p. 1741-53.
90. Sriram, G., et al., *Phosphorylation of Crk on tyrosine 251 in the RT loop of the SH3C domain promotes Abl kinase transactivation.* Oncogene, 2011. **30**(46): p. 4645-55.
91. Nishihara, H., et al., *DOCK2 associates with CrkL and regulates Rac1 in human leukemia cell lines.* Blood, 2002. **100**(12): p. 3968-74.
92. Miller, C.T., et al., *Increased C-CRK proto-oncogene expression is associated with an aggressive phenotype in lung adenocarcinomas.* Oncogene, 2003. **22**(39): p. 7950-7.
93. Takino, T., et al., *Tyrosine phosphorylation of the CrkII adaptor protein modulates cell migration.* J Cell Sci, 2003. **116**(Pt 15): p. 3145-55.
94. Crawford, M., et al., *MicroRNA-126 inhibits invasion in non-small cell lung carcinoma cell lines.* Biochem Biophys Res Commun, 2008. **373**(4): p. 607-12.
95. Sattler, M., et al., *The proto-oncogene product p120CBL and the adaptor proteins CRKL and c-CRK link c-ABL, p190BCR/ABL and p210BCR/ABL to the phosphatidylinositol-3' kinase pathway.* Oncogene, 1996. **12**(4): p. 839-46.
96. Deininger, K., et al., *The Rab5 guanylate exchange factor Rin1 regulates endocytosis of the EphA4 receptor in mature excitatory neurons.* Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12539-44.
97. Oda, T., et al., *Crkl is the major tyrosine-phosphorylated protein in neutrophils from patients with chronic myelogenous leukemia.* J Biol Chem, 1994. **269**(37): p. 22925-8.
98. Salgia, R., et al., *CRKL links p210BCR/ABL with paxillin in chronic myelogenous leukemia cells.* J Biol Chem, 1995. **270**(49): p. 29145-50.

99. Hamilton, A., et al., *Optimization of methods for the detection of BCR-ABL activity in Philadelphia-positive cells.* Exp Hematol, 2009. **37**(3): p. 395-401.
100. Burton, E.A., R. Plattner, and A.M. Pendergast, *Abl tyrosine kinases are required for infection by Shigella flexneri.* EMBO J, 2003. **22**(20): p. 5471-9.
101. Vepachedu, R., et al., *Unc119 protects from Shigella infection by inhibiting the Abl family kinases.* PLoS One, 2009. **4**(4): p. e5211.
102. Deng, Q., J. Sun, and J.T. Barbieri, *Uncoupling Crk signal transduction by Pseudomonas exoenzyme T.* J Biol Chem, 2005. **280**(43): p. 35953-60.
103. Suzuki, M., et al., *Interaction of CagA with Crk plays an important role in Helicobacter pylori-induced loss of gastric epithelial cell adhesion.* J Exp Med, 2005. **202**(9): p. 1235-47.
104. Heikkinen, L.S., et al., *Avian and 1918 Spanish influenza a virus NS1 proteins bind to Crk/CrkL Src homology 3 domains to activate host cell signaling.* J Biol Chem, 2008. **283**(9): p. 5719-27.
105. Hrincius, E.R., et al., *CRK adaptor protein expression is required for efficient replication of avian influenza A viruses and controls JNK-mediated apoptotic responses.* Cell Microbiol, 2010. **12**(6): p. 831-43.
106. Grzesiek, S. and A. Bax, *An Efficient Experiment for Sequential Backbone Assignment of Medium-Sized Isotopically Enriched Proteins.* Journal of Magnetic Resonance, 1992. **99**(1): p. 201-207.
107. Kay, L.E., et al., *3-Dimensional Triple-Resonance Nmr-Spectroscopy of Isotopically Enriched Proteins.* Journal of Magnetic Resonance, 1990. **89**(3): p. 496-514.
108. Marintchev, A., D. Frueh, and G. Wagner, *NMR methods for studying protein-protein interactions involved in translation initiation.* Methods Enzymol, 2007. **430**: p. 283-331.
109. Grzesiek, S. and A. Bax, *Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins.* J Biomol NMR, 1993. **3**(2): p. 185-204.
110. Lin, Y. and G. Wagner, *Efficient side-chain and backbone assignment in large proteins: application to tGCN5.* J Biomol NMR, 1999. **15**(3): p. 227-39.
111. Gaetano T. Montelione, B.A.L., S. Donald Emerson, Mitsuru Tashiro, *An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins.* J. Am. Chem. Soc, 1992. **114**(27): p. 10974-10975.
112. Wuthrich, K., *NMR studies of structure and function of biological macromolecules (Nobel Lecture).* J Biomol NMR, 2003. **27**(1): p. 13-39.
113. Marion, D., et al., *Overcoming the overlap problem in the assignment of 1H NMR spectra of larger proteins by use of three-dimensional heteronuclear 1H-15N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta.* Biochemistry, 1989. **28**(15): p. 6150-6.
114. Zuiderweg, E.R. and S.W. Fesik, *Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a.* Biochemistry, 1989. **28**(6): p. 2387-91.
115. Marion, D., Lewis E. Kay, Steven W. Sparks, Dennis A. Torchia, Ad Bax, *Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins.* J. Am. Chem. Soc., 1989. **111**(4): p. 1515-1517.
116. Wand, A.J., M.R. Ehrhardt, and P.F. Flynn, *High-resolution NMR of encapsulated proteins dissolved in low-viscosity fluids.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15299-302.
117. Sattler, M. and S.W. Fesik, *Use of deuterium labeling in NMR: overcoming a sizeable problem.* Structure, 1996. **4**(11): p. 1245-9.

118. Gardner, K.H. and L.E. Kay, *The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins.* Annu Rev Biophys Biomol Struct, 1998. **27**: p. 357-406.
119. Kanelis, V., J.D. Forman-Kay, and L.E. Kay, *Multidimensional NMR methods for protein structure determination.* IUBMB Life, 2001. **52**(6): p. 291-302.
120. Clore, G.M. and J. Iwahara, *Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes.* Chem Rev, 2009. **109**(9): p. 4108-39.
121. Clore, G.M., C. Tang, and J. Iwahara, *Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement.* Curr Opin Struct Biol, 2007. **17**(5): p. 603-16.
122. Iwahara, J., C. Tang, and G. Marius Clore, *Practical aspects of (1)H transverse paramagnetic relaxation enhancement measurements on macromolecules.* J Magn Reson, 2007. **184**(2): p. 185-95.
123. Battiste, J.L. and G. Wagner, *Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data.* Biochemistry, 2000. **39**(18): p. 5355-65.
124. Bertoncini, C.W., et al., *Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein.* Proc Natl Acad Sci U S A, 2005. **102**(5): p. 1430-5.
125. Fawzi, N.L., et al., *Mechanistic details of a protein-protein association pathway revealed by paramagnetic relaxation enhancement titration measurements.* Proc Natl Acad Sci U S A, 2010. **107**(4): p. 1379-84.
126. Popovych, N., et al., *Structural basis for cAMP-mediated allosteric control of the catabolite activator protein.* Proc Natl Acad Sci U S A, 2009. **106**(17): p. 6927-32.
127. Jennings-Antipov, L.D., L. Song, and R.J. Collier, *Interactions of anthrax lethal factor with protective antigen defined by site-directed spin labeling.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 1868-73.
128. Cornilescu, G., F. Delaglio, and A. Bax, *Protein backbone angle restraints from searching a database for chemical shift and sequence homology.* J Biomol NMR, 1999. **13**(3): p. 289-302.
129. Shen, Y., et al., *TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.* J Biomol NMR, 2009. **44**(4): p. 213-23.
130. Lipsitz, R.S. and N. Tjandra, *Residual dipolar couplings in NMR structure analysis.* Annu Rev Biophys Biomol Struct, 2004. **33**: p. 387-413.
131. Tjandra, N. and A. Bax, *Measurement of dipolar contributions to 1JCH splittings from magnetic-field dependence of J modulation in two-dimensional NMR spectra.* J Magn Reson, 1997. **124**(2): p. 512-5.
132. Hansen, M.R., L. Mueller, and A. Pardi, *Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions.* Nat Struct Biol, 1998. **5**(12): p. 1065-74.
133. Ruckert, M., *Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments.* J Am Chem Soc, 2000. **122**(32): p. 7793-7797.
134. Losonczi, J.A. and J.H. Prestegard, *Improved dilute bicelle solutions for high-resolution NMR of biological macromolecules.* J Biomol NMR, 1998. **12**(3): p. 447-51.

135. Ottiger, M., F. Delaglio, and A. Bax, *Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra.* J Magn Reson, 1998. **131**(2): p. 373-8.
136. Sattler, M., *Introduction to biomolecular NMR spectroscopy.* 2004(EMBL Heidelberg).
137. Schwieters, C.D., J.A. Alford, and J.B. Delos, *Semiclassical scattering in a circular semiconductor microstructure.* Phys Rev B Condens Matter, 1996. **54**(15): p. 10652-10668.
138. Mittermaier, A. and L.E. Kay, *New tools provide new insights in NMR studies of protein dynamics.* Science, 2006. **312**(5771): p. 224-8.
139. Ishima, R. and D.A. Torchia, *Protein dynamics from NMR.* Nat Struct Biol, 2000. **7**(9): p. 740-3.
140. Akke, M., *NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis.* Curr Opin Struct Biol, 2002. **12**(5): p. 642-7.
141. Zuiderweg, E.R., *Mapping protein-protein interactions in solution by NMR spectroscopy.* Biochemistry, 2002. **41**(1): p. 1-7.
142. Takeuchi, K. and G. Wagner, *NMR studies of protein interactions.* Curr Opin Struct Biol, 2006. **16**(1): p. 109-17.
143. Nakanishi, T., et al., *Determination of the interface of a large protein complex by transferred cross-saturation measurements.* J Mol Biol, 2002. **318**(2): p. 245-9.
144. Ikura, M., et al., *Solution structure of calmodulin and its complex with a myosin light chain kinase fragment.* Cell Calcium, 1992. **13**(6-7): p. 391-400.
145. Vaynberg, J. and J. Qin, *Weak protein-protein interactions as probed by NMR spectroscopy.* Trends Biotechnol, 2006. **24**(1): p. 22-7.
146. Gordon, G.W., et al., *Quantitative fluorescence resonance energy transfer measurements using fluorescence microscopy.* Biophys J, 1998. **74**(5): p. 2702-13.
147. Weiss, S., *Fluorescence spectroscopy of single biomolecules.* Science (New York, N Y), 1999. **283**(5408): p. 1676-83.
148. Adams, S.R., et al., *Fluorescence ratio imaging of cyclic AMP in single cells.* Nature, 1991. **349**(6311): p. 694-7.
149. Nagy, P., et al., *Intensity-based energy transfer measurements in digital imaging microscopy.* Eur Biophys J, 1998. **27**(4): p. 377-89.
150. Zimmermann, T. and F. Siegert, *4D confocal microscopy of Dictyostelium discoideum morphogenesis and its presentation on the Internet.* Dev Genes Evol, 1998. **208**(7): p. 411-20.
151. Perozzo, R., G. Folkers, and L. Scapozza, *Thermodynamics of protein-ligand interactions: history, presence, and future aspects.* J Recept Signal Transduct Res, 2004. **24**(1-2): p. 1-52.
152. Jelesarov, I. and H.R. Bosshard, *Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition.* J Mol Recognit, 1999. **12**(1): p. 3-18.
153. Luque, I. and E. Freire, *Structure-based prediction of binding affinities and molecular design of peptide ligands.* Methods Enzymol, 1998. **295**: p. 100-27.
154. Guris, D.L., et al., *Dose-dependent interaction of Tbx1 and Crkl and locally aberrant RA signaling in a model of del22q11 syndrome.* Dev Cell, 2006. **10**(1): p. 81-92.
155. Moon, A.M., et al., *Crkl deficiency disrupts Fgf8 signaling in a mouse model of 22q11 deletion syndromes.* Dev Cell, 2006. **10**(1): p. 71-80.
156. Park, T.J., K. Boyd, and T. Curran, *Cardiovascular and craniofacial defects in Crk-null mice.* Mol Cell Biol, 2006. **26**(16): p. 6272-82.

157. Feng, L. and J.A. Cooper, *Dual functions of Dab1 during brain development.* Mol Cell Biol, 2009. **29**(2): p. 324-32.
158. Park, T.J. and T. Curran, *Crk and Crk-like play essential overlapping roles downstream of disabled-1 in the Reelin pathway.* J Neurosci, 2008. **28**(50): p. 13551-62.
159. Seo, J.H., et al., *Structural and functional basis of a role for CRKL in a fibroblast growth factor 8-induced feed-forward loop.* Mol Cell Biol, 2009. **29**(11): p. 3076-87.
160. Kobashigawa, Y., et al., *Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK.* Nat Struct Mol Biol, 2007. **14**(6): p. 503-10.
161. Donaldson, L.W., et al., *Structure of a regulatory complex involving the Abl SH3 domain, the Crk SH2 domain, and a Crk-derived phosphopeptide.* Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14053-8.
162. Tjandra, N., et al., *Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation.* Journal of the American Chemical Society, 1995. **117**(50): p. 12562-12566.
163. Sarkar, P., et al., *Structural basis for regulation of the Crk signaling protein by a proline switch.* Nat Chem Biol, 2011. **7**(1): p. 51-7.
164. Felschow, D.M., et al., *The adapter protein CrkL associates with CD34.* Blood, 2001. **97**(12): p. 3768-75.
165. Mintz, P.J., et al., *An unrecognized extracellular function for an intracellular adapter protein released from the cytoplasm into the tumor microenvironment.* Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2182-7.
166. Senechal, K., J. Halpern, and C.L. Sawyers, *The CRKL adaptor protein transforms fibroblasts and functions in transformation by the BCR-ABL oncogene.* J Biol Chem, 1996. **271**(38): p. 23255-61.
167. Fischer, G. and T. Aumuller, *Regulation of peptide bond cis/trans isomerization by enzyme catalysis and its implication in physiological processes.* Rev Physiol Biochem Pharmacol, 2003. **148**: p. 105-50.
168. Dalgarno, D.C., et al., *1H NMR studies on bovine cyclophilin: preliminary structural characterization of this specific cyclosporin A binding protein.* Biochemistry, 1986. **25**(22): p. 6778-84.
169. Walsh, C.T., L.D. Zydowsky, and F.D. McKeon, *Cyclosporin A, the cyclophilin class of peptidylprolyl isomerases, and blockade of T cell signal transduction.* J Biol Chem, 1992. **267**(19): p. 13115-8.
170. Liu, J., et al., *Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes.* Cell, 1991. **66**(4): p. 807-15.
171. Lee, J. and S.S. Kim, *An overview of cyclophilins in human cancers.* J Int Med Res, 2010. **38**(5): p. 1561-74.
172. Ke, H., D. Mayrose, and W. Cao, *Crystal structure of cyclophilin A complexed with substrate Ala-Pro suggests a solvent-assisted mechanism of cis-trans isomerization.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3324-8.
173. Jin, Z.G. and B.C. Berk, *Role of secreted oxidative stress-induced factors (SOXFs) in the pathogenesis of atherosclerosis.* Arch Mal Coeur Vaiss, 2004. **97**(12): p. 1256-9.
174. Smith, J.L., et al., *Structural heterogeneity in protein crystals.* Biochemistry, 1986. **25**(18): p. 5018-27.
175. Wohlfarth, C. and T. Efferth, *Natural products as promising drug candidates for the treatment of hepatitis B and C.* Acta Pharmacol Sin, 2009. **30**(1): p. 25-30.
176. Towers, G.J., et al., *Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors.* Nat Med, 2003. **9**(9): p. 1138-43.

177. Satoh, K., H. Shimokawa, and B.C. Berk, *Cyclophilin A: promising new target in cardiovascular therapy.* Circ J, 2010. **74**(11): p. 2249-56.
178. Satoh, K., P. Nigro, and B.C. Berk, *Oxidative stress and vascular smooth muscle cell growth: a mechanistic linkage by cyclophilin A.* Antioxid Redox Signal, 2010. **12**(5): p. 675-82.
179. Campa, M.J., et al., *Protein expression profiling identifies macrophage migration inhibitory factor and cyclophilin a as potential molecular targets in non-small cell lung cancer.* Cancer Res, 2003. **63**(7): p. 1652-6.
180. Howard, B.A., et al., *Translating biomarkers into clinical practice: prognostic implications of cyclophilin A and macrophage migratory inhibitory factor identified from protein expression profiles in non-small cell lung cancer.* Lung Cancer, 2004. **46**(3): p. 313-323.
181. Howard, B.A., et al., *Stable RNA interference-mediated suppression of cyclophilin A diminishes non-small-cell lung tumor growth in vivo.* Cancer Res, 2005. **65**(19): p. 8853-60.
182. Li, M., et al., *Effect of cyclophilin A on gene expression in human pancreatic cancer cells.* Am J Surg, 2005. **190**(5): p. 739-45.
183. Li, M., et al., *Cyclophilin A is overexpressed in human pancreatic cancer cells and stimulates cell proliferation through CD147.* Cancer, 2006. **106**(10): p. 2284-94.
184. Mikuriya, K., et al., *Expression of glycolytic enzymes is increased in pancreatic cancerous tissues as evidenced by proteomic profiling by two-dimensional electrophoresis and liquid chromatography-mass spectrometry/mass spectrometry.* Int J Oncol, 2007. **30**(4): p. 849-55.
185. Zheng, J., et al., *Prolyl isomerase cyclophilin A regulation of Janus-activated kinase 2 and the progression of human breast cancer.* Cancer Res, 2008. **68**(19): p. 7769-78.
186. Hathout, Y., et al., *Differential protein expression in the cytosol fraction of an MCF-7 breast cancer cell line selected for resistance toward melphalan.* J Proteome Res, 2002. **1**(5): p. 435-42.
187. Lou, J., et al., *Proteomic profiling identifies cyclooxygenase-2-independent global proteomic changes by celecoxib in colorectal cancer cells.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(9): p. 1598-606.
188. Melle, C., et al., *Identification of proteins from colorectal cancer tissue by two-dimensional gel electrophoresis and SELDI mass spectrometry.* Int J Mol Med, 2005. **16**(1): p. 11-7.
189. Wong, C.S., et al., *Identification of 5-fluorouracil response proteins in colorectal carcinoma cell line SW480 by two-dimensional electrophoresis and MALDI-TOF mass spectrometry.* Oncol Rep, 2008. **20**(1): p. 89-98.
190. Choi, K.J., et al., *Overexpressed cyclophilin A in cancer cells renders resistance to hypoxia- and cisplatin-induced cell death.* Cancer Res, 2007. **67**(8): p. 3654-62.
191. Schubert, M., et al., *A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on 13C chemical shift statistics.* J Biomol NMR, 2002. **24**(2): p. 149-54.
192. Nicholson, L.K. and K.P. Lu, *Prolyl cis-trans Isomerization as a molecular timer in Crk signaling.* Mol Cell, 2007. **25**(4): p. 483-5.
193. Piotukh, K., et al., *Cyclophilin A binds to linear peptide motifs containing a consensus that is present in many human proteins.* J Biol Chem, 2005. **280**(25): p. 23668-74.
194. Bosco, D.A., et al., *Catalysis of cis/trans isomerization in native HIV-1 capsid by human cyclophilin A.* Proc Natl Acad Sci U S A, 2002. **99**(8): p. 5247-52.

195. Theriault, Y., et al., *Solution structure of the cyclosporin A/cyclophilin complex by NMR.* Nature, 1993. **361**(6407): p. 88-91.
196. Howard, B.R., et al., *Structural insights into the catalytic mechanism of cyclophilin A.* Nat Struct Biol, 2003. **10**(6): p. 475-81.
197. Ubersax, J.A. and J.E. Ferrell, Jr., *Mechanisms of specificity in protein phosphorylation.* Nat Rev Mol Cell Biol, 2007. **8**(7): p. 530-41.
198. Luban, J., et al., *Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B.* Cell, 1993. **73**(6): p. 1067-78.
199. Thali, M., et al., *Resistance to neutralization by broadly reactive antibodies to the human immunodeficiency virus type 1 gp120 glycoprotein conferred by a gp41 amino acid change.* J Virol, 1994. **68**(2): p. 674-80.
200. Yoo, S., et al., *Molecular recognition in the HIV-1 capsid/cyclophilin A complex.* J Mol Biol, 1997. **269**(5): p. 780-95.
201. Song, F., et al., *Cyclophilin A (CyPA) induces chemotaxis independent of its peptidylprolyl cis-trans isomerase activity: direct binding between CyPA and the ectodomain of CD147.* J Biol Chem, 2011. **286**(10): p. 8197-203.
202. Seeliger, M.A., et al., *High yield bacterial expression of active c-Abl and c-Src tyrosine kinases.* Protein Sci, 2005. **14**(12): p. 3135-9.
203. Farrow, N.A., et al., *Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation.* Biochemistry, 1994. **33**(19): p. 5984-6003.
204. Bosco, D.A. and D. Kern, *Catalysis and binding of cyclophilin A with different HIV-1 capsid constructs.* Biochemistry, 2004. **43**(20): p. 6110-9.
205. Yao, L., J. Ying, and A. Bax, *Improved accuracy of 15N-1H scalar and residual dipolar couplings from gradient-enhanced IPAP-HSQC experiments on protonated proteins.* J Biomol NMR, 2009. **43**(3): p. 161-70.
206. Mandel, A.M., M. Akke, and A.G. Palmer, 3rd, *Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme.* J Mol Biol, 1995. **246**(1): p. 144-63.
207. Tang, C., J. Iwahara, and G.M. Clore, *Visualization of transient encounter complexes in protein-protein association.* Nature, 2006. **444**(7117): p. 383-6.

8 Curriculum Vitae

Tamjeed saleh

1981	**Born March 4, Dhaka Bangladesh**
1999	**Graduated from Highschool, Scholastica, June 1999**
2004	**Graduated Brown University. Sc.B Biology, A.B Economics.**
2004-2005	**Research Assistant. Dept. of Biology MIT.**
2005-2008	**Teaching Assistant Department of Biology, Rutgers Newark**
2008-2010	**Teaching Assistant, Department of Molecular Biology and Biochemisty, Rutgers New Brunswick**
2010-2012	**Graduate Assistant, Department of Molecular Biology and Biochemisty, Rutgers New Brunswick**
2012	**Ph.D in Molecular Biology and Biochemisty, Rutgers, Rutgers New Brunswick**

Publications

P. Sarkar, C. Reichman, T. Saleh, R.B. Birge & C.G. Kalodimos (2007) "Proline cis-trans isomerization controls autoinhibition of a signaling molecule" *Molecular Cell*, 25, 413-426

P. Sarkar, T. Saleh, H-S. Tzeng, R.B. Birge & C.G. Kalodimos (2011) "Structural basis for regulation of the Crk signaling protein by a proline switch" *Nature Chemical Biology*, 7, 51-57.

G Sriram, C Reichman, A Tunceroglu, N Kaushal, T Saleh, K Machida, B Mayer, Q Ge, J Li, P Hornbeck, C G Kalodimos, R B Birge (2011) "Phosphorylation of Crk on tyrosine 251 in the RT loop of the SH3C domain promotes Abl kinase transactivation." Oncogene. Nov 17;30(46):4645-55

W. Jankowski,T. Saleh, M.-T. Pai, S. Ganapathy, R. B. Birge & C.G. Kalodimos (2012) "Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII" *Nature Chemical Biology*, 8, 590-596.