Structural basis for the function of the Bcr-Abl protein substrate CrkL and human-CrkII as a novel partner for Cyclophilin A

by

WOJCIECH JANKOWSKI

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 Chemistry

Written under the direction of

Professor Charalampos G. Kalodimos

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2012

ABSTRACT OF THE DISSERTATION

Structural basis for the function of the Bcr-Abl protein substrate CrkL and human-CrkII as a novel partner for Cyclophilin A

by Wojciech Jankowski

Dissertation Director:
Professor Charalampos G. Kalodimos

Adaptor proteins are known to play an essential role in assembling protein-protein complexes which result in cellular signal propagation. Crk belongs to a family of adaptor proteins and was originally identified as an oncogene product of the CT10 retrovirus (v-Crk). Cellular homologues of v-Crk include CrkI, CrkII and CrkL. CrkI and CrkII are different splice variants, whereas CrkL is encoded by a distinct gene. Crk proteins contain one Src homology 2 (SH2) and one or two Src homology 3 (SH3) domains. Specific domain organization can assemble and activate a number of different ligands, including Abl. It remains poorly understood why CrkII and CrkL have distinct physiological roles despite showing similar domain structures, high sequence identity, and identical binding partners. Unlike CrkII, CrkL was found to be a key signaling molecule to interact with Bcr-Abl, which is a tyrosine kinase that plays a major role in Chronic Myeloid Leukemia (CML) pathogenesis. The interaction of CrkL with Bcr-Abl and its potent tyrosine phosphorylation (higher as compared to CrkII) is commonly used as a hallmark of Bcr-Abl kinase activity and response to tyrosine kinase inhibitors used in CML

treatment. Knowing the differences in 3-dimensional structures between CrkII and CrkL would help to understand how these adaptors alter key signaling partners. However, the structure of CrkL is not available.

Using NMR spectroscopy methodologies complemented by many other biochemical and biophysical techniques, we show that CrkL and phosphorylated CrkL structures are radically different from the corresponding structures of CrkII. The phosphorylation of Tyr221 (CrkII) and Tyr207 (CrkL) by Abl induces intramolecular binding to the SH2 domain, which in the case of phosphorylated CrkII was shown to completely abrogate signal transduction. In phosphorylated CrkL, however, the $SH3^N$ domain remains accessible and can form complexes. The data show that CrkL, unlike CrkII, forms a constitutive complex with Abl hence explaining the preference of Bcr-Abl for CrkL over CrkII. The results also highlight how the structural organization of the modular domains in adaptor proteins can control signaling outcome.

In the second part we show that the Gly219−Pro220 motif of CrkII binds into the active site cleft of CypA. In contrast, CrkL does not contain this GP motif and therefore, is not susceptible to CypA regulation. The interaction between CypA and CrkII occurs both *in vitro* and *in vivo*. CypA is recruited to the CrkII phosphorylation site (Tyr221), and delays phosphorylation by Abl. 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. David A. Case

Prof. Wilma K. Olson

Prof. Raymond Birge

Tamjeed Saleh

Sriram Genapathy

Lab Members, Current and Past

Dr Xuinjun Ai, Dr Akash Bhattacharya, Dr Anusarka Bhaumik, Dr Chih-Hong Chen, Dr Li Chen, Dr Yiannis Gelis, Dr Soumyasri Das Gupta, Dr Xiao Guan, Dr Dimitra Keramisanou, Nandish Kharna, Soma Mandal, Dr. Mandar Naik, Dr Ming-Tao Pai, Dr Krishna Mohan Poluri, Dr Tomohide Saio, Dr Paramita Sarkar, Pragati Sharma, Dr Hsiu-Ru Tseng

NMR managers (New Brunswick):

Dr Seho Kim and Dr Nagarajan Murali

Department of Chemistry and Chemical Biology Faculty and Staff

Biomedical Engineering Faculty and Staff

Family and Friends

Table of Contents

List of abbreviations

3BP-1	SH3 binding protein
3D	three-dimensional
Abl	Abelson kinase
Abl-1a	isoform 1a of Abl
Abl-1b	isoform 1b of Abl
Abl^{KD}	Abl kinase domain
ADP	adenosine diphosphate
Ala	alanine (A)
A-MuLV	Abelson murine leukaemia virus
Arg	Abl related gene -Abl2
Arg	arginine (R)
Asn	asparagine (N)
Asp	aspartic acid (D)
ATP	adenosine triphosphate
AVS-1	avian sarcoma virus 1
Bcr	breakpoint cluster region protein
Bcr-Abl	fused BCR-Abl protein
BME	β-mercaptoethanol
B-Raf	v-raf murine sarcoma viral oncogene homolog B1
c-	cellular
C3G	guanine nucleotide-releasing factor
c-Abl	cellular Abelson tyrosine kinase- Abl1
CagA	virulence factor of *Helicobacter pylori*
CAS or Cas	protein called Crk-associated substrate
CBR	Crk binding region
cCrk	cellular Crk
CD147	membrane receptor
CDC25	cell division cycle 25
CFP	cyan fluorescent *protein*

CML	chronic myeloid leukemia
CML	myelogenous leukemia
CPMG	Carr-Purcell-Meiboom-Gill methods
Crk	chicken tumor virus no. 10 regulator of kinase
CrkL	Crk-like
CRKL	gene encoded Crk-like protein
Crm1	an export receptor for leucine-rich nuclear export signals
CsA	cyclosprin A
CT10	chicken tumor virus no. 10
CypA	cyclophilin A
Cyps	cyclophilins
Cys	cysteine (C)
D_{AB}	dipolar coupling in solution,
Dbl	oncogene isolated from a human diffuse B-cell lymphoma
DH	Dbl homology
DHR-1	DOCK Homology Region-1
DHR-2	DOCK Homology Region-2
DNA	deoxyribonucleic acid
DOCK 180	180-kDa protein downstream of Crk
DOCK1	dedicator of cytokinesis 1
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
ELMO	engulfment and cell motility
EphB4	ephrin type-B receptor 4
ERBB	epidermal growth factor receptor
ERKs	extracellular-signal-regulated kinases
FAK	focal adhesion kinase
Fyn	tyrosine-specific phospho-transferase
FGFR	fibroblast growth factor receptor
FKBPs	FK506-binding proteins
FRET	Förster resonance energy transfer

FRNK	FAK-related-non-kinase
Gab1	Grb2-associated binder
gag	group-specific antigen (*protein)*
GDP	guanosine diphosphate
GEFs	guanine-nucleotide-exchange factors
Gln	glutamine (Q)
Glu	glutamic acid (E)
Gly	glycine (G)
GP	glycoprotein
Grb2	growth-factor-receptor-bound 2
GST	glutathione S-transferase
GTP	guanosine trihosphate
h	Planck's constant
h-	human
HetNOE	heteronuclear ^{15}N-nuclear Overhouse effect
His	histidine (H)
HIV	human immunodeficiency virus
HRP	horseradish peroxidase
HSC-3	head and neck squamous cell carcinoma cell line
hSOS	human homologs of SOS
HSQC	heteronuclear single quantum coherence
Ile	isoleucine (I)
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISC	Inter-SH3 core
ITC	isothermal titration calorimetry
Itk	IL2-inducible T-cell *kinase*
JNK	c-Jun N-terminal kinases
Ka	association constant
LB	lysogeny broth
Leu	leucine (L)
LIM	protein structural domain

Lys	lysine (K)
MALDI	matrix-assisted laser desorption/ionization
MALDI-TOF	matrix-assisted laser desorption/ionization-Time Of Flight
MALLS	multi-angle laser light scattering
MAPK	mitogen-activated protein kinase
MCAS	mucinous cystadenocarcinoma
Met	methionine (M)
N	stoichiometry
Nck	non-catalytic region of tyrosine kinase adaptor protein 1
NcoI	restriction enzyme
NES	nuclear export sequence
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NSCLC	*non-small-cell lung carcinoma*
P- GST	phosphorylated GST
p130Cas	protein called Crk-associated substrate
PBS	phosphate buffered saline
pCrk	phosphorylated Crk
PEG	polyethylene glycol
PEST	C-terminal PEST motif in PTP
PH	Pleckstrin homology
Phe	phenylalanine (F)
PI3K	Phosphatidylinositol 3-kinases
PLCγ	phospholipase C-γ
PP2A	Ser/Thr phosphatase 2A
PPIase	peptidyl−prolylcis−trans isomerase
PPII	polyproline type II
PPII	polyproline Type II
PRE	paramagnetic relaxation enhancement
PRE	paramagnetic relaxation enhancement

Pro	proline (P)
PTB	pTyr binding
PTKs	protein tyrosine kinases
PTP-1B	protein-tyrosine phosphatase 1B
PTPA	PP2A activator
PTPs	protein tyrosine phosphatases
PxxP	proline rich motif
Pyk2	nonreceptor tyrosine kinase of the Fak family
r	the time-averaged internuclear distance
R2	transverse relaxation rate
R7	central photoreceptor
Rac	Rho family of GTPases
Raftk	related adhesion focal tyrosine kinase
Rap	GTP-binding protein
RDC	residual dipolar couplings
RDC	residual dipolar coupling
REM	Ras exchanger motif
R_{ex}	exchange rate
RNA	ribonucleic acid
S	the generalized order parameter
S^2	measures the magnitude of the angular fluctuation of a chemical bond vector
SAP	Serum amyloid P component
SDS	sodium dodecyl sulfate
Ser	serine (S)
SH2	Src Homology 2 domain
SH3	Src Homology 3 domain
$SH3^C$	C-terminus SH3
$SH3^N$	N-terminus SH3
SHIPTP2	adaptor protein in platelet-derived growth factor receptor signaling

Socs2	suppressor of cytokine signaling 2
Socs4	suppressor of cytokine signaling 4
SOS	protein called Son of Sevenless
Src	proto-oncogene tyrosine-protein kinase
T	temperature
T2	transverse relaxation
TBST	mixture of Tris-Buffered Saline and Tween 20
TCEP	tris(2-carboxyethyl)phosphine
Tel-Abl	an oncogen which encodes an ETS family transcription factor (ETV6)
Tev	Tobacco Etch Virus
Thr	theronine (T)
TK	tyrosine kinase domain
Tris	tris(hydroxymethyl)aminomethane
Trp	tryptophan (W)
Tyr	tyrosine (Y)
v-	viral
Val	valine (V)
vCrk	viral Crk
WB	Western Blot
Xho1	restriction enzyme
YFP	yellow fluorescent protein
γA, γB	gyromagnetic ratios of two nuclei
ΔG°	free energy
ΔH	enthalpy change
ΔS	entropy change
θ	the angle between the internuclear bond vector and the z-axis of the alignment tensor
D	donor
A	acceptor
$\mu 0$	the permeability in vacuum

τ_c	correlation time
φ	the angle between the projection of the internuclear bond vector onto the x-y plane and the x-axis

List of figures

Chapter 1. Introduction

1.1 Crk Adaptor Proteins

1.1.1 Crk history

Cellular signals are mediated by many genes and proteins. Adaptor proteins are important molecules in many signaling pathways, and play a significant role despite the absence of any intrinsic catalytic activity.

The Crk adaptor protein family exhibits a dominant role in signal transduction pathways including cell adhesion, motility, phagocytosis, differentiation, proliferation, transformation and apoptosis[1 2 3]. The Crk gene was first identified by Mayer[4,5] in 1988 from a chicken tumor sample as an oncogene *gag-crk*. *Gag-crk* is a fusion of the viral *gag* gene encoding structural proteins of the virion and a cellular gene *crk* encoding two domains. There was increase of tyrosine-phosphorylated proteins in chicken embryo fibroblast transformed by this gene despite the absence of a catalytic domain. This initial observation suggested that crk is a cellular tyrosine kinase activator and[6,7] was named crk for chicken tumor virus no. 10 [CT-10] regulator of kinase. One year later Tsuchie[8] identified and characterized an independent avian retrovirus called ASV-1 (avian sarcoma virus 1) isolated from a tumor in an adult chicken and found that both oncogenes differ only in a few amino acids. Following the isolation of v-Crk, Reichman et al.,[9] have cloned and sequenced the cellular cDNA of chicken Crk. They found that cCrk and vCrk have similar structural organization, although vCrk possesses an additional proline-rich region of roughly 100aa.

Shortly after that Matsuda and colleagues isolated CRK cDNAs, from human embryonic lung cells and named them CRK-I and CRK-II[10]. Finally, Crk-like[11] (CrkL), a product of a different gene and the last member of the Crk family, was identified and found to have 60% homology to Crk.

Discovery of the viral Crk oncoprotein followed by cloning of its cellular homologs has significantly changed our understanding of intracellular signal transduction. Early experiments showing the interactions of v-Crk with phosphorylated proteins and finding of multiple Crk/CrkL binding proteins has provided many functional details of this extraordinary signaling machinery.

1.1.2 Crk Family members

The c-Crk gene encodes two splice variants named CrkI and CrkII, while a homologous but distinct gene encodes Crk-like (CRKL). All members of the Crk adaptor family are composed of Src homology 2 (SH2) and Src homology 3 (SH3) domains separated by a flexible linker (Fig. 1).

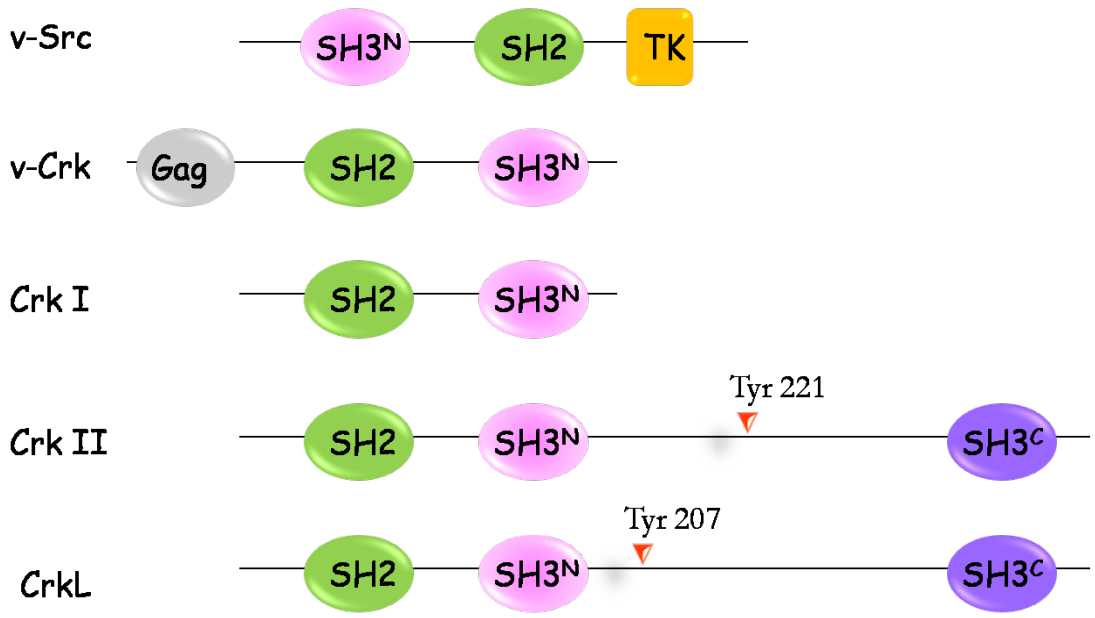

Figure 1. Structure of the Crk family of proteins. The domains are boxed: SH2 (Src homology 2); SH3 (Src homology 3); Gag, viral group specific antigen; Y221 or Y207, negative regulatory phosphorylation site. The structure of Src is shown at the top of the figure to indicate its spatial arrangements compared to Crk. TK, tyrosine kinase domain.

Selective formation of multi-protein signaling complexes by the Crk and CrkL proteins depends on specific motifs recognized by their SH2 and SH3 domains, which can trigger normal as well as oncogenic cell transformation[12]. On the N-terminus these proteins have an SH2 domain which is made of ~100 amino acids and mediates binding by interacting with a pY-X-X-L[13] motif-containing tyrosine phosphorylated proteins such as paxillin, p130Cas, Cbl or Gab1. The $SH3^N$ domain, which is made up of ~60 amino acids, can interact with proline-rich P-X-L-P-X-P binding motifs of c-Abl, SOS, and CG3. In contrast to CrkI, CrkII and CrkL have an additional C-terminus SH3 ($SH3^C$) domain attached to the first two domains by a ~50 amino acid proline-rich linker. These can be negatively

regulated by their tyrosine phosphorylation via the autoinhibitory mechanism of the intramolecular linker binding to the SH2 domain. Phosphorylation of Y221 in CrkII or Y207 in CrkL prevent the SH2 and $SH3^N$ domains from binding to their target proteins. The lack of a $SH3^C$ domain makes CrkI unable to be negatively regulated, therefore giving it higher transformation activity than CrkII or CrkL[14]. Studies suggest that the main physiological function of the $SH3^C$ domain is focused on stabilizing CrkII and CrkL into a conformation that regulates their biological function. The $SH3^C$ domain may regulate accessibility of the Crk proteins to their binding partners by intramolecular interaction with the $SH3^N$ domain and by proper linker folding between the two. $SH3^N$ has been well characterized and many binding partners have been identified[2]. The $SH3^C$ domain in both CrkII and CrkL has not been shown to have interacting partners except for one, nuclear exporter Crm1[15], which has been reported to bind to the $SH3^C$ domain of CrkII. Mutation of the $SH3^C$ domain of CrkII causes increased phosphorylation of many cellular proteins[16] and also affects the $SH3^N$ domain's interaction with its partners[17]. This suggests that the $SH3^C$ domain negatively regulates the binding of signaling proteins to the $SH3^N$ domain[18,19]. In CrkL the SH3c domain has no interacting proteins that have been identified, questioning the importance of its presence. Although CrkII and CrkL have similar domain organization and homology CrkL seems to play a more important role in hematopoietic cells[11] and was shown to be the most prominent substrate for the Bcr-Abl oncoprotein[20].

1.1.3 Crk in signaling

By using their SH2 and SH3 domains, Crk family adaptors control selective formation of multi-protein signaling complexes and regulate a vast number of biological processes including cell migration, morphogenesis, invasion, phagocytosis, survival and regulation of gene expression[21]. More than 20 years of research has significantly increased our knowledge and helped us to understand the biological functions, as well as the mechanistic properties, of Crk adaptor signal transmission. Since then over 40 cellular proteins have been shown to bind with Crk family proteins. Such diversity clearly indicates that we cannot assign one specific cellular function to Crk and shows the complexity of Crk signaling transduction.

As mentioned before, the major tools for Crk to control particular binding activity are its SH2 and SH3 domains. Crk signaling pathways can be divided into two groups: the input pathway and the output pathway. The input pathway involves the SH2 domain, which binds to a specific phosphorylated motif, inducing an on/off mechanism controlled by phosphorylation and dephosphorylation of the Crk tyrosine. The second mechanism involved $SH3^N$ domain, where the protein complexes occurs constitutively by proline-rich motifs and are not an effect of post-translational modifications.

One of the representative Crk signaling pathways (Fig. 2) is the mechanical stretching induced by phosphorylation of p130Cas[22] . In response to mechanical force, p130Cas is phosphorylated by Src or FAK tyrosine kinase, followed by binding of the phosphorylated YXXP motif to the SH2 domain of CrkII[23]. CrkII

binds GEFs for the Rap 1 (C3G) via its $SH3^N$ domain. This regulates cell adhesion and cytoskeletal organization[21]. The inter-$SH3^C$ linker region of CrkII undergoes Abl induced phosphorylation on Y221, resulting in intramolecular binding with the SH2 domain as a consequence CrkII dissociates from Cas. This pathway has been shown to play a crucial role in breast cancer cell tumorigenicity[24].

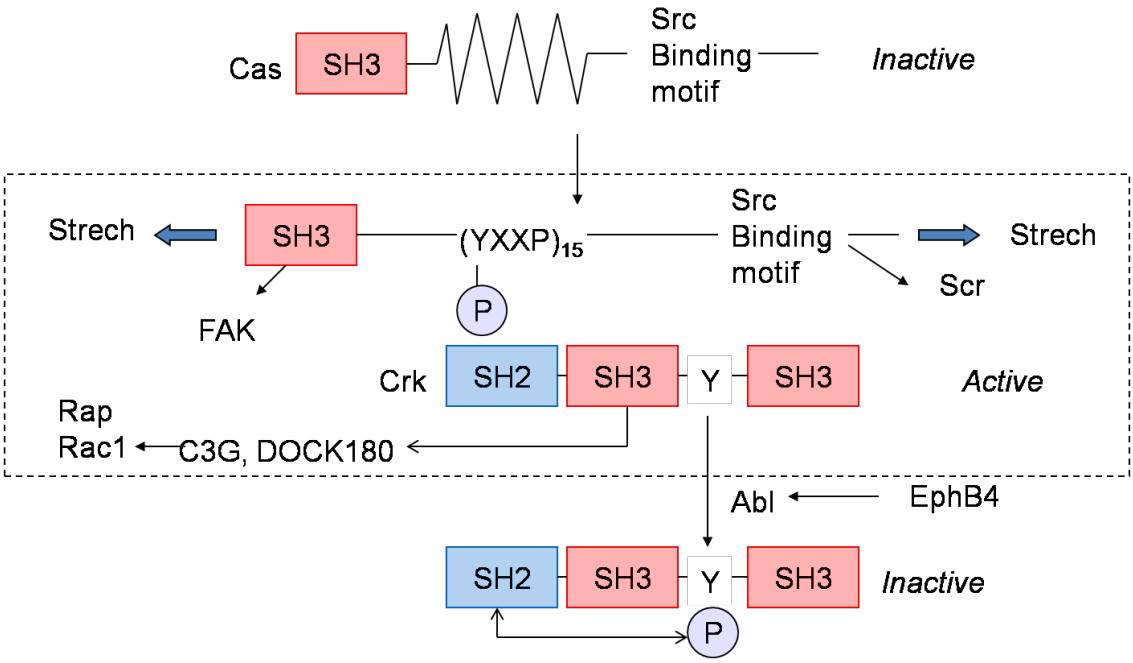

Figure 2. Model for the regulation of the Cas docking protein and the Crk adaptor. The central region of Cas, with 15 YXXP motifs, is protected from phosphorylation by intramolecular interactions. Stretching forces lead to extension of Cas and exposure of the YXXP motifs for phosphorylation by Src family tyrosine kinases, and these phosphorylated motifs then recruit the SH2 domain of the Crk adaptor, and effectors associated with the Crk N-terminal SH3 domain. Crk binding activity is suppressed by tyrosine phosphorylation at Tyr (Y) 221, for example by EphB4–Abl signaling, leading to intramolecular interaction with the SH2 domain. The dashed box shows the active Cas–Crk complex.

Despite detailed information of how the Crk can interact with its substrates, there is still lack of information how these complexes can form in real time and space.

1.1.3.1 SH2 domain

Post-translational modification by tyrosine phosphorylation is one of the most important mechanisms used in cells to convey signals in response to external and internal stimuli. It plays a major role in many cellular and developmental processes including cell proliferation and differentiation. Protein phosphorylation involves several factors. First are protein tyrosine kinases (PTKs), which are used to phosphorylate substrates. Next are protein tyrosine phosphatases (PTPs), which are used for dephosphorylation. Last are the specific protein domains, which can recognize the phosphorylated ligand, and, by doing so, initialize signaling events. Among all pTyr binding ligands (proteins/domains) which includes: SH2 domains, pTyr binding (PTB) domains and one C2 domain, SH2 is the major target.

Recently Liu and co-workers examined SH2 domain containing proteins from the genomes of 28 different eukaryotic organisms including 111 SH2 domains containing proteins from the human genome, highlighting the importance of SH2 domains in biological systems[25]. SH2 domains can be classified into different families according to their functionality[26]

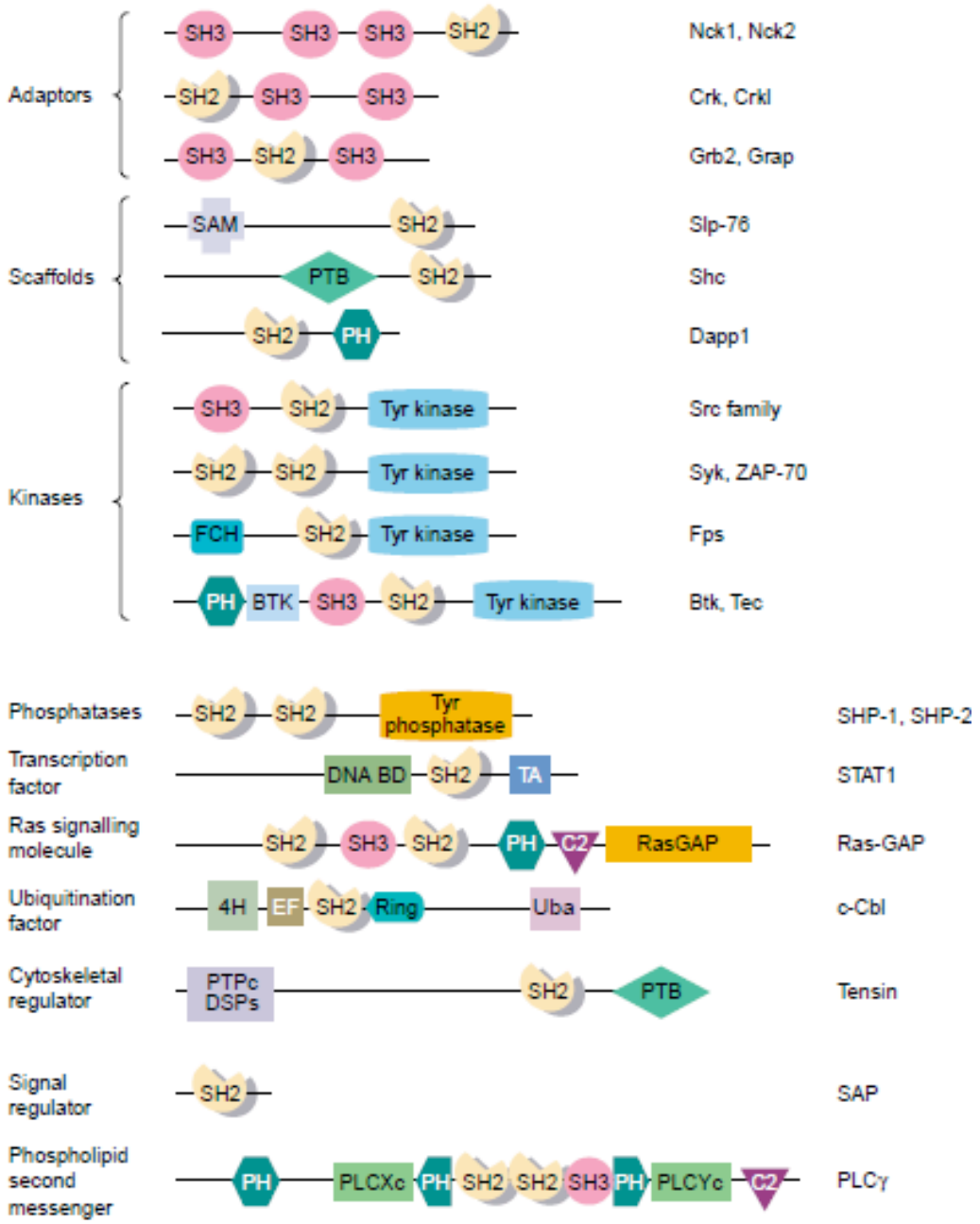

Figure 3. SH2 domain-containing proteins classification based on their functionality[27].

The SH2 domain was originally found in 1986 in the arcane retroviral oncoprotein v-Fps with an active tyrosine kinase domain on its C-terminus. It was described

as a polypeptide sequence containing approximately 100 residues and named SH2[28] for src homology 2. Authors have found that the SH2 domain is located N-terminal to the catalytic domain and is not required for kinase reactivity but is used rather as an activity regulator.

Later the SH2 domain was found to interact with tyrosine phosphorylated partners[4] and finally, in 1992, John Kuriyan and Gabriel Waksman solved the Src-SH2 domain structure bound to phosphopeptide[29]. They have shown in molecular detail how the SH2 domain recognizes the pTyr motif, and have identified the features, which are common for many interaction domains. This analysis uncovered several features that are now recognized as common themes of many interaction domains.

The Sh2 domain contains anti-parallel β-sheets surrounded by two α-helices (Fig.4)

Figure 4. SH2 domain (PDB ID 1SHB) of v-src complexed with tyrosine-phosphorylated peptides.

The SH2 domain binds to proteins containing phosphorylated tyrosine followed by specific residues. Crk SH2 domains recognize pTyr-Asp-x-Pro sequence and binds to the pocket where it is captured by a conserved Arg residue located in the SH2 domain. Arg play very important role by coordinating the phoshate oxygens from the phospho-Tyr (the basic residue is recognized by a negatively charged pocket). The way the SH2 domain interacts with pTyr proteins can be categorized into two major groups: intramolecular regulation, where phosphorylated site of the protein binds to its own SH2 domain (CRkII-pTyr221 or CrkL-pTyr207), and intermolecular, where the phosphorylated protein recruits its partner to form a complex.

SH2 domains have been shown to recognize specific sets of targets, for instance Socs2, which recognizes growth hormone receptor, or Socs4, which recognizes phosphorylated EGFR[30]. In contrast SH2 domains of protein families like Crk, NCK or GRB2 share a common core set of binding peptides and are highly conserved[31].

The discovery of SH2 domains was a major breakthrough in understanding how signaling networks are controlled by protein phosphorylation or by domain-domain interactions.

1.1.3.2 The SH3 domain

Another important protein-protein interaction member in context of cell signaling is Src homology 3 (SH3) domain. There are over 300 SH3 domains in the human

genome[32], making the SH3 domain one of the most prevalent families of protein modules found in nature showing a wide range of cellular functions tehes proteins arecapable of performing. SH3 domains are involved in many important cellular functions including intracellular signaling and cell-environment communication, cytoskeletal rearrangements and cell movement, cell growth and differentiation, protein trafficking and degradation, and immune responses[33,34,35]. SH3 domains were first characterized as polypeptide regions conserved between different signaling proteins such as the Src family tyrosine kinase, Crk, and phospholipase C-γ[4]. All SH3 domains consist of 50-70 residues and feature a five-stranded anti-parallel beta-sheet structure[36,37].

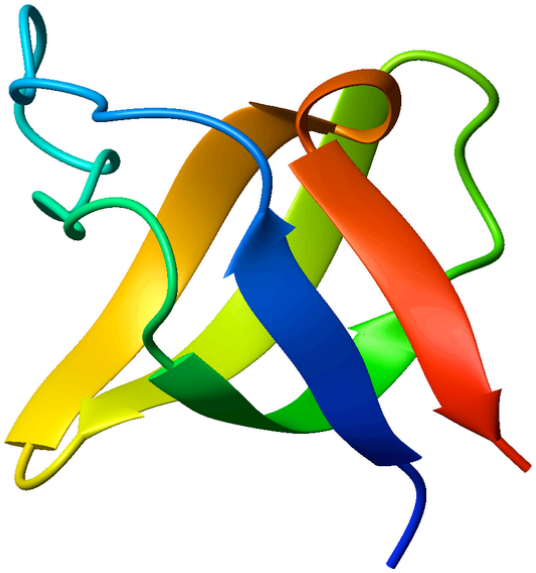

Figure 5. SH3 domain (1SHG)

The first report involving an SH3 domain in context of protein-protein interaction was described between protein 3BP-1 bound to the SH3 domain of Abl[38]. A year later it was shown that only the short, proline-rich fragment of 3BP-1 is required

for Abl SH3 domain binding and was identified to be PxxP[39]. Soon after identification the PxxP motif Chen et al. classified SH3-binding motifs into two similar yet distinct groups called classes I and II[40]. Class I is represented as [R/K]xXPxXP and class II as XPxXPx[R/K][41], where the 'X' signifies a non-glycine, hydrophobic residue while the 'x' denotes any naturally occurring amino acid. SH3 domain complex structures with the class I and II peptides revealed that these ligands bind in opposite directions. In each case the peptide forms a left-handed helix which is called polyproline type II (PPII) with three residues per turn. Some SH3 domains prefer one peptide orientation over the other but others make very little distinction between the two. According to the work published by Wu[42], which tested binding of 686 class I and 686 class II peptides to 12 SH3 domains, each SH3 domain can bind either class of peptides however class II with higher selectivity.

An Arg or a Lys residue that precedes or follows the PxxP core motif determines the orientation of a peptide with respect to the binding site. In either case, the basic residue is recognized by a negatively charged pocket (or specificity pocket) on the SH3 domain (Fig. 6). Apart from electrostatic interactions with acidic residues lining the specificity pocket, the side chain of an Arg or a Lys makes favorable charge-aromatic interactions and/or Van der Waals contacts with a Trp residue that is conserved at the binding site for all SH3 domains.

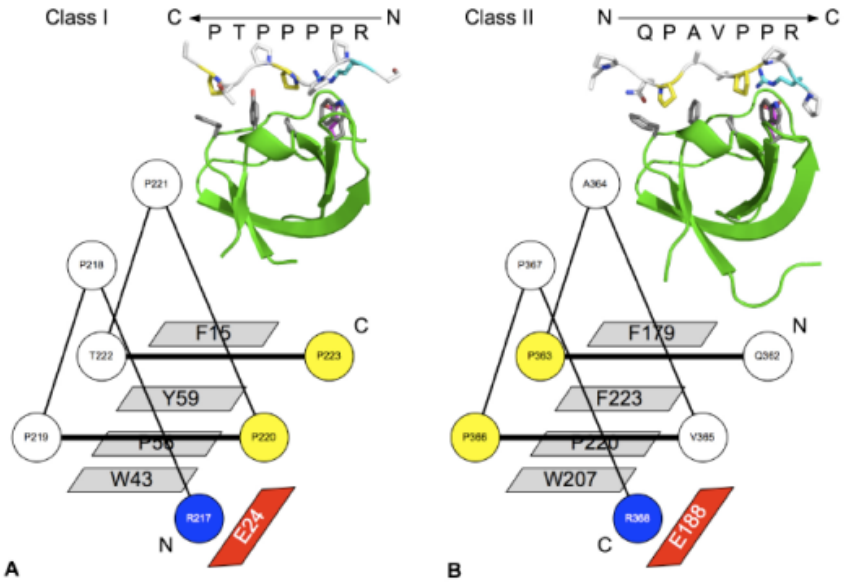

Figure 6. Structural basis of an SH3 domain binding to a Class I or II peptide. An schematic representation of SH3 domain recognition of a peptide in a 'plus' (C_N) (left) or 'minus' (N_C orientation (right) (9). The left-handed PPII helix of the peptide is shown as circles (residues) connected by sticks (amide bonds). Bold lines represent the XP dipeptide units. Conserved residues found at the ligand-binding site of the SH3 domain are shown in rectangles. (A) The beta-PIX SH3 domain in complex with a peptide derived from AIP-4 (PDB code: 2P4R) (14). The peptide contains the class I motif ([R/K]xXPxXP) and is bound in the plus orientation. (B) The p40phox SH3 domain in complex with a p47phox-derived peptide (PDB code: 1W70). The peptide contains the class II motif (XPxXPx[R/K]) and is bound to the SH3 domain in the minus orientation[43].

In the context of CrK, the first protein shown to bind with the $SH3^N$ domain was Crk SH3-domain-binding guanine-nucleotide releasing factor (C3G)[44].

Although many proteins have been identified to bind to the $SH3^N$ domain, no cellular molecules have been shown to interact with the Crk $SH3^C$ domain, with

the single exception of the nuclear export factor crm1[15]. Authors have shown that $SH3^C$ domain of Crk contains a binding site (nuclear export sequence - NES) for the nuclear export factor Crm1 and that a mutant lacking the NES sequence promotes apoptosis. Both CrkII and CrkL contain an atypical C-terminal SH3 domain, which is unable to bind Polyproline Type II (PPII) motifs[19]. In both proteins, the domain adopts core structural characteristics of SH3 domains, comprising of a five-stranded beta barrel. However, unlike the $SH3^N$ domain, where the aromatic amino acids – F141, W169, Y186 line the canonical PPII binding pocket, the $SH3^C$ domain contains polar residues instead – Q244, Q274 and H290. Therefore SH3c cannot bind typical PXXP ligands or binds them with reduced affinity.

1.1.4 Binding partners via Crk/CrkL SH2 and $SH3^N$ domains

1.1.4.1 p130Cas family

As previously mentioned, nearly all proteins which were shown to bind to SH2 domains of Crk used YxxP motifs (where Y undergoes phosphorylation. This general rule also applies to a protein called Crk-associated substrate (p130Cas), which is controlled by phosphorylation on both tyrosine and serine/threonine residues and was first identified as a tyrosine phosphorylated protein in cells transformed by v-Crk[7] and v-Src[45]. Due to the presence of many conserved sequence motifs and tyrosine and serine/threonine phosphorylation, p130Cas controls formation of multiprotein complexes via protein–protein interactions and

therefore plays an essential role in intracellular signaling events. Cas/Crk is one such complex, which activates downstream effectors such as Rac and C3G and reorganizes the actin cytoskeleton[46]. Overexpression of Cas proteins is involved in the development of many human cancers. For instance in human breast cancer overexpression of p130Cas and ERBB2 (p130CAS is necessary for transformation by several oncogenes like ERBB2) is associated with increased proliferation, metastasis and poor prognosis[3,47]. Mice lacking Cas ($Cas^{-/-}$) exhibited cardiovascular anomalies and died in utero[48].

The large (130 kDa), multi-domain Cas molecule contains an N-terminal SH3 domain, proline-rich domain, a large substrate-binding domain, a serine-rich region, and a C-terminal region (Fig. 7).

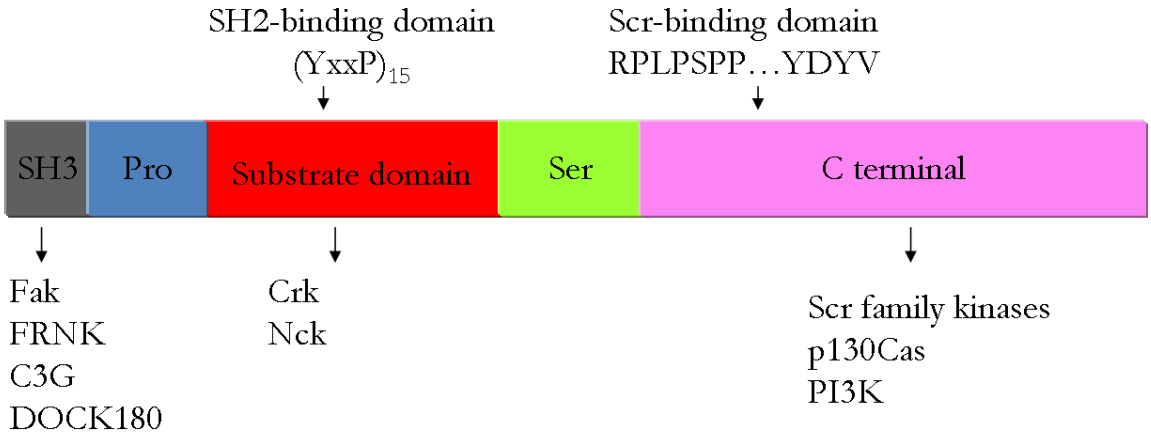

Figure 7. The structural characteristics and interacting proteins of p130Cas.

p130Cas tyrosine phosphorylation occurs in the substrate-binding domain, which contains 15 repeats of a YxxP sequence. Tyrosine phosphorylation in the YXXP motif, which is regulated by many growth factors and hormones[46,49], creates binding sites for the SH2 domains of downstream target molecules, including

Crk. As an example, p130Cas phosphorylation by Src kinase leads to assembly of the p130Cas/Crk/DOCK180 protein scaffold at the adhesion sites and was shown to play an important role in cell migration[50,51]. The C-terminal, polyproline region of p130Cas seems to have crucial function in terms of its activity. It activates Src family kinases by its SH3 domain displacement and anchors Cas to Src, allowing more efficient phosphorylation of Cas. It was shown that Cas-Crk complexes function to promote cytoskeletal rearrangements through activation of Rac1. Molecules that may link Cas-Crk interactions to Rac1 activation include C3G[44] and DOCK180[52]. The Cas-Crk pathway has also been shown to be important for chemotaxis induced by insulin and epidermal growth factor (EGF)[53,54]. Recently another interesting work by Sawada[22] shed a light on the Cas protein tyrosine phosphorylation level induced by mechanical extension.

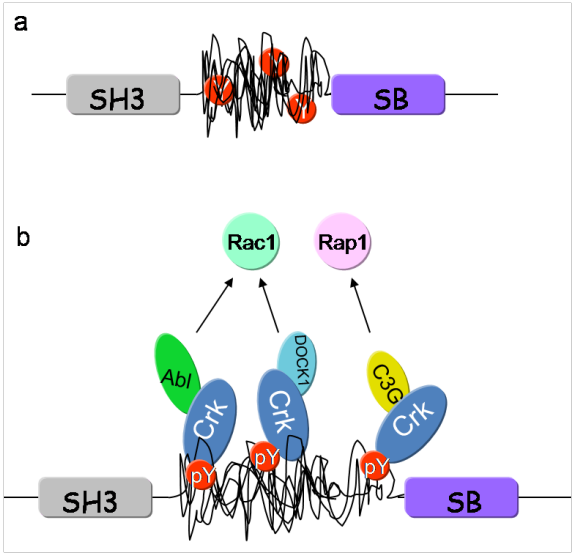

Figure 8. Model of extension of p130Cas and signaling. (a) Represent a Cas molecule with unextended configuration where none of the Y (Tyr in red) is phosphorylated. (b) The extension-dependent phosphorylation of the Cas substrate domain by SFK and enhancement of its signaling.

Similarly to Cas, another multiple sequence repeat containing proteins like tatin[55] were shown previously to have mechanical functions. By analyzing this mechanism authors found that cell stretching has significantly increased Cas substrate-domain phosphorylation by Src with no observable change in Src activity and could serve as an important signaling switch by controlling kinase or phosphatase. In another recent work, mutants containing single or multiple YXXP mutations were phosphorylated processively by Src, suggesting that individual sites are dispensable[56]. The results also suggest that there is no defined order to the Cas phosphorylation events so the multiple YXXP sites could be used to diversify the signaling pathways. All these studies taken together could help shed a light on the reason for duplicating PXXP motif in substrate-domain of p130Cas 15 times. This indicates many molecular complexes (Fig. 8) (like p130Cas/Crk/DOCK180, p130Cas/Crk/C3G, p130Cas/Crk/Abl, p130Cas/Crk/JNK, and p130Cas/Crk/PI3K) could exist simultaneously, because substrate-domain would have many accessible phosphorylated motifs available at any given time.

Abl kinase was also shown to negatively regulate the migratory function of p130Cas. On one hand, by phosphorylating Crk on Tyr221, it dissociates p130Cas-Crk complex[57]. On the other hand, Tyr221 of Crk is dephosphorylated by the tyrosine phosphatase PTP-1B[58], promoting p130Cas–Crk coupling and cell migration. Taken all together this shows p130Cas diversity and it's potential in cell signaling regulation.

1.1.4.2 Paxillin

Paxillin is a 68-kDa substrate to various tyrosine kinases. It was initially identified as a member of the cytoskeletal proteins where it was shown to be tyrosine phosphorylated when co-expressed with v-src[59]. Soon after that initial finding, paxillin was identified as a vinculin binding partner, which is known to be involved in focal adhesion and actin binding[60]. Because of its multidomain structure and the absence of enzymatic motifs/patterns paxillin is thought to be an adaptor protein[61]. Paxillin plays a major role in controlling signals from integrins and growth factor receptors, which result in efficient cellular migration.

On its N-terminus are five leonine-rich sections, called paxillin LD motifs (LDXLLXXL). They serve as protein recognition domains and were identified as binding sites for the proteins vinculin and FAK[62]. Beside the LD motifs the N-terminus contains a short proline-rich motif charasteristic of the SH3 domain and many tyrosine phosphorylation sites (Fig. 9). FAK, together with Src, phosphorylates paxillin at two major sites Y31 and Y118[63,64].

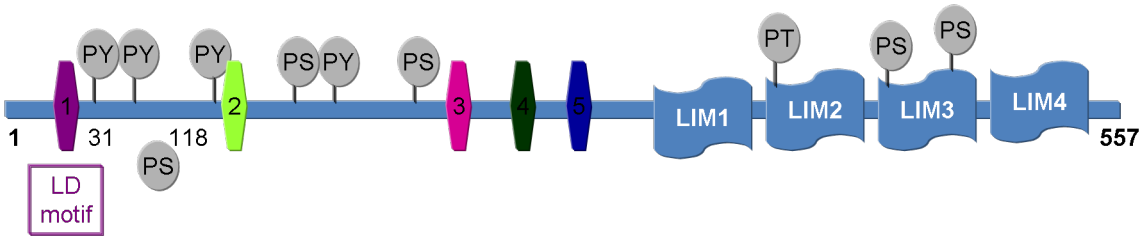

Figure 9. Domain structure of the Paxillin. Paxillin contains 5 leucine-rich LD motifs (consensus LDXLLXXL), 4 double zinc finger LIM domains, and two major phosphorylation sites Y31 and Y118.

Paxillin phosphorylation results in functional binding site generation for SH2 domains of Crk family members[65,63,64,66]. It has been demonstrated that by

introducing a nonphosphorylatable tyrosine mutant (Y31F/Y118F) on paxillin the migration was blocked[67]. These events are important in the context of integrin-mediated cell motility that is controlled by assembly and disassembly of signaling complexes like Crk/ P130Cas/paxillin[67,68].

Paxillin C-terminus is comprised of four LIM domains, which are double-zinc finger motifs[69,70,71], and has also been shown to be very important for mediating protein-protein interaction[69]. For example, LIM3 and LIM4 domains would facilitate dephosphorylation of p130Cas by direct binding to protein tyrosine phosphatase–PEST (PTP–PEST). This may control the aforementioned Crk/ P130Cas/paxillin complex[72 73] by disassembling it. Paxillin can also bind and be phosphorylated by the non-receptor tyrosine kinase Abl[74,75], which regulates Crk proteins binding activity.

1.1.4.3 C3G

GEFs (guanine-nucleotide-exchange factors) are the regulators of the GDP-GTP exchange process that activates small GTPases (20-25kDa). This controls a variety of cellular events, including cell differentiation, proliferation or apoptosis[76]. One of the GEFs that had originally been isolated as one of the two major binding partner of the Crk adaptor protein is C3G (Crk SH3-binding guanine-nucleotide-exchange factor)[77].

C3G has two functionally important regions; the central domain and the catalytic domain[44] (Fig. 10). The central domain, also known as Crk binding region (CBR),

contains three proline-rich sequences that bind to the $SH3^N$ domain of Crk and one sequence that binds to p130Cas[78]. The catalytic domain (CDC25 homology domain), located on the C-terminus, is responsible for Rap1 protein activation.

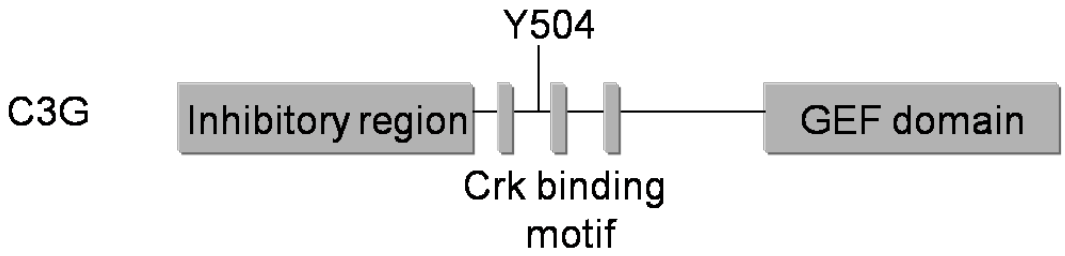

Figure 10. Structural characteristic of C3G

Depending on the cell type, C3G can be activated by several different signals, including integrins[79] and reelin stimulation[80], and can itself also mediate activation of several pathways including JNK[81,82] and ERKs (extracellular-signal-regulated kinases)[83].

Ichiba et al. demonstrated that C3G phosphorylation (on Tyr504) and binding to Crk are necessary for C3G activation[84]. Crk/C3G complex formation, and its translocation to the cytoplasmic membrane, cause the GTP exchange on Rap1[85] and activate signaling. When the formation of this complex is inhibited, Rap activation is blocked[86]. Rap1 signaling mediates different cellular events like adhesion, migration, phagocytosis, inflammation and differentiation[87]. It has been shown that p130Cas/Crk/C3G complex formation mediates Rap1, activates B-Raf in neurons and activates the mitogen-activated protein kinase (MAPK) pathway to control neuroblast differentiation[88].

In another study, C3G knockout mice died before embryonic day 7.5, which shows that C3G activation of Rap1 is required for the early embryogenesis mice[85].

In signaling pathways C3G was found to integrate variable signals and mediate many cellular functions due to its catalytic activity as well as direct protein interaction.

1.1.4.4 SOS

SOS (Son of Sevenless) protein was discovered over two decades ago as a Ras activator controlling the process of central photoreceptor (R7) cell development in *Drosophila melanogaster* eye[89]. Homologues of SOS were subsequently discovered and characterized in other organisms from *Caenorhabditis elegans* to mammalian cells[90,91,92,93]. There are 2 human homologs of SOS, hSOS1 and hSOS2[94].

The hSOS1 protein (150 kDa) is composed of several important functional domains (Fig. 11).

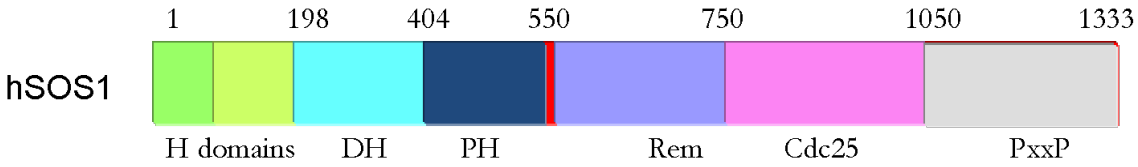

Figure 11. Organization of hSOS 1

The C-terminus has a proline-rich (PXXP) domain and was shown to interact with SH3 domains of proteins like E3B1 and Grb2 (a Rac1 GEF)[95]. The central region

of hSOS1 has two domains: CDC25 (cell division cycle 25) and REM (Ras exchanger motif). CDC25 and REM were shown to bind and stimulate Ras-GTP exchange. The structural data indicate that upon CDC25/REM interaction with Ras-GDP, the SOS can adopt a fully functional structural state[96]. The SOS1 N-terminal contains two domains typical for nucleotide exchange factors, the Dbl homology (DH) and the Pleckstrin homology (PH). The catalytic DH domain allows Rac activation and the PH domain regulates the DH domain, through an unknown mechanism. In general the PH/DH duet functions to exchange GTP for GDP on Ras.

Several studies show that the Crk SH3 domain binds to SOS[97,98]. However the cellular function of the SOS/Crk complex and its implications are not well understood.

1.1.4.5 DOCK180 (DOCK1)

DOCK180 (180-kDa protein downstream of CRK) was identified as one of the most prominent Crk binding partner in `Far Western' overlay blot[99]. It was then cloned and described as a cell structure regulator upon the activation of tyrosine kinases[100]. DOCK180 is known to function as GEFs for various GTPases. In mammalian cells, the Crk–ELMO–DOCK180 complex activates several Rac-dependent pathways, including the JNK kinase cascade, actin remodeling, cell migration and engulfment of apoptotic cells[101,102,53,103,52]. As with all GEFs[104,105], DOCK180 promotes the exchange of protein bound GDP for GTP. In all signaling

systems GEFs act as upstream regulators of Rho and Ras family GTP binding proteins. Unlike other previously described C3G exchange factors DOCK180[106] lacks the typical Dbl homology (DH) domain or pleckstrin-homology (PH) domain, which are known as nucleotide exchange promoters in all known mammalian Rac GEFs. Instead it uses a novel DHR-2 (DOCK Homology Region-2) domain that specifically recognizes nucleotide-free Rac[107] and induces the GTP loading of Rac both in vitro and in vivo. However, proper activation involves the PH domain of ELMO[108], which facilitates Dock180-mediated Rac activation by helping to stabilize Rac in its nucleotide-free transition state and by binding the DOCK-Rac complex in trans. It was shown that by inactivation of the DHR-2 domain in Dock180, Rac activation was blocked[109]. This highlights the importance of this domain in the biological function of the protein

Another conserved domain in all members of the DOCK180 related superfamily is DHR-1 (DOCK Homology Region-1), located upstream of DHR-2 domain. DHR-1 was shown to mediate specific interaction with phosphatidylinositol (3,5)-bisphosphate and PtdIns(3,4,5)P3 signaling lipids in vitro and in vivo. This results in DOCK180 membrane translocation[110].

The importance of the DHR-1 domain in DOCK180-mediated signaling was proved by expression of DOCK180 lacking DHR-1 (DOCK180ΔDHR-1) in LR73 cells (used for cell migration studies). In cells where DOCK180ΔDHR-1, ELMO1 and CrkII were coexpressed enhanced motility was not observed. This demonstrates that the DHR-1 domain is essential for DOCK180-mediated cell movement. Despite not being able to promote cell migration, the

DOCK180ΔDHR-1 was unexpectedly shown to be able to promote Rac GTP-loading to the same extent as wild-type DOCK180 in both the presence and absence of ELMO1 and CrkII. These results highlight the difference between Rac activation and signaling, support the notion that GTP-loading of Rac alone is not sufficient to promote cell migration, and emphasize the importance of DHR-1. Rac was shown to play a key role in not only controlling actin dynamics, integrin-mediated cell adhesion and cell shape changes, but also in the engulfment of dead cell bodies.

Although DOCK180 was identified as a Crk binding partner its role in DOCK mediated Rac activation is not quite clear. It has been shown that CrkII binds PxxPx(K,R) in the C-terminal region of DOCK180 via the $SH3^{N}$ domain[111,112] and promote translocation of DOCK180 to the plasma membrane[52,113]. In one study[17] authors show how the assembly and function of the DOCK180/ELMO/Rac can be regulated by introducing mutation in CrkII $SH3^{C}$ domain, suggesting that Crk can control assembly of the complex. Contrary to these findings other study[114] have shown that a direct interaction of CrkII with DOCK180 was not essential for both engulfment of apoptotic cells or for the recruitment of DOCK180. Since the expression of a DOCK180 mutant lacking the proline-rich sequence was sufficient for Rac1 activation, direct physical association between CrkII and Dock180 is not essential for these proteins to function during engulfment. Despite the fact that the Crk-DOCK180 interaction involves a specific binding motif (PxxPxK) it was shown that DOCK2[115], a homolog of DOCK180 where PxxPxK is not present, can bind CrkL. This suggests different binding specificity

of SH3 domains of these highly homologous proteins and also that the CrkL-DOCK2 complex may regulate cell motility (through Rac). Thus the role CrkII plays in the function of the aforementioned complex is still elusive.

1.1.4.5 Abl Kinase

Abl is a 140kDa non-receptor tyrosine kinase localized at several sites inside the cell including the nucleus, cytoplasm, mitochondria, endoplasmic reticulum and the cell cortex. Abl was identified as a homolog of the v-Abl oncogene – the retroviral oncoprotein from Abelson murine leukemia virus (A-MuLV)[116]. There are two members of the Abl family; c-Abl (Abelson tyrosine kinase- Abl1) and its paralogue, Arg (Abl related gene -Abl2)[117]. C-Abl was found to interact with various cellular proteins including kinases, phosphatases, transcription factors, cytoskeletal and signaling proteins[118]. It has also been shown to play an important role in many cellular processes, including actin dynamics, cell growth, survival and migration. There are two isoforms of c-Abl: types 1a and 1b. The difference is a myristoyl group covalently linked to the N-terminus of the 1b type. The myristate moiety was found to be a critical for autoinhibition by binding to a unique binding pocket in the c-Abl kinase domain[119]. Abl contains on its N-terminus SH3, SH2 and tyrosine kinase domains (TK) but also a large (60 kDa) C-terminal domain containing cytoskeletal protein and DNA-binding domains, which are unique to Abl and Arg (Fig. 12). This long C-terminal extension contains multiple interaction sites like proline-rich motifs that are used to

communicate with $SH3^N$ domains of adaptor proteins such as Crk, Nck and Grb2 (growth-factor-receptor-bound 2)[120 121].

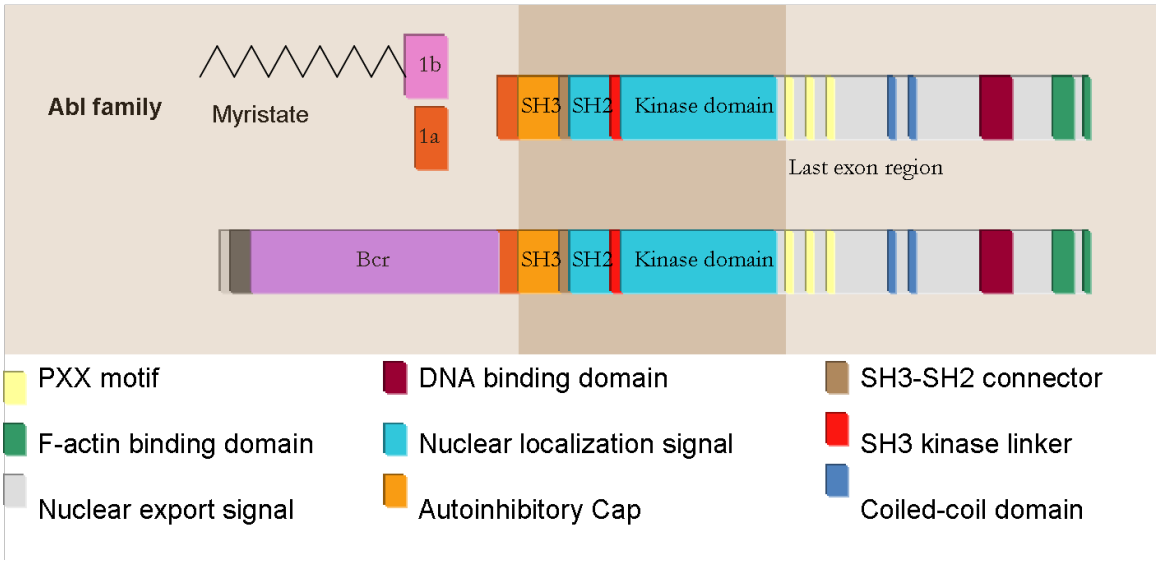

Figure 12. Abl family domains organization.

The Abl activity is controlled by a mechanism of autoinhibition, which is achieved by many complex intramolecular interactions including a catalytic kinase domain with activation loop, SH3/SH2 domain interactions, and other segments in the N-terminal site such as the myristoyl group and the SH3–SH2/SH2–kinase-domain linkers[122,123] (Fig.13).

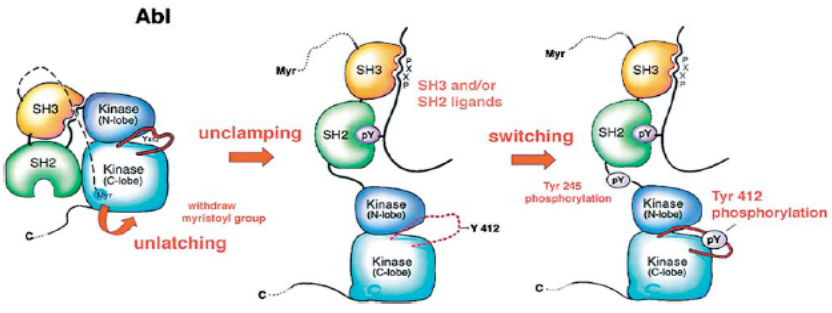

Figure 13. Three-step mechanism of ABL tyrosine kinase activation[124].

Abl is known to bind numerous partners via its SH2 and SH3 domains. The Abl SH2 domain can engage phosphotyrosine, but the phosphorylated tyrosine on Abl can then bind to other SH2 domain. Analogously the PXXP motif on Abl can bind to partner SH3 domain whereas Abl SH3 domain can engage PXXP motif of its partners. Among direct Abl binding partners, the CAS family and the CRK-family are key regulators of cell attachment and motility[125]. In a report by Shishido et al., it was shown for the first time that interaction of the $SH3^N$ domain of Crk and the proline-rich motifs (P545xxP and P589xxP)[126] of c-Abl are essential for the phosphorylation of Crk and activation of c-Abl. They have shown that Abl activation is negatively regulated by phosphorylation of the tyrosine 221 of c-Crk[127]. This phosphorylation can disrupt Crk-CAS complexes[57]. Crk expression with Abl induces Abl phosphorylation on its two major sites (Tyr 245 and Tyr 412)[128,129] suggesting that Crk enhances enzymatic activity of Abl. Another observation is that Crk transactivation is much more efficient when either the Y222F or the P225A Crk mutants are used to prevent Crk phosphorylation and intramolecular binding to its own SH2 domain[130].

Abl was also identified as a part of a fusion oncoprotein, BCR-ABL[131] (which is known as the Philadelphia chromosome, a result of Abl deregulation by translocation between chromosomes 22 and 9), found in more than 95% of human adults with chronic myeloid leukemia[132]. Although Abl is known to play an essential role in CML pathogenesis, the detailed mechanism by which it can regulate the disease remains unclear. However, some studies have shown that it induces tyrosine phosphorylation of downstream signaling molecules, including

CrkL. CrkL was shown to be the major tyrosine-phosphorylated protein detected in the peripheral blood cells of patients with CML[133,134].

The specificity of the Bcr-Abl/CrkL complex, has led to the acceptance of Y207 phosphorylation of CrkL as a CML diagnostic tool[135].

1.1.5 Regulation of Crk by $SH3^C$ domain

CrkII and CrkL, in addition to the $SH3^N$ domain, contain the $SH3^C$ domain which is unable to bind to Polyproline Type II (PPII) motifs[19]. More broadly, there are no cellular molecules that have been shown to interact with the Crk $SH3^C$ domain, with the single exception of the nuclear export factor crm1, which binds to LALEVGELVKV sequence of the $SH3^C$ domain. However, the solution structure of Crk $SH3^C$ domain shows the LVK motif to be buried in the hydrophobic core of the $SH3^C$ domain. This could be partially explained in case of the CrkL $SH3^C$ domain. The $SH3^C$ domain monomer-dimer transition caused by its partial unfolding can expose the LVK motif, thereby providing a binding site for Crm1[136]. As discussed previously, the $SH3^C$ domains in both CrkII and CrkL adopt core structural characteristic common to all SH3 domains, composed of a five-stranded beta barrel. Despite similar structural characteristic, aromatic residues on the $SH3^N$ domain, which line the canonical PPII binding pocket, are replaced by polar residues on the $SH3^C$ domain surface (Fig. 14). Subsequently the $SH3^C$ domain cannot activate multiple transmitting signals.

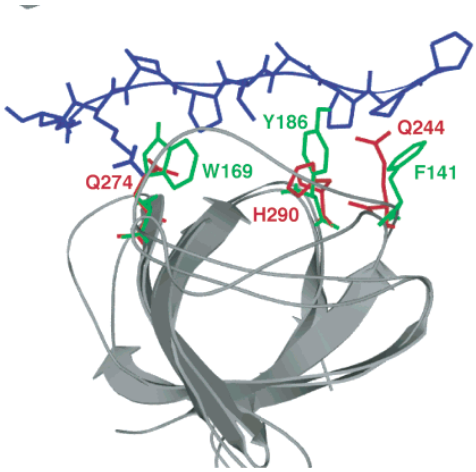

Figure 14. Overlaid structures of the $SH3^N$ and $SH3^C$ domains of CrkII are shown. The polyproline ligand for the $SH3^N$ domain is also shown (in blue). The aromatic residues that line the ligand-binding pocket in the $SH3^N$ domain (in green) and the corresponding residues occupying the same position in the $SH3^C$ domain (in red) are also indicated.[19]

Another interesting aspect of $SH3^C$ domain is its ability to regulate CrkII functions by constraining binding to the $SH3^N$ domain. Recent studies on chicken CrkII have revealed that the binding activity of this protein is regulated by cis-trans isomerization at Pro238[137,138]. Authors show that Crk construct comprising $SH3^N$-linker-$SH3^C$ exists in two conformations in solution. The major conformation (90%), where Gly237-Pro238 adopts the cis conformation, and a minor (10%), where Gly237-Pro238 adopts the trans conformation. In the dominant cis conformation the $SH3^N$ and $SH3^C$ domains interact in an intramolecular fashion, causing an autoinhibitory mechanism by masking the canonical PPII-binding site. In the subordinate trans conformation, representing an uninhibited state, the $SH3^N$ and $SH3^C$ domains do not interact. In this case the $SH3^N$ domain is not occluded by the $SH3^C$ and remains accessible for the binding of the PPII ligands. The human CrkII sequence around Pro238 is not conserved and it remains

unclear whether a cis-trans isomerization mechanism negatively regulates human CrkII. However, the solution structure of human CrkII has revealed another interesting mode of regulation. The linker between the two SH3 domains containing residues 224-237, called the Inter-SH3 core (ISC) was shown to form contacts with the SH2 and both SH3 domains to control the assembly of human CrkII[139]. Phosphorylaton of human CrkII on Y221 negatively regulates its function. Intramolecular binding of pY221 by the SH2 domain prevents the SH2 and $SH3^N$ domains from binding to their substrates. A recent study shows that Y251 located on the $SH3^C$ domain, can be phosphorylated by Abl along with Y221[140]. Phosphorylated Y251 binds specifically to the Abl SH2 domain and transactivates Abl. In addition to the above the $SH3^C$ domain was shown to have a unique role in phagoctosis and Rac avctivation. Mutation in the CrkII $SH3^C$ domain results in the stabilization of a ternary complex of Crk, DOCK180 and ELMO[17]. This indicates that the $SH3^C$ domain regulates the assembly or disassembly of the Crk complex. Interestingly, the $SH3^C$ domain has also been proposed to enhance the ability to activate c-Abl. It was also shown that either the mutation (W276K) or truncation of the $SH3^C$ domain increased c-Abl binding to $SH3^N$ domain of Crk and increased Y222 phosphorylation in chicken CrkII[130]. These results suggest that the $SH3^C$ domain of CrkII could serve as its negative regulatory element and as an Abl activator. In any case, it is not fully understood how these seemingly diverse functions are accomplished.

1.1.6. Solution structure of CrkII and its regulation by tyrosine phosphorylation

The solution structures of CrkI, CrkII and phosphorylated CrkII (pCrkII) solved by Inagaki and collegues revealed novel regulatory mechanisms of the Crk adaptor protein. CrkI (residues 1-204) was shown to have an extended structure in which both SH2 and $SH3^C$ domains are accessible for interactions with its substrates. In contrast, CrkII (residues 1-304) revealed a compact structure where the binding site of SH2 is exposed, but $SH3^N$ is covered by SH2 (Fig. 15).

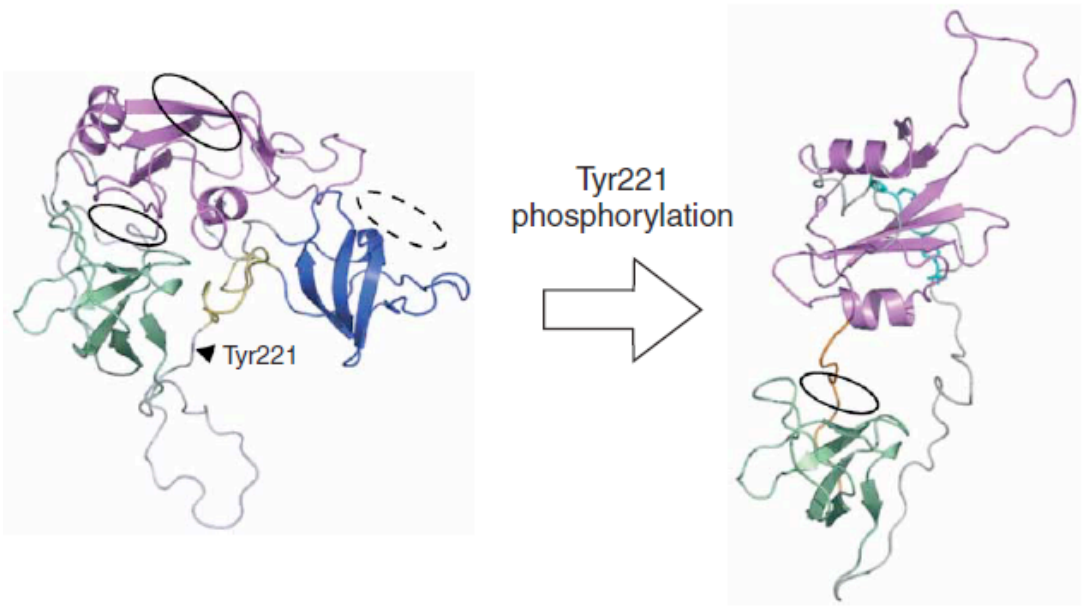

Figure 15. Ribbon representation of CRKII (left) and $pCRKII_{1–228}$ (right), SH2 (magenta), $SH3^N$ which is presented in the same orientation in both structures (green), $SH3^C$ (blue); ligand-binding sites in SH2 and SH3 are circled (dotted circle in CRKII represents putative binding site of $SH3^C$). In pCRKII, the inter-SH2-nSH3 linker (121–133) is colored orange and the phosphorylation site (221–224) is cyan[139].

In CrkII, there are many interactions between the SH2 and $SH3^N$ domains and between the SH2 and $SH3^C$ domains, however $SH3^N$ and $SH3^C$ do not interact. The CrkII structure is stabilized via the inter-SH3 core (ISC), which contains

residues 224-237. The ISC region makes contacts with all three domains (Fig. 16). Pro224 and Val226 of the ISC forms a hydrophobic core with Val184, Lys189, Trp170 and Lys164 of SH3N; the second is formed between Leu230, Pro231, Ile236 and Pro237 of the ISC and Ile269, Trp275, Val266, Lys265, Tyr239 and Ile263 of $SH3^C$. The mutation in the ISC region (CRKII-m226–237) or truncation (CRKII-Δ226–237) is indeed responsible for formation of the assembled structure as judged by NMR spectroscopy. In CrkII the binding site in the $SH3^N$ domain is masked by the SH2 domain, which partially mimics the binding of the PXXP peptide.

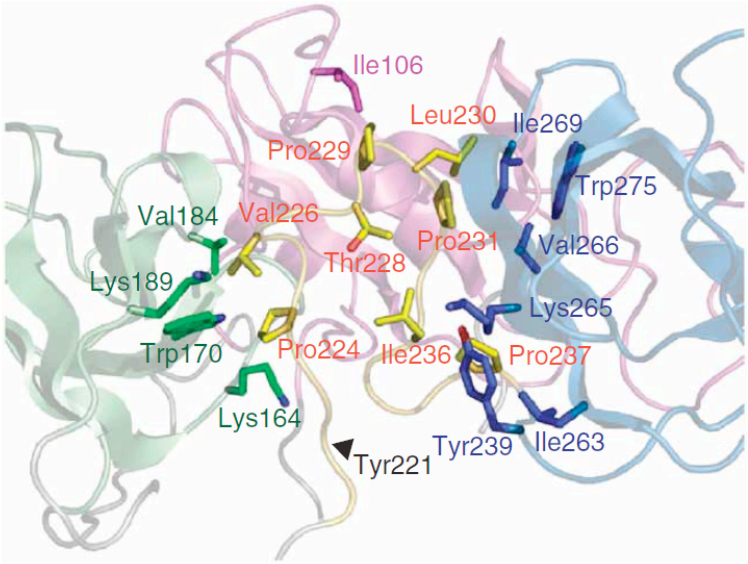

Figure 16. Ribbon model of hydrophobic core of CRKII, between the ISC (yellow) and the SH2 (magenta), $SH3^N$ (green) and $SH3^C$ (blue)[139]

The Inagaki study of pCrkII (resudues 1-228) reveals intramolecular interactions between the SH2 domain and the pTyr221, which agree with previous findings[141] (Fig. 15). In the phosphorylated state the ligand-binding pocket of the $SH3^N$

domain is also masked by the SH2-$SH3^N$ linker (Arg122–Glu133) suggesting that neither domain is accessible for the peptide.

These structural analyses reveal how intramolecular interaction of the SH domains with the specific ISC region can control the structure assembly, and also how phosphorylation can act as a intramolecular switch controling Crk activity. Authors have shown that the binding site of CRKII SH2 is exposed, while that on the $SH3^N$ is partially masked by interaction with SH2. However authors have speculated that there could be an equilibrium between these two states (Fig. 17a and b). In the case of pCrkII, the phosphorylation of Tyr221, by tyrosine kinases induces intramolecular binding to SH2. This detaches CrKII from SH2-mediated complexes and completely abrogates the signal transduction controlled by CrkII (Fig. 17 c).

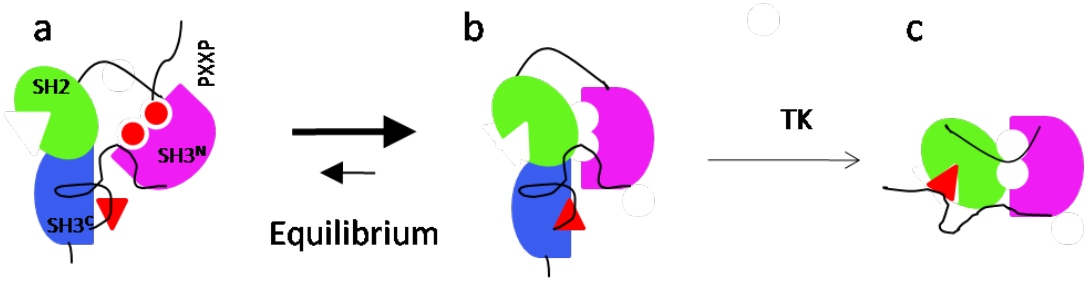

Figure 17. Schematic representation of the domain structure of CrkII. SH2 is green, $SH3^N$ is magenta and $SH3^C$ is blue. (a) The binding sites of CRKII SH2 and $SH3^N$ is exposed. (b) The binding site of CRKII SH2 is exposed but its $SH3^N$ is masked by interaction with SH2. Therefore, there may be an equilibrium between (a) and (b) based on the fact that CRKII and pCRKII had 6 times and 16 times lower affinity for PXXP peptide than CRKI. (c) pCrkII, the binding sites are not accessible.

1.1.7 Crk and CrkL: differences and similarities

Since Crk and CrkL can associate with identical proteins in vivo, they were expected to have similar functions. Indeed, some recent studies suggest that CrkII and CrkL can compensate to each other in cellular signalling. One of the examples comes from the study of the Reelin pathway, where authors demonstrate that mutation of either Crk or CrkL did not compromise Reelin signaling, indicating their overlapping roles[142]. Morover, studies using shRNA against the Crk family adaptors showed that knockdown of both proteins are required for defects in neuronal positioning during embryogenesis[143].

Peterson at al. have shown that CrkL is not absolutely required for T cell development or function[144].

However, in contrast to the above observations showing overlapping roles for both proteins, other studies suggest that CrkII and CrkL play distinct roles during embryonic development. The 'knockout' of murine *Crk* or *CrkL* genes resulted in cardiac and skeletal development and partial lethality[145], suggesting that each of these two genes is essential and cannot by replaced by the other. In very elegant study, Imamoto and colleagues showed that mice homozygous for a null mutation of CrkL exhibit defects in multiple cranial and cardiac neural crest derivatives[146]. Finally, from a very recent study by Yanagi et al., it appears that in HSC-3 (head and neck squamous cell carcinoma cell line) cells the dominant binding partner for C3G was CrkL, nor Crk[147].

How multiple species of Crk family adaptors differ in their activity still remains an active area of investigation.

1.1.8 Crk and diseases

1.1.8.1. Crk and human cancers

A vast number of studies have shown that Crk is overexpressed in many human cancers. A study by Miller at al. has provided evidence that CRKI and CRKII are involved in the development and dissemination of lung adenocarcinomas[148]. Genomic and functional analysis presented by Kim et al. defines a role of *CRKL* amplification in NSCLC[149]. Breast cancer also revealed strong correlation between Crk phosphorylation and cancer staging. In a report by Rodrigues et al., CrkI and CrkII protein levels were eleveted by 60% when analyzing primary breast tumors[150]. Phosphorylation of CrkII on Tyr221 by Abl was shown to be a major factor for induction of apoptosis and suppression of migration of the human malignant breast cancer cell line MDA-MB-435c by EphrinB2[24]. Crk was also found in glioblastoma and several sarcomas[151,152]. For example a report by Takino et al. found high expression levels of CrkI, but not CrkII, by analyzing samples of tissues from human glioblastomas[58]. Ovarian cancer is another aggresive tumor in which Crk has been shown to be overexpressed. Studies by Linghu *et al*., suggest that the Crk is a key component of focal adhesion and is involved in cell growth, invasion, and dissemination of an MCAS (mucinous cystadenocarcinoma) human ovarian cancer cell line derived from an ovarian cancer patient.[153] The Crk family adaptors also appear to be highly expressed in myeloid cells[154]. Notably, CrkL was shown to be involved to a greater extent than

Crk. Indeed, the interaction of CrkL with Bcr-Abl and its high tyrosine phosphorylation on Tyr207 is considered a hallmark of Bcr-Abl activity in CML[154]

1.1.8.2. Crk and bacterial diseases

Over the last decade there were many reports suggesting that Crk may be involved in bacterial pathogenesis. The first report, which demonstrated the role for Crk in bacterial uptake pathways, came from Bouton and collegues[155]. Their study demonstrates that the formation of CAS/Crk complexes and subsequent Rac1 activation are important elements of the process of *Yersinia pseudotuberculosis* uptake into human epithelial cells. In studies by Pendergast and collegues, Crk was shown to be implicated in *Shigella flexneri* infection[156]. This study showed that the activation of Rho during *Shigella* internalization is preceded by the activation of the Abl family kinases, and that Crk phosphorylation and Crk Y221F mutation completely blocks *Shigella flexneri* infection. The novel role for Alb and Crk phosphorylation was also found in the *Pseudomonas aeruginosa* internalization pathway[157]. Another study on CagA protein, which is a major virulence factor of *Helicobacter pylori*, shows that the phosphorylated CagA forms a complex with Crk proteins and this interaction play a major role in *H. pylori* infection[158].

1.2 Nuclear Magnetic Resonance (NMR) spectroscopy

1.2.1 NMR as a powerful tooll in biochemistry

Nuclear magnetic resonance (NMR) is, next to X-ray crystallography, the only biophysical method that can be applied to the study of three-dimensional molecular structures of biomolecules such as proteins, nucleic acids and their complexes at atomic resolution[159]. NMR has been widely used as a major structural biology tool to study biomolecular structure, dynamics and function[160] [161]. In contrast to other methods, NMR spectroscopy is able to investigate chemical properties of molecules by studying individual nuclei. The conditions used in NMR are very close to the conditions found in the cell. Moreover, NMR can be used to study flexible parts of a protein. In general the flexible linkers could interfere with crystallization so cannot by studied by X-ray crystallography[162]. Partially folded proteins which are difficult to crystallize (because they can not adopt one unique 3D structure but fluctuate over an ensemble of conformations) can be nicely characterized by NMR[163]. NMR spectroscopy since the first structure determination[164] has drastically changed. Major improvements in NMR hardware (magnetic field strength, cryoprobes) and NMR methodology, combined with the availability of molecular biology and biochemical methods for preparation and isotope labeling of recombinant proteins have dramatically increased the use of NMR for the characterization of structure and dynamics of biological molecules in solution. These improvements

are on-going and are designed to overcome the two main problems with NMR analysis of biomacromolecules, namely signal to-noise ratio and spectral overlap. Higher magnetic field strengths and cryoprobes combined with specific isotopic labeling schemes have enabled to study larger and more complex biological systems[165].

1.2.2 ^{15}N-HSQC: A Protein's fingerprint

Heteronuclear Single Quantum Coherence (HSQC) is a simple 2D NMR Spectrum, and usually the first experiment performed on proteins that generates cross peaks correlating ^{1}H and ^{15}N[166]. Magnetization is transferred from hydrogen to the attached ^{15}N nuclei via J-coupling. After the chemical shift evolves on nitrogen the magnetization is back transferred to the hydrogen for detection. An HSQC spectrum is sometimes called a "fingerprint" of the protein because each peak on the spectrum corresponds to the NH of an individual residue (Fig. 18).

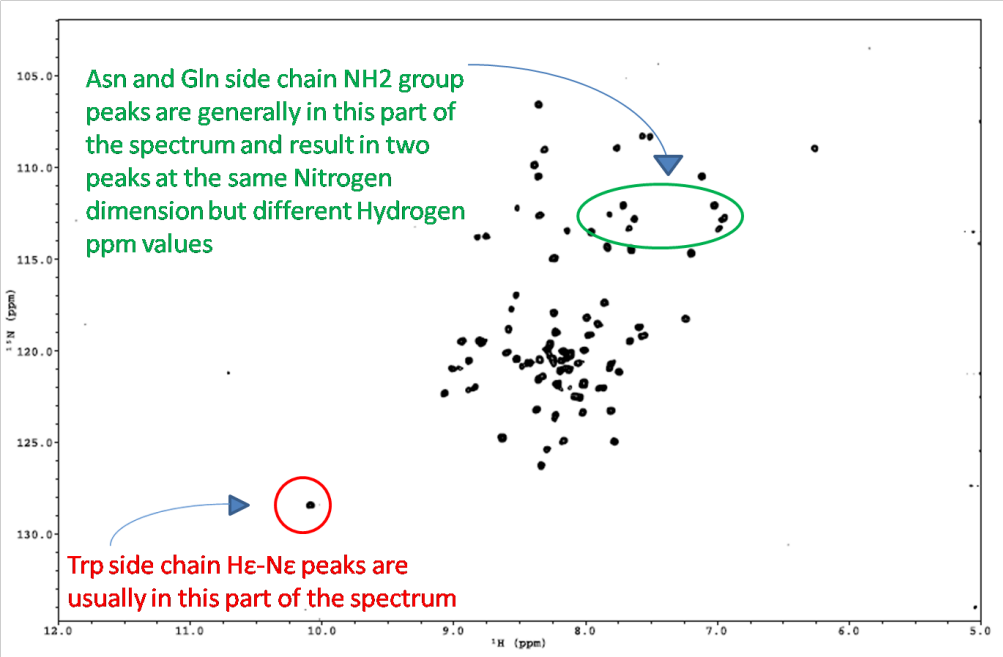

Figure 18. HSQC spectra. The number of peaks on an HSQC spectrum will correspond to the number of residues, excluding pralines, which lack an amide proton attached to nitrogen in the peptide bond. The Trp side-chain Nε-Hε groups and Asn/Gln side-chain Nδ-Hδ2/Nε-Hε2 groups are also visible.

HSQC is a powerful technique to quickly map the interacting sites of a protein and its ligand[167]. The observed change in the chemical shift in the HSQC of a complex compared to the HSQC of an uncomplexed protein will give information on the binding interface and the conformational changes of the protein.

1.2.3 Protein assignment

1.2.3.1 Protein backbone assignment

To investigate the structure and dynamics of the protein of intrest, it is necessary to perform assignment of the resonance peaks in the spectrum. The most common strategy for resonance assignments is the use of triple resonance (3D)

experiments (^{1}H, ^{13}C, ^{15}N)[168]. This requires uniform labeling of the protein under investigation with ^{13}C and ^{15}N isotopes by using ^{13}C glucose and ^{15}N ammonium. The 3D experiments are based on the ability to transfer magnetization between NMR active nuclei through J couplings (Fig. 19)

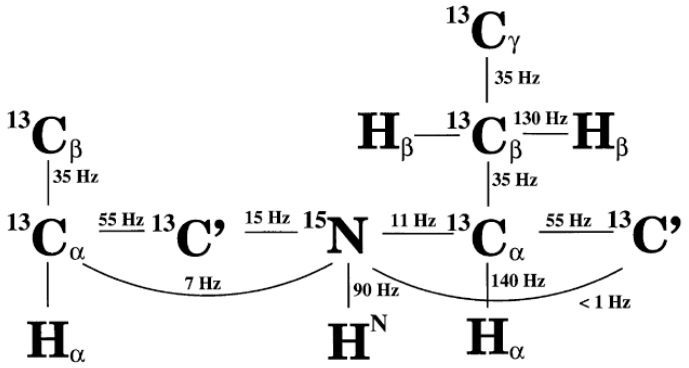

Figure 19. Spin system of the peptide backbone and the size of the ^{1}J and ^{2}J coupling constants that are used for magnetization transfer in ^{13}C-, ^{15}N-labelled proteins.

The assignment protocols that are routinely used are based on the 3D sets of HNCA with HN(CO)CA (Fig. 20), HNCACB with CBCA(CO)NH and HNCO with HN(CA)CO. Each set of experiments correlates the chemical shifts of the amide group (NH) of a spin system with both inter- and intra- residue chemical shifts of main chain (CO, Cα, N). For example, in the HNCA experiment (Fig. 20A), the magnetization is transferred from ^{1}H to ^{15}N and then via the N-Cα J-coupling to the $^{13}C\alpha$ (i-1) and then back again to ^{15}N and ^{1}H hydrogen for detection. Similarly in HN(CO)CA experiment (Fig. 20B), magnetization is transferred from ^{1}H to ^{15}N and then to ^{13}CO. From here it is transferred to $^{13}C\alpha$ (i-1) and the chemical shift

is evolved. The magnetization is then transferred back via ^{13}CO to ^{15}N and ^{1}H for detection

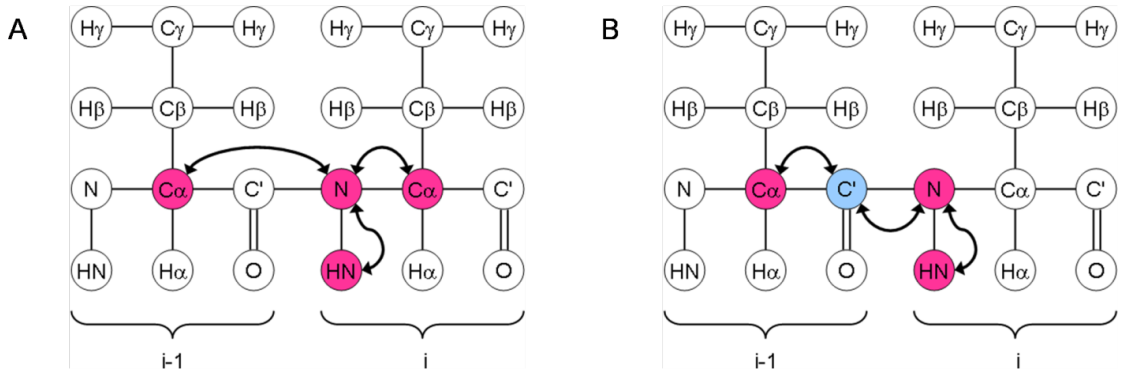

Figure 20. Magnetization transfer through bond in HNCA (a) and HN(CO)CA (b) triple resonance experiments. Red circles indicate the recorded chemical shifts and blue circled atoms mediate the transfer of magnetization but their chemical shift is not recorded

1.2.3.2 Protein side-chain assignment

Most NMR structure determination techniques use ^{1}H-^{1}H NOE distances as the main source of constraints to elucidate the high-resolution structure of a target protein. Therefore, an almost complete set of both backbone and side-chain resonance assignments are required to assign inter-proton NOE distance restraints from NOESY spectra. The assignment of the ^{1}H and ^{13}C side-chain resonances can be achieved by a combination of 3D H(CCO)NH, CC(CO)NH, ^{15}N and ^{13}C NOESY-HSQC experiments[169]. The H(CCO)NH and CC(CO)NH are routinely used to correlate the ^{1}H and ^{13}C side-chain atoms of the residue i-1 with the amide of the residue i (Fig. 21)

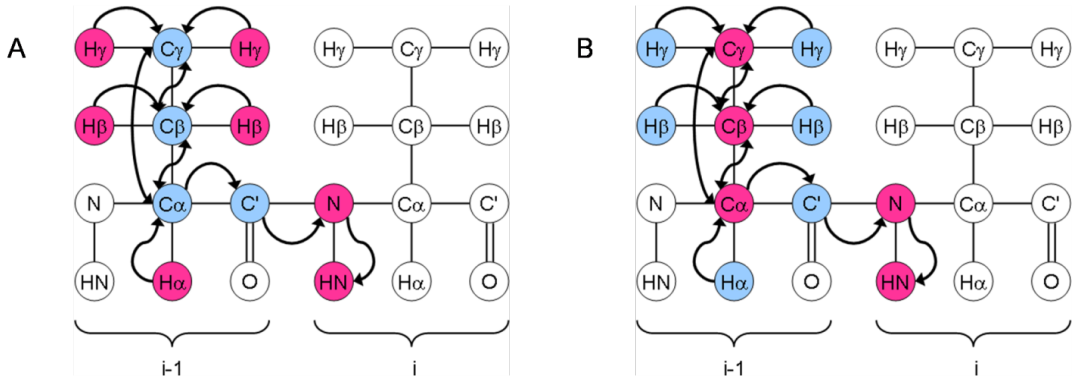

Figure 21. The magnetization transfer pathway in 3D H(CCO)NH (a) and CC(CO)NH (b) experiments. Red circles indicate the recorded chemical shifts and blue-circled atoms mediate the transfer of magnetization but their chemical shift is not recorded

In the H(CCO)NH experiment[170], the magentization is transferred from the side-chain ^{1}H nuclei to the attached ^{13}C nuclei. Then isotropic ^{13}C mixing is used to transfer magnetisation between the carbon nuclei. Later, magnetisation is transferred to the ^{13}CO, on to the amide nitrogen and finally the amide hydrogen for detection. The chemical shift is evolved simultaneously on all side-chain hydrogen nuclei, as well as on the amide nitrogen and hydrogen nuclei, resulting in a 3D spectrum with two proton dimensions and one nitrogen dimension. The CC(CO)NH experiments have similar mechanisms to those of H(CCO)NH, except that the observed nuclei are ^{13}C[171].

1.2.4 Nuclear Overhauser Effect Spectroscopy (NOESY)

The Nuclear Overhauser Effect (NOE) is a common tool within molecular biology used for the generation of distance restraints in structure calculations. Despite the large amount of work done on other structure restraints development like

RDC or PRE, NOE is still the major source of structural information, providing distance information between two protons close in space (usually within 5-6 Å)[172]. The NOE arises from dipole–dipole relaxation between two spin-½ nuclei. The cross-relaxation between two spins is distance-dependent and the rate of relaxation that produces the NOE is proportional to r^{-6}, where r is the internuclear distance[173]. There are two commonly used NOE experiments: ^{15}N-NOESY-HSQC[174] and ^{13}C-NOESY-HSQC[175]. In both ^{15}N and ^{13}C NOE experiments the magnetisation is exchanged between all hydrogens using the NOE. The magnetisation is then transferred to neighboring ^{15}N or ^{13}C nuclei and back to ^{1}H for detection. In addition, 3D HMQC-NOESY-HMQC experiments can be used for distance restraints in the case of heavily overcrowded spectral regions[176].

1.2.5 Protein deuturation

The application of NMR to study proteins larger than 30 kDa has been an ongoing challenge[177]. Larger molecules give poor-quality spectra because of the resonance line's tendency to broaden. Line broadening is the result of fast transverse relaxation (T2) due to the slower tumbling of large molecular weight molecules (Fig. 22)

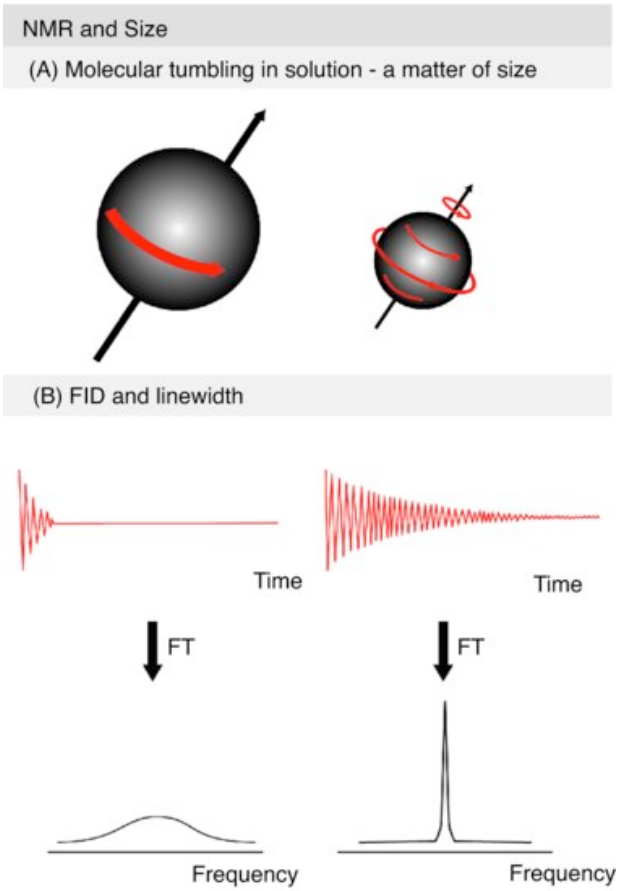

Figure 22. Molecular weight dependent tumbling.

The line width of the spectrum is proportional to T2 ($\Delta v = 1/\pi T2$)[178]. The major source of relaxation in proteins comes from the high density of protons present, which results in

proton-proton dipolar relaxation. By exchanging the protons to the deuterons, the T2 relaxation is reduced; hence the sensitivity and the resolution are significantly improved[179]. However, deuteration results in complete elimination of protons, which are the source of NOE. Therefore, it is often necessary to use site-directed protonation of uniformly deuterated proteins at specific positions, either by introducing the specific amino acids (e.g., Phr, Tyr, Lys)[180] or selectively protonating CH3 groups of Val, Leu, Ile and Ala[181,182].

1.2.6 Paramagnetic Relaxation Enhancement (PRE)

The NOE is the main source of the geometric information used to further calculate structures. However, the structures of large molecular weight proteins remain a challenge due to the fact that proteins must be deuterated, leaving most of the side-chain protons for NOE analysis is inaccessible[183]. To supplement the NOE data, distances derived from paramagnetic relaxation enhancement (PRE) were employed[184]. The PRE effect is based on the magnetic dipolar interaction between the nucleus (e.g., 1H) and the unpaired electrons of the paramagnetic center. For an electron-nucleus distance r, the magnitude of the paramagnetic effect is proportional to r^{-6}. Because of the large magnetic moment of the electron, the observed PRE effects are also large, providing long-range distance restraints up to ~35Å[185]. MTSL, which can be covalentely attached to the Cys side-chain, has been used successfully as a source of PRE (Fig. 23). Research on eIF4E (25 kDa translation initiation protein) showed that the distance-dependent line broadening was consistent with the three-dimensional structure of this protein for all spin label substitutions[184].

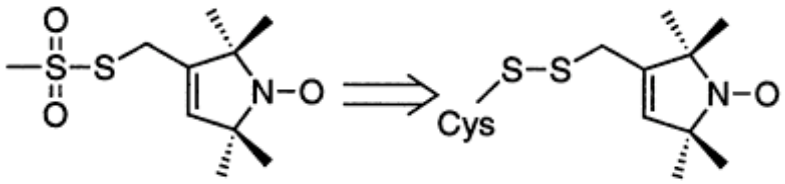

Figure 23. Chemical structure of the spin label reagent MTSL (left) and covalently bound to Cys (right)

The PRE quantification can be done by comparing two ^{15}N-HSQC experiments, one in the paramagnetic state and one in the diamagnetic state. The relaxation rates can be calculated from the HSCQ experiment using the cross peak intensity ratio from both states using following formula:

$$\frac{I_{ox}}{I_{red}} = \frac{R2\ \exp(-R2^{sp}t)}{R2 + R2^{sp}}$$

where t is the total INEPT evolution time of HSQC, R2 and $R2^{sp}$ are the intrinsic and spin contribution relaxation rates, respectively. The distances between the nuclei of interest and a paramagnetic center can be calculated as follows:

$$r = \left[\frac{K}{R2^{sp}}\left(4\tau_c + \frac{3\tau_c}{1+\omega_h{}^2\tau_c{}^2}\right)\right]^{1/6}$$

where r is the distance between the electron and nuclear spins, τ_c is the correlation time for the electron-nuclear interaction, ω_h is the Larmor frequency of the nuclear spin (proton), and K[186] is the 1.23 x10^{-32} cm^6 s^{-2}.

The PRE method has been successfully applied to study of unfolded and partially unfolded states[187], analysis of protein-protein[188] , mambrane proteins[189] and also to ivestigate the nature of, low population transient intermediates[190].

1.2.7 Residual Dipolar Coupling (RDC)

The mid 1990s brought the first reports to demonstrate the use of residual dipolar couplings (RDC) in the analysis of biomolecular structures [191]. Since the

recognition of the possible advantages of RDCs in protein structure detrmination, applications have been extended to nucleic acid structure[192], protein-ligand interactions[193], and studies of protein motions[194].

The relative orientations of interacting biomolecules, or separate domains of a large biomolecule, can be derived by measuring residual dipolar couplings (RDCs) between NMR active nuclei in partially oriented media[161]. Dipolar coupling is an interaction through space, which provides information about on how each dipole is related to the molecular coordinate frame and how the dipoles are related to each other. In solution, the dipolar couplings average to zero due to Brownian motion effects. As a result, they can be only observed under anisotropic conditions. For a pair of dipole-coupled nuclei, A and B, the observable dipolar coupling in solution, D_{AB}, can be expressed as:

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

and`

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

where A_a^{AB} and R are the axial and rhombic components, respectively, of the molecular alignment tensor; θ is the angle between the internuclear bond vector and the z-axis of the alignment tensor; φ is the angle between the projection of the internuclear bond vector onto the x-y plane and the x-axis; μ_0 is the

permeability in vacuum, *h* is Planck's constant (6.626*10-34 J*s); S is the generalized order parameter; γ_A, γ_B are the gyromagnetic ratios of two nuclei; *r* is the time-averaged internuclear distance; and A_a is the aforementioned axial component of the molecular alignment tensor.

To obtain RDCs in solution, the alignment solution is required to induce the partial alignment of protein molecules. Commonly used alignment media include, micelles[195], Pf1 filamentous phage[196] or polyethylene glycol (PEG)[197]. The dipolar couplings can be extracted from many available experiments. For example, N-HN dipolar couplings can be measured in a simple IPAP 15N-HSQC where spectra are recorded both in-phase and antiphase. Therefore Dipolar couplings are obtained from differences in the splittings measured between in the approach oriented and isotropic phases[198].

In order to use the RDC data for structure refinement, good estimates for A_a and R must be available. In the method demonstrated by Clore at al.[199], the measured RDCs are plotted on a histogram (Fig. 24).

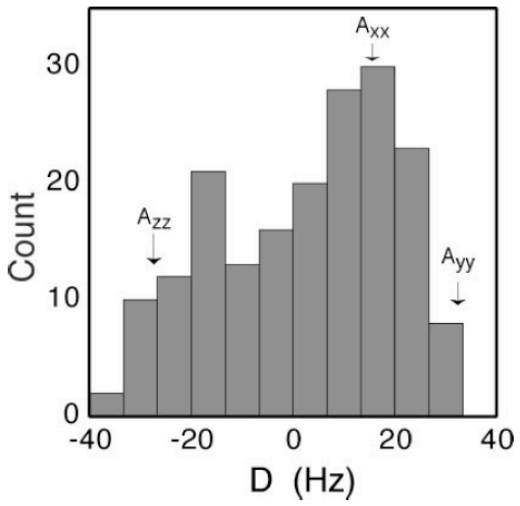

Figure 24. A histogram of RDC[200].

Values from the three exrema (Azz, Ayy, and Axx) can be used with the following equations to calculate Aa and R

$$\text{A-B lies along } D_{zz}: \theta = 0^\circ \quad A_{zz} = 2A_a,$$

$$\text{A-B lies along } D_{yy}: \theta = 90^\circ, \phi = 90^\circ \quad A_{yy} = -A_a\left\{1 + \frac{3}{2}R\right\}$$

$$\text{A-B lies along } D_{xx}: \theta = 90^\circ, \phi = 0^\circ \quad A_{xx} = -A_a\left\{1 - \frac{3}{2}R\right\}$$

1.2.8 Protein dynamics by NMR

Proteins have been shown to have very dynamic personalities[201]. Protein dynamics play an essential role in catalysis, ligand binding, molecular recognition and allostery[202]. NMR spectroscopy is an excellent tool to investigate the protein dynamics over a wide range of motional time scales (Fig. 25)

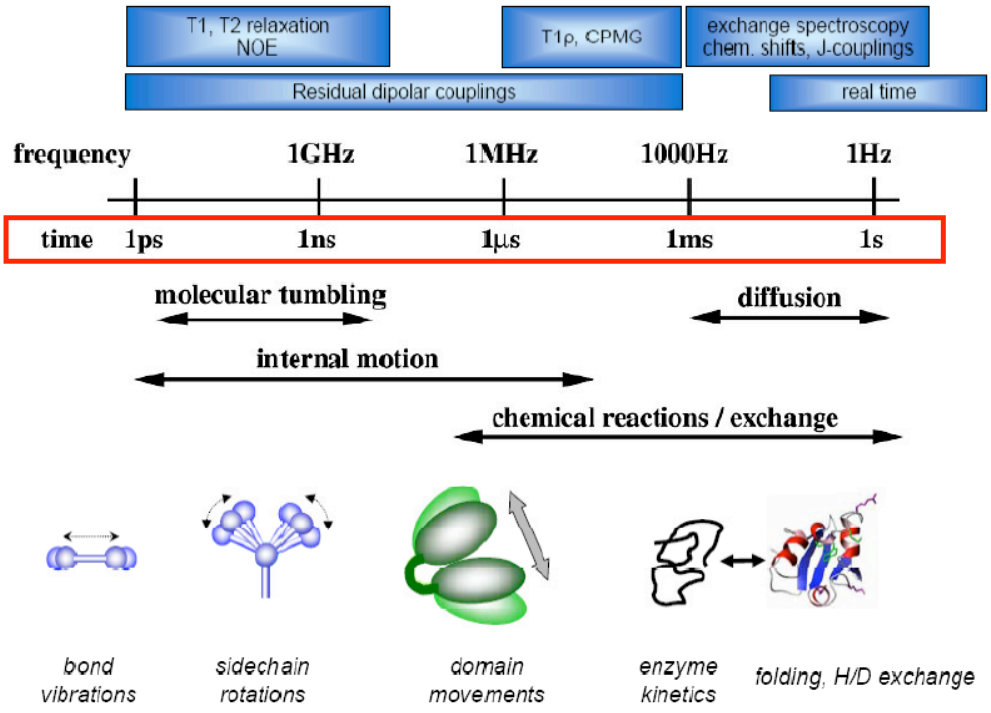

Figure 25. NMR time scales and dynamics in biology[178].

NMR can describe a great number of biological processes, from rapid bond fluctuations in the pico- to nanosecond range (pn-ns) to slower collective motions in the micro- to millisecond (μs-ms) range.

1.2.8.1 Fast protein dynamics (ps-ns)

There are great numbers of examples proving NMR as an excellent method to characterize molecular motions[202]. The longitudinal relaxation rate (R1), the transverse relaxation rate (R2), and the heteronuclear 15N-Nuclear Overhouse Effect (HetNOE) are the most commonly measured NMR parameters for decribing the dynamics on fast timescales (ps-ns). Using these three different relaxation components one can calculate the order parameter, S2 , which is used to measure the magnitude of the angular fluctuation of a chemical bond vector

(such as that of the NH bond) in a protein [203]. The value of S2 ranges from zero to one, where zero represents a bond vector rapidly sampling multiple orientations and one represents a bond without internal motions. R2 is also sensitive to motions on μs-ms times scale, however only the ps-ns motions are related to S2.

1.2.8.2 Slow protein dynamics (μs-ms)

Many biological processes occur also on the μs–ms timescale, including protein folding, substrate binding, and catalysis. This necessitates the need to study slower timescale motion[204]. This can be monitored by measuring R2 as a function of the effective radiofrequency field strength using using spin-lock (R1ρ) or Carr-Purcell-Meiboom-Gill (CPMG) methods[203].

1.2.9 Förster resonance energy transfer (FRET)

Energy transfer is essential for biological reactions. The general scheme of these processes is as follows (15):

$$D^* + A \rightarrow D + A^* \qquad (15)$$

Förster resonance energy transfer (FRET) is a spectroscopic process by which energy is passed nonradiatively between molecules. When the donor chromophore (D) is excited by incident light and its emission spectrum overlaps with the absorption wavelength of an acceptor (A), the excited state energy from the donor may transfer to the acceptor through dipole-dipole coupling over distances 10-100 Å (Fig. 26)[205].

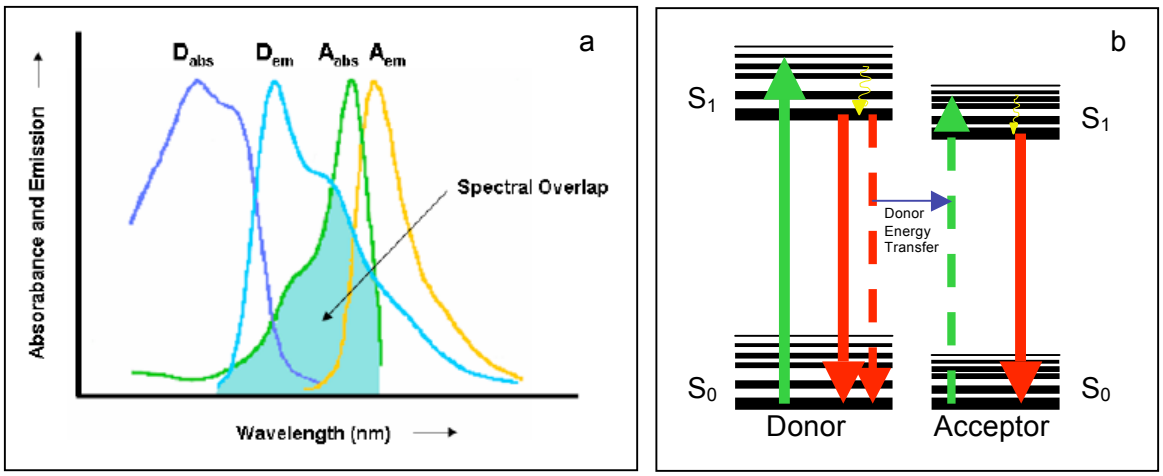

Figure 26. (a) Absorbance end emission spectrum of donor, acceptor and D-A spectral overlap region in FRET[206] (b) Fluorescence resonance energy transfer Jablonski diagram.[207]

As a consequence of this process the donor fluorescence intensity and excited state lifetime decrease, and the acceptor becomes electronically excited. If the acceptor is fluorescent it may then emit a photon of light.

FRET can also provide a new perspective on the conformations, dynamics and interactions of proteins, since allow measuring the distance between the dyes attached to molecules.[205]

The rate of the energy transfer is determined by the distance between the donor and acceptor, and the extent of spectral overlap[208].

$$k_T(r) = \frac{1}{\tau_0}(R_0/r)^6$$

R_0 is the Förster radius, ranges from 20 to 80Å for most of dye pairs and is the distance at which half of the energy is transferred (Fig 27). It depends on the orientation factor of the dye κ^2, the donor quantum yield Φ_D, the D-A spectral overlap *J* and refractive index *n* of the medium[209].

$$\mathbf{Ro\ [Å] = (8.79 \times 10^{23}\ \kappa^2\ \Phi_D\ \mathit{J}\mathit{n}^{-4})^{1/6}}$$

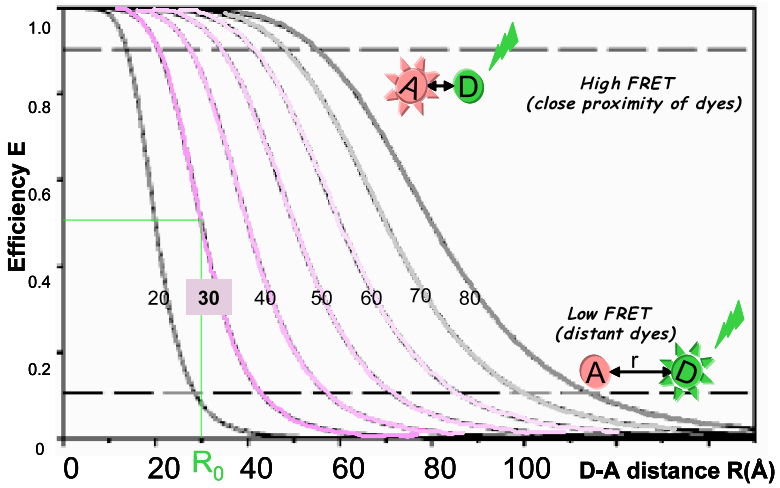

Figure 27. Dependence of dynamic range of FRET on the Förser radius R_0. The dotted lines delineate the regime of maximum sensivity for each pair with differenf R_0.[210]

The efficiency of energy transfer (E) depends on the donor-acceptor separation distance, r, with an inverse 6th-power law:

$$E = \frac{R_0^6}{(R_0^6 + r^6)} = \frac{k_T}{k_r + k_{nr} + k_T} = 1 - \frac{\tau}{\tau_0}$$

where τ_0 and τ is the donor lifetime in the absence or in the presence of acceptor, respectively ($\tau_0=1/(k_r+k_{nr})$; $\tau=1/(k_r+k_{nr}+k_T)$) and k_r and k_{nr} are the rate coefficients for radiative and non-radiative deexcitation.

The FRET efficiency value, E, is usually calculated for individual molecules from a lifetime or intensity measurement of the donor (I_D) and the acceptor (I_A) as follow:

$$E = \frac{I_A}{(\gamma I_D + I_A)}$$

where, γ is a correction parameter for different detection efficiencies of the two channels (η) and quantum yields of the dyes (Φ)[211].

$$\gamma = \frac{\eta_A \Phi_A}{\eta_D \Phi_D}$$

1.2.10 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is a technique used to measure the thermodynamics and the kinetics of binding interactions. The ITC is able to measure heat effects as small as 0.1 µcal and corresponding binding constants, K's, as large as 10^8–$10^9 M^{-1}$. Currently ITC is capable of detecting heat rates as small as 0.1 µcal/sec and precisely determining reaction rates in the range of 10^{-12} mol/sec.

An ITC instrument contains two identical cells made from a highly efficient thermal conducting material (usually gold) placed in an adiabatic jacket. The cell is controlled to have a constant temperature by applying constant cooling. The heating power from both the chemical reaction and the control heater are kept at a constant level to compensate for each other. The raw ITC signal is simply the power (µcal/sec) applied to the control heater to keep the calorimetric cell temperature constant. A typical ITC is shown in Fig. 28.

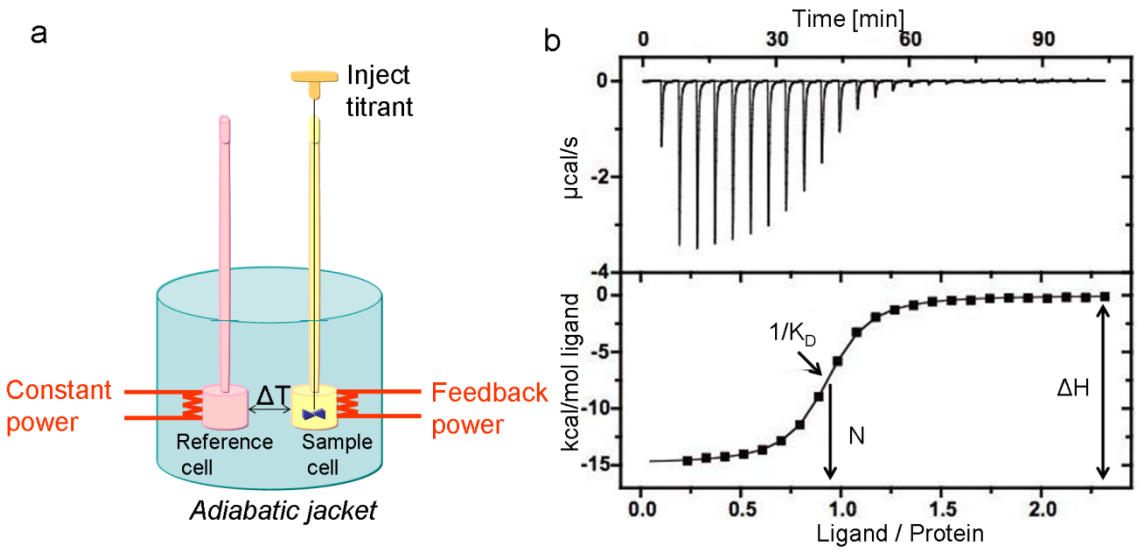

Figure 28. (a) Representative diagram of typical power compensation ITC. (b upper) Heat trace over time, each peak corresponds to an individual injection.(b lower) plot of the integration of the peak area versus the molar ratio of ligand and protein

To perform an ITC experiment, first a protein solution of known concentration is placed in the sample cell. The reference cell contains buffer or water without protein. Before the ligand titration starts, a constant power is applied to the reference cell, which activates the heater located on the sample cell. Upon the injection of the titrant ligand into the sample cell, heat is taken up or evolved depending on whether the binding process is endothermic or exothermic.

ITC enables the direct determination of the enthalpy change (ΔH), the association constant (Ka) and stoichiometry (N) during one experiment. Since temperature (T) is held constant throughout the entire experiment the free energy (ΔG°) of the binding reaction, which determines the stability of any given biological complex, can be calculated by:

$\Delta G^0 = -RT\ln K_a$ where Kd = 1/Ka and R is the gas constant.

ITC directly measures ΔH°, so the change in entropy (ΔS°) can be determined by:

$$\Delta S^0 = (\Delta H^0 - \Delta G^0)/T$$

Quantification of these thermodynamic parameters reveals the energetics of the physical processes involved in the binding reaction.

Chapter 2. Research Outline

CrkII and CrkL are prototypical members of a family of SH2 and SH3 domain containing adaptor proteins that participate in the co-ordinate assembly of protein-protein complexes in signal transduction. They share high sequence homology within their modular domains and have been shown to have overlapping binding partners. However, they seem to have distinct non-overlapping physiological roles as highlighted by the fact they are both required for embryonic development. Also, CrkL has been shown to be a preferred substrate for BCR-ABL, the oncogenic fusion protein that drives Chronic Myelogenous Leukemia (CML). The present study sheds light on the specific domain organization differences between the two adaptor proteins. The goal of the current research is as follows:

1) Structually delineate the organization of the modular SH2 and SH3 domains in CrkL.

Lack of structural information on CrkL has been a hurdle in understanding specific differences in signaling properties between CrkII and CrkL. We have obtained the structure of CrkL by NMR, which reveals unique features of CrkL in terms of the organization of its modular domains. Strikingly, the binding pocket of the SH2 domain of CrkL is partially occluded by the $SH3^N$ whereas the PPII site on the $SH3^N$ is freely accessible. Further, unlike CrkII, the $SH3^C$ does not participate in autoinhibition of the $SH3^N$.

2) Obtain structural insights into changes in domain organization upon phosphorylation and autoinhibition of CrkL.

Similar to CrkII, upon phosphorylation at the negative regulatory Tyr207, the SH2 domain of CrkL binds intramolecularly to phospho-Tyr207. However, there is little effect of this re-organization on the $SH3^N$ and $SH3^C$ domains of CrkL. The $SH3^N$ domain remains freely accessible to PPII ligands such as DOCK180 or Abl upon phosphorylation at Tyr207, which is strikingly different from what has been shown for CrkII. Importantly, this could provide an explanation for why CrkL is a preferred substrate for BCR-ABL.

3) Analysis of the dynamic features of CrkL/pCrkL.

The relaxation analysis demonstrates that the SH2 and $SH3^N$ domains of CrkL tumble as a unit, whereas the $SH3^C$ domain tumbles much faster and in an independent fashion. This shows that the SH2 and $SH3^N$ domains interact with each other, while the $SH3^C$ domain does not interact with any of the other domains. The present results indicate that $SH3^C$ has no role in regulating the activity of the SH2 and $SH3^N$ domains in CrkL, in sharp contrast to CrkII proteins, wherein $SH3^C$ was shown to act as a regulatory structural element by stabilizing the autoinhibitory conformation.

4) **Discern the role of CypA as a regulator of CrkII activity.**

CypA was shown to affect the kinetics of *cis-trans* isomerization of chicken CrkII, whereby the SH3C domain negatively regulates binding of ligands to the $SH3^N$ by physical capping of the $SH3^N$ domain. However, cis-trans isomerization has not been shown to regulate human CrkII in a similar manner. The present study unravels a novel mode of regulation of human CrkII actvity by CypA wherein CypA delays phosphorylation of CrkII at Y221 by binding to Gly219-Pro220. This partially negates autoinhibition of human CrkII thereby possibly maintaining SH2 and $SH3^N$ mediated protein complexes in cells.

Chapter 3. Structure and Dynamic NMR Studies of CrkL

3.1 Introduction

The members of the Crk family of adaptor proteins are important signaling molecules that function downstream of a wide number of receptors and regulate key cellular processes, including cell adhesion, motility, phagocytosis, differentiation, proliferation, transformation and apoptosis[1,2,3]. Crk proteins are implicated in many human cancers, including lung adenocarcinoma and glioblastoma[212], and prostate[213], ovarian[214], gastric[215] and breast cancers[24]. The Crk family consists of two alternatively spliced protein forms, CrkI and CrkII, that are expressed by a single gene (*CRK*)[216], and the Crk-like (CrkL) protein, which is expressed by a distinct gene (*CRKL*)[217]. The CrkII (304 residues) and CrkL (303 residues) proteins each consist of a single Src homology 2 (SH2) domain, an N-terminal Src homology 3 ($SH3^N$) domain and a C-terminal SH3 domain ($SH3^C$) (Fig. 29). The $SH3^N$ and $SH3^C$ domains are tethered by an approximately 50-residue-long linker, which contains a tyrosine residue (Tyr221 in CrkII and Tyr207 in CrkL) that is phosphorylated by the Abl kinase[120,218].

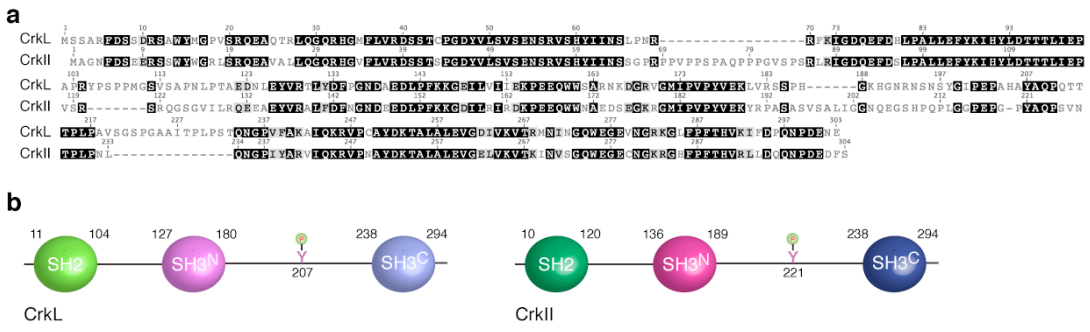

Figure 29. Sequence and domain organization of CrkL and CrkII. (a) Sequence alignment of human CrkL and human CrkII. Domain organization of CrkL and

CrkII, each consisting of one SH2 and two SH3 ($SH3^N$ and $SH3^C$) modular domains. Tyr207 in CrkL and Tyr221 in CrkII are phosphorylated by the Abl kinase.

Crk proteins link activated receptors to specific downstream signaling cascades using their SH2 and $SH3^N$ domains, which selectively bind phosphorylated tyrosine (pTyr)-Xaa-Xaa-Pro and Pro-Xaa-Leu-Pro-Xaa-Lys motifs (where Xaa is any amino acid), respectively[219,220]. A large variety of proteins have been identified as binding partners of Crk proteins[1,221]. In contrast, the $SH3^C$ domain lacks the binding determinants of typical polyproline II (PPII)-binding SH3 domains[19] and was shown to function as an autoregulatory element in CrkII[137,138,139] or even to promote certain signaling pathways[222].

CrkL has received particular attention primarily because it is a preferred substrate of Bcr-Abl[223,224], an oncogenic kinase that causes chronic myelogenous leukemia (CML)[225]. CrkL is indispensable for mediating the aberrant activity of Bcr-Abl[226,227] and is constitutively phosphorylated in human CML cells[228,229]. In fact, the level of CrkL phosphorylation is being used as a predictor of clinical outcome in patients treated for CML[230]. In addition to the relevance of CrkL to cancer, deletion of the *CRKL* gene causes cardiovascular and craniofacial defects resulting in embryonic lethality[231]. Notably, CrkL seems to be a rather unique adaptor protein, as it has been reported to function in the cytoplasm, the nucleus and the extracellular milieu[213].

Although CrkL and CrkII have been shown to compensate for each other's absence in certain cases[1], numerous studies have demonstrated that the two Crk

proteins have distinct, nonoverlapping functions[1,232,233]. Because of the high sequence identity between CrkL and CrkII (56%, up to 72% in the modular domains; (Figure. 30), it is hypothesized that the two proteins adopt very similar structures.

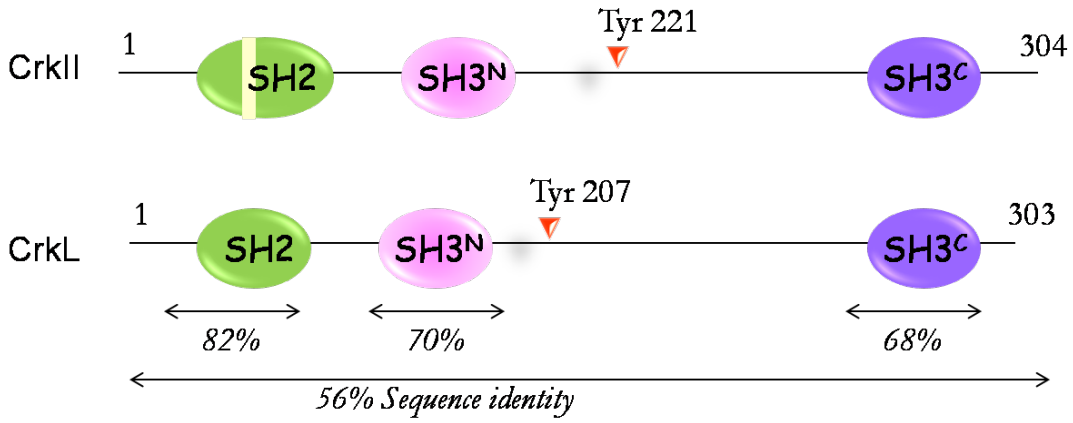

Figure 30. Sequence identity comparison between CrkL and CrkII.

Taking also into account the identical binding preferences of their SH2 and $SH3^N$ domains, it has been difficult to account for the different functional roles and signaling output of the CrkL and CrkII proteins.

Here we report the structure of CrkL in its unliganded and unphosphorylated state as well as in the phosphorylated (pTyr207) form (pCrkL). The data show that the SH2 and SH3 modular domains in CrkL are organized in a considerably different architecture than in CrkII[139,138]. (i) the pTyr-binding site of SH2 is partially masked in CrkL, whereas it is accessible in CrkII; (ii) the PPII-binding site of $SH3^N$ is accessible in CrkL but is completely occluded in CrkII; and (iii) the $SH3^C$ domain is mobile and does not interact with any of the other domains in CrkL, whereas it interacts extensively with the SH2 and $SH3^N$ domains in CrkII, thereby stabilizing its overall structure. We also show that, upon Tyr207

phosphorylation, the linker region interacts in an intramolecular fashion with the CrkL SH2 domain, thereby inhibiting the binding of pTyr ligands. Remarkably, this intramolecular association has little effect on the overall structure of CrkL, in sharp contrast to CrkII, wherein phosphorylation of Tyr211 results in $SH3^N$ autoinhibition. Notably, the present data demonstrate that CrkL forms a constitutive complex with Abl. Thus, despite the very high sequence identity of CrkL and CrkII, the structural organization of the two adaptors is substantially different and may account for their different functional roles as well as for the preference of Bcr-Abl to interact with CrkL rather than CrkII.

3.2 Results

3.2.1 Structure determination of CrkL and pCrkL

Human CrkL (~34 kDa) consists of 303 residues. Gel filtration and multi angle laser light scattering (MALLS) data show that the protein exists as a monomer in solution (Fig. 31).

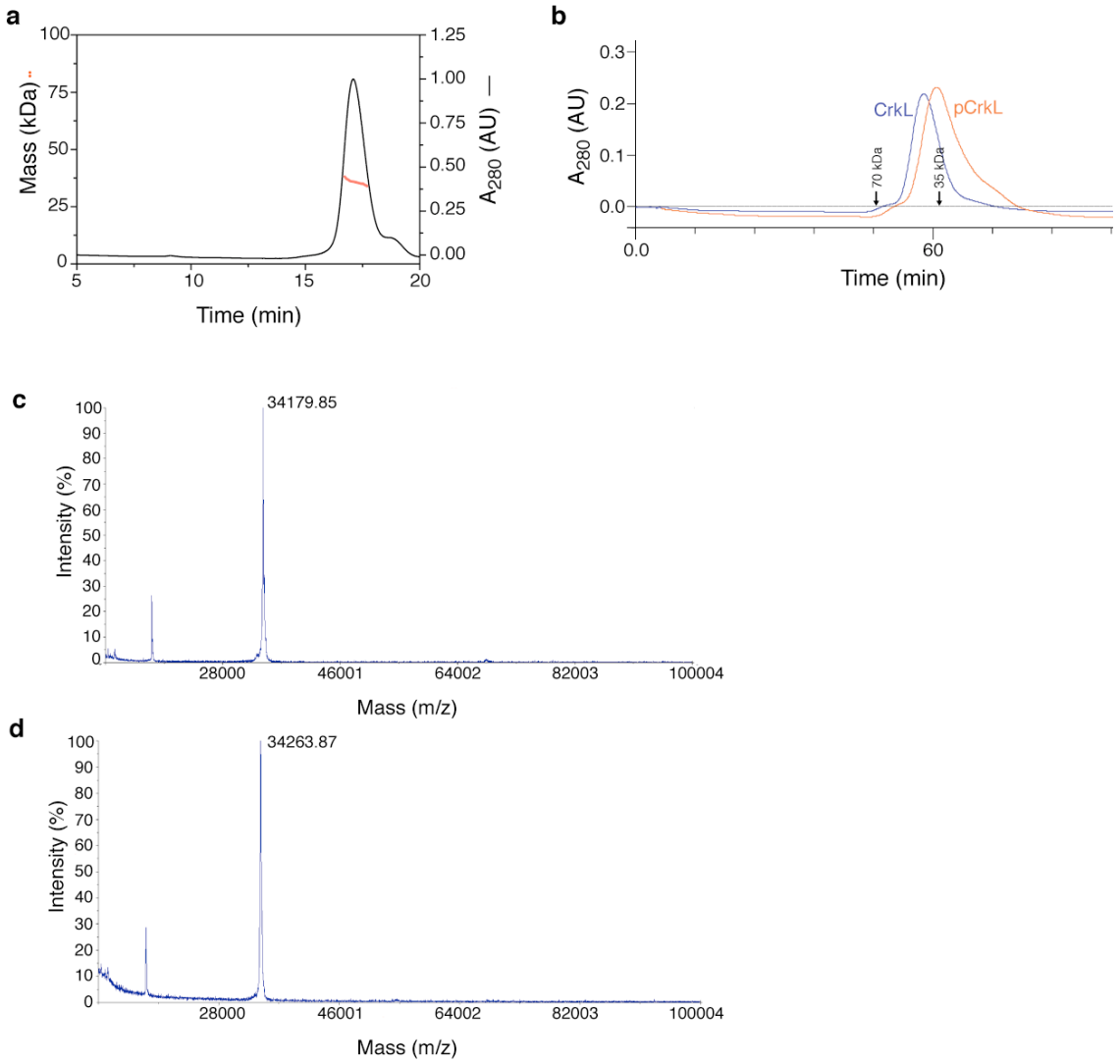

Figure 31. Molecular mass determination of CrkL and pCrkL. (a) Multi-angle laser light scattering (MALLS) of CrkL showing that CrkL is a monomer in solution. CrkL remains monomeric even at concentrations as high as ~1 mM. A Superdex 75 HR 10/30 column was used. (b) Sizeexclusion chromatography (Superdex 200) of CrkL and pCrkL showing that both proteins elute as monomers. Because of the intramolecular folding, pCrkL is a little more compact than CrkL and elutes slower than CrkL. (c) MALDI-TOF mass spectrum of CrkL. (d) MALDI-TOF mass spectrum of pCrkL. Only one site (Tyr207) is phosphorylated.

Despite the relatively large size of CrkL, the NMR spectra are of outstanding sensitivity and resolution (Fig. 33). Assignment of CrkL was facilitated by preparing and assigning the isolated SH2 and SH3 modular domains (Fig. 32).

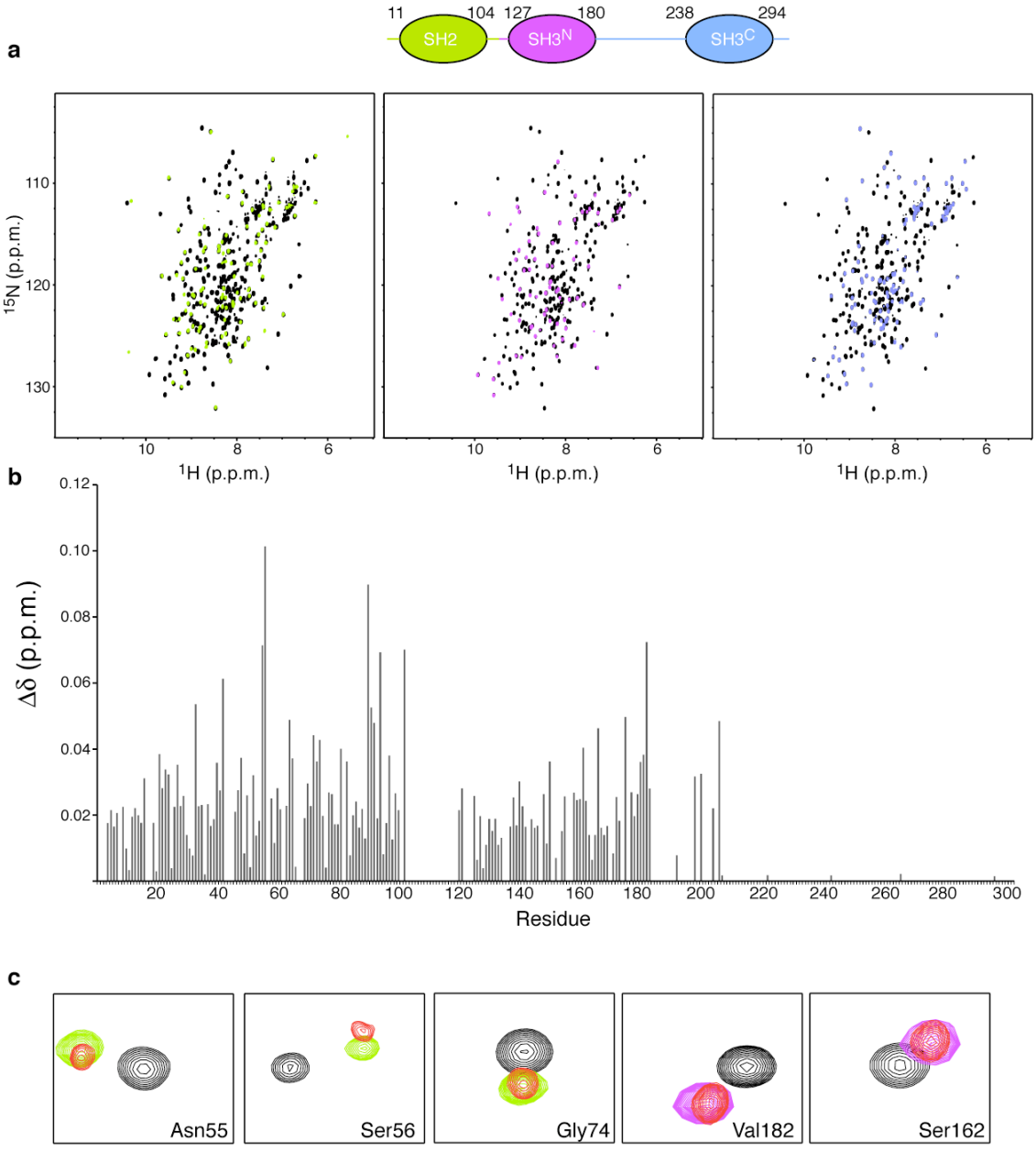

Figure 32. (a) ^{1}H-^{15}N HSQC NMR spectra of the three modular domains of CrkL overlaid with the spectrum of the full-length CrkL (black). The cross-peaks of the domains are colored using the color code in the schematic. Chemical shift analysis demonstrates that the SH2 and $SH3^{N}$ domains interact with each other,

whereas the linker-$SH3^C$ region does not interact with any of the other domains in CrkL. (b) Chemical shift difference (Δδ), plotted as a function of the residue number, between the isolated domains and full-length CrkL. (c) Overlaid ^1H-^{15}N HSQC spectra of select residues whose chemical shift is different in the full-length CrkL (black) and the isolated SH2 (green) and $SH3^N$ (magenta) domains. Chemical shift analysis shows that the D94K substitution (red cross-peaks) disrupts the SH2-$SH3^N$ interface in CrkL resulting in unrestricted binding of the pTyr-peptide to the SH2 domain of CrkL, as measured by ITC (Kd ~7 μM).

Notably, overlay of the NMR spectra of the isolated domains with that of full-length CrkL revealed substantial chemical shift differences only for a relatively small number of residues, located at the SH2 and $SH3^N$ domains, whereas no differences were observed for the $SH3^C$ domain (Fig. 32b,c). The pCrkL sample for NMR and structural characterization was prepared by adding catalytic amounts of the Abl kinase. Similarly to CrkL spectra, pCrkL's spectra are also of high quality, allowing for complete backbone and side chain assignment (Fig. 33).

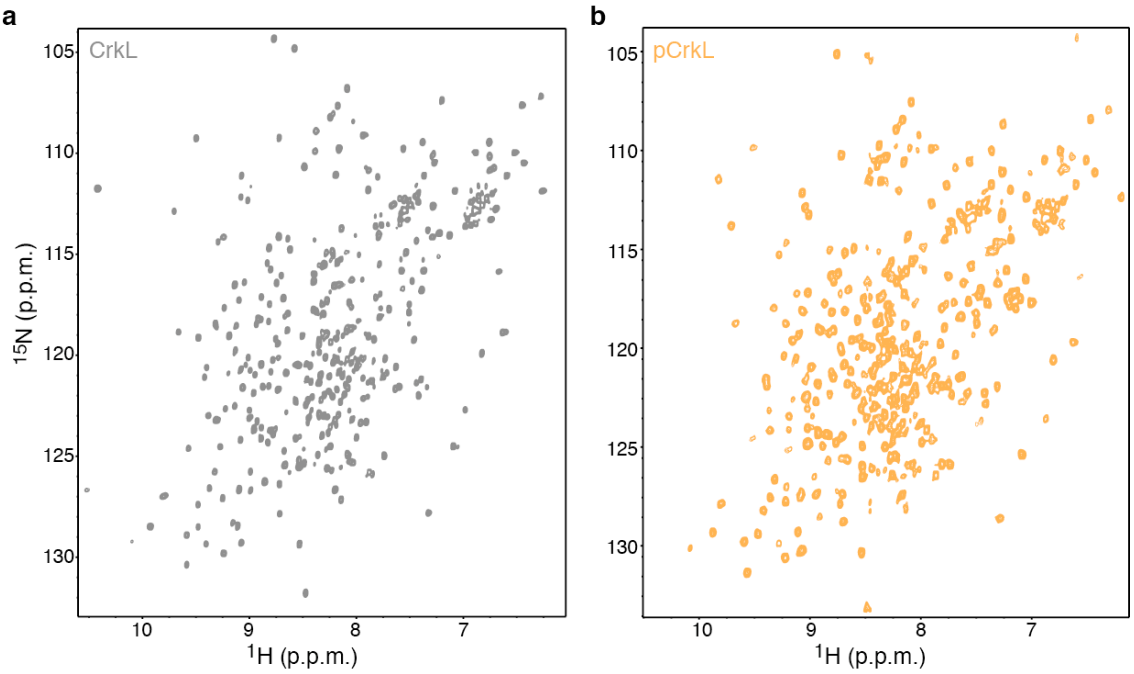

Figure 33. ^{1}H-^{15}N HSQC NMR spectra of (a) CrkL and (b) phosphorylated CrkL (pCrkL). Spectra were recorded at 32 °C.

The structures of both CrkL and pCrkL were determined by combining NOE, paramagnetic relaxation enhancement (PRE) and residual dipolar coupling (RDC) restraints.

Nitroxide spin labels (MTSL) 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 (Fig. 34d). ^{15}N-HSQC experiments were acquired at 32 °C at either 600 or 700 MHz (Fig. 34a,b,c). All experiments were repeated after the spin label was reduced with 2-3-fold excess ascorbic acid, adding <5 μL from a concentrated stock (dilution ~0.5%). Samples were placed in the magnet

at 32 °C and kept for 1 h to ensure complete reduction of the spin label. CrkL double mutant (C44S_C249S) was used as a negative controll to ensure that MTSL is not experiencing non-specific binding (Fig. 34c)

Distance-dependent line broadening consistent with the three-dimensional structure of SH2 domain (PDB 2EO3) of CrkL was observed for all spin-label substitution we have used.

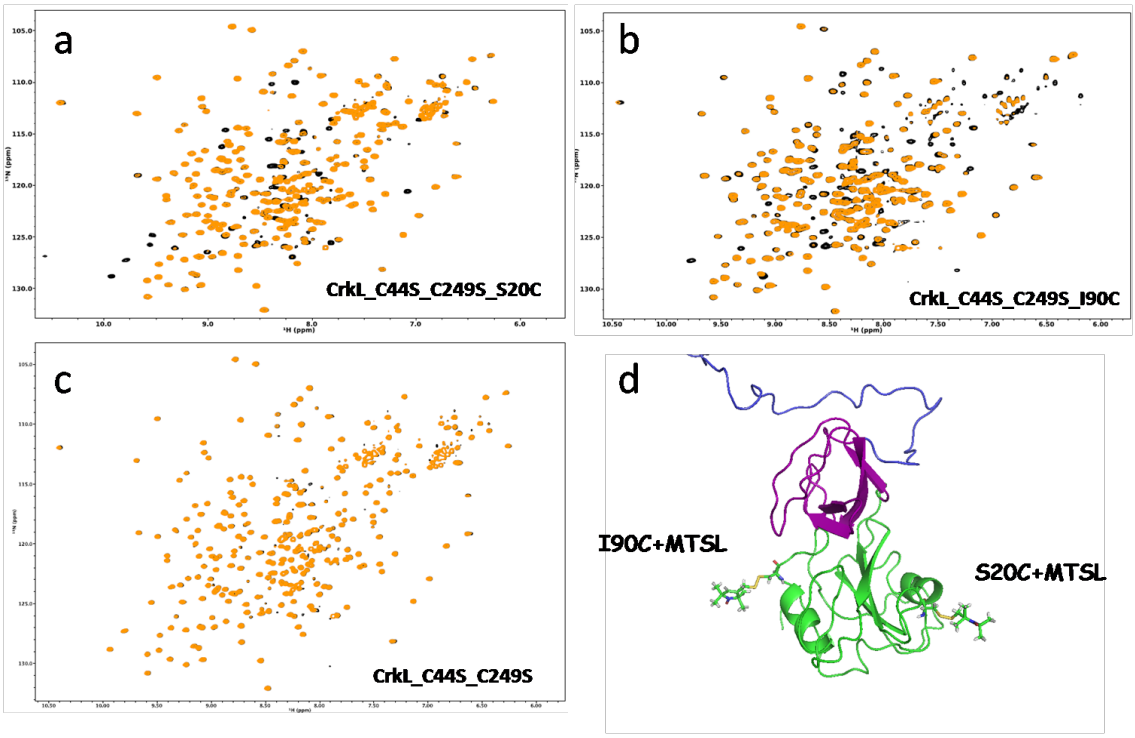

Figure 34. PRE experiments. ^{15}N-HSQC spectra of CrkL_C44S_C249S_S20C (a), CrkL_C44S_C249S_I90C (b), CrkL_C44S_C249S. In (a),(b) and (c) diamagnetic-black, paramagnetic-orange. (d) shows MTSL positions.

3.2.2 Structural architecture of CrkL

The lowest-energy structure of CrkL is shown in Figure 37a, and the conformational ensemble is shown in Figure 35a. The structures of the CrkL

individual SH2 and SH3 domains are very similar to those of the corresponding domains of the CrkII protein (Fig. 35).

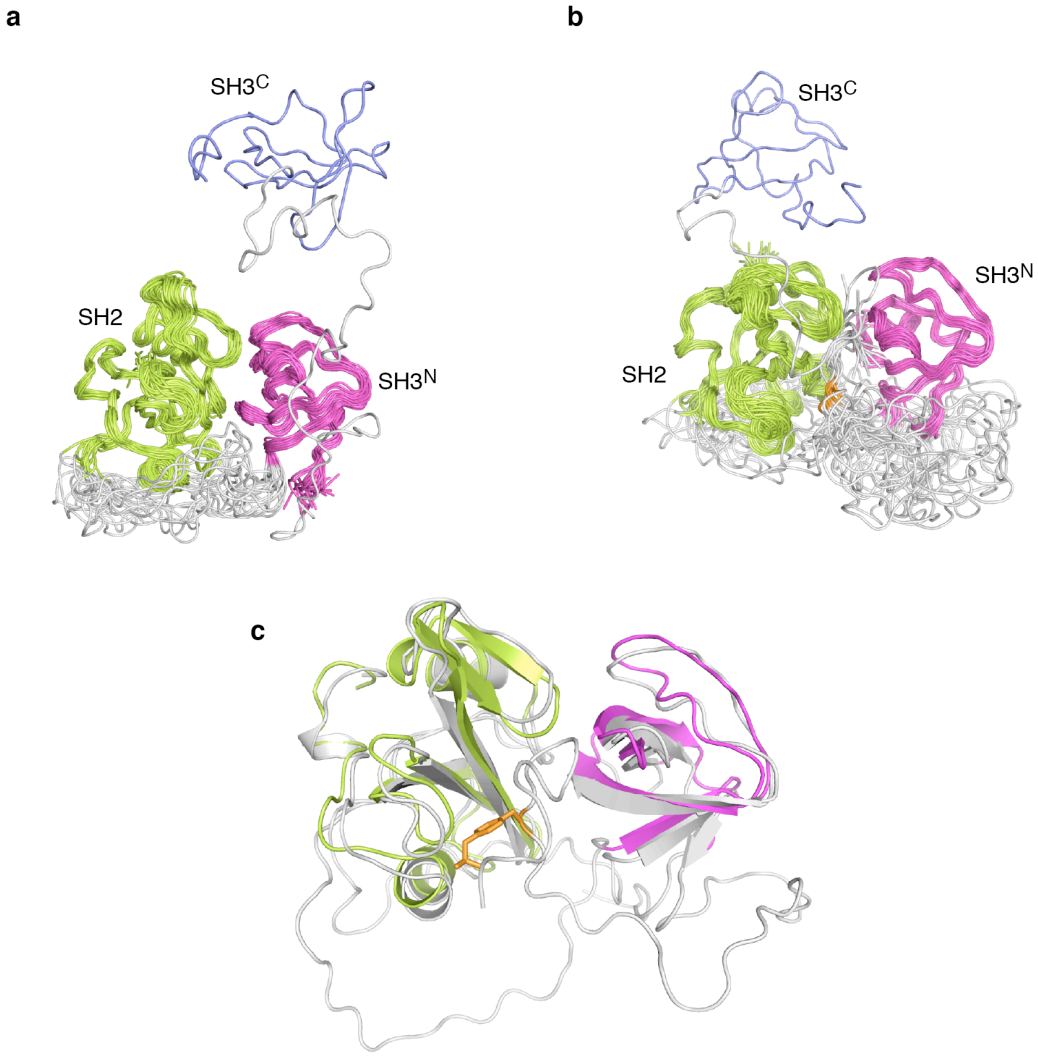

Figure 35. Overlay of the 20 lowest-energy conformers of (a) CrkL and (b) pCrkL. The linker-SH3^{C} region (188-303 in CrkL and 214-303 in pCrkL) is displayed only in one conformer as the mobility of this region results in very poor overlap. The pTyr207 region in CrkL is colored in orange. (c) Superposition of the CrkL (colored as in a) and pCrkL (colored grey) structures. Only the SH2-SH3^{N} region and the phosphorylated Tyr region are shown.

As expected on the basis of sequence conservation (Fig. 29a), the binding pockets of SH2 and $SH3^N$ are almost identical in CrkL and CrkII (Fig. 36), thus accounting for the similar recognition preferences of CrkL and CrkII. The CrkII SH2 domain has a 17-residue-long insertion between β-strands D and E that forms a flexible loop (DE loop; Fig. 36a).

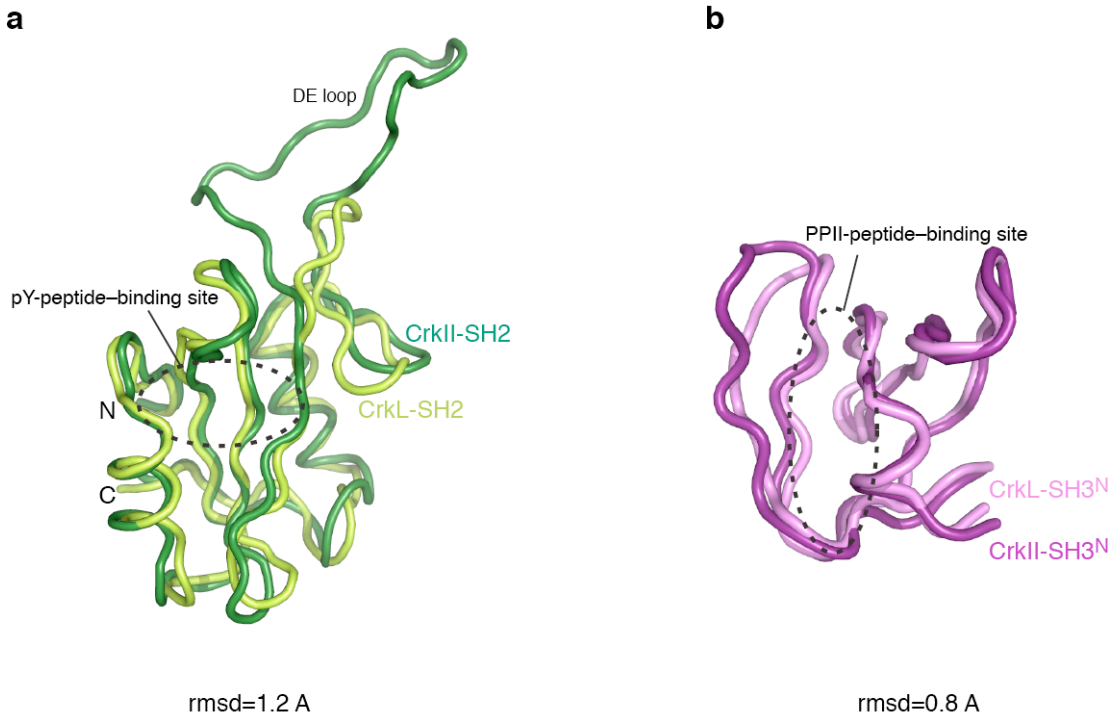

Figure 36. (a) Overlay of the structures of CrkL SH2 (this work) and CrkII SH2(ref1) domains. The DE loop, which was shown to bind to the SH3 domain of Abl, is present only in CrkII. (b) Overlay of the structures of CrkL $SH3^N$ (this work) and CrkII $SH3^N$ domains. Overall the structures are very similar other than some structural heterogeneity in the loops.

The DE loop is enriched in proline residues and has been shown to bind the SH3 domain of the Abl kinase[234]. This loop is absent in CrkL SH2 (Figs. 37a and 6a), and thus the binding mode of CrkL and CrkII to Abl is expected to be

different. In agreement with previous studies[136], the structural data show that the CrkL $SH3^C$ domain cannot bind PPII-type sequences because of the lack of aromatic residues at its canonical binding site.

The structural organization of CrkL is rather unique within the family of adaptor proteins: the SH2 domain and the first SH3 domain ($SH3^N$) interact directly with each other, whereas the second SH3 domain ($SH3^C$) does not interact with any of the modular domains (Fig. 37a). In agreement with the chemical shift analysis, the SH2-$SH3^N$ interaction is mediated by burying ~600 $Å^2$ of surface. The binding surfaces on both SH2 and $SH3^N$ domains consist primarily of polar and charged residues, and thus the interaction is mediated almost exclusively by polar contacts (Fig. 38b). Substitution of residues located at the SH2-$SH3^N$ interface disrupts the interaction between the two domains in CrkL (Fig. 32c).

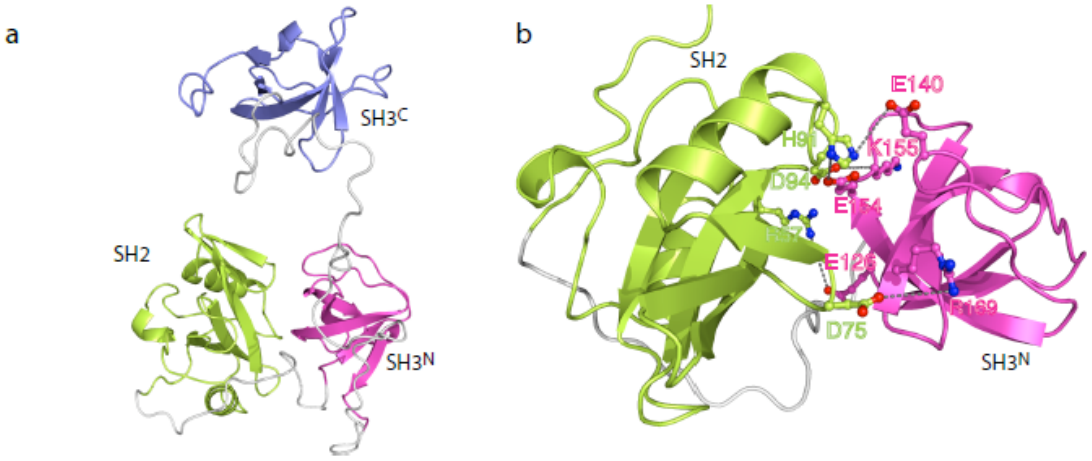

Figure 37. Structural and dynamic properties of CrkL. (a) Structure of CrkL. The SH2, $SH3^N$ and $SH3^C$ domains are colored green, magenta and blue, respectively. The linker regions are colored gray. The $SH3^C$ domain does not interact with the other domains. (b) Close-up view of the SH2-$SH3^N$ interface in CrkL. Only polar or charged residues mediate the interaction between the two domains.

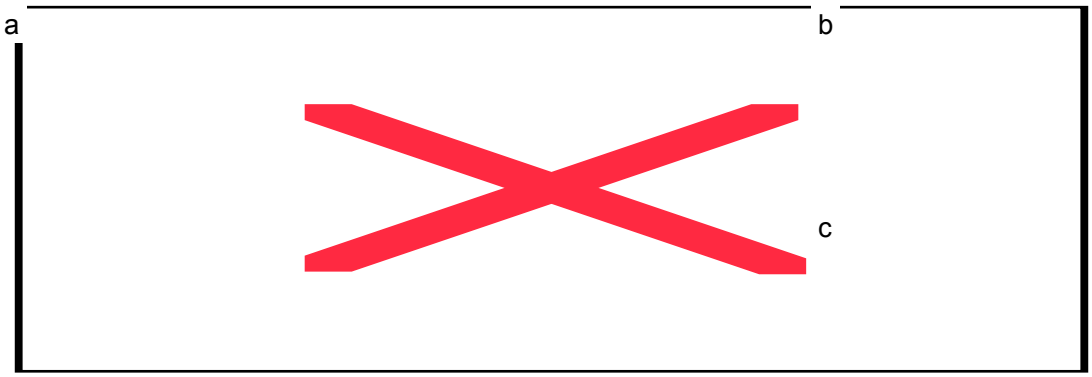

Figure 38. (a) Plot of the R_2/R_1 ratio. ^{15}N relaxation rates of the CrkL backbone as a function of residue number. The R_2/R_1 ratio provides information about the tumbling of the molecule, with higher values indicating slower tumbling. (b) Correlation times (τ_c) for the tumbling of CrkL. The SH2-$SH3^N$ module tumbles as a rigid unit, whereas the $SH3^C$ domain tumbles much faster and independently of the other domains. (c) Residues undergoing substantial μs–ms motions, as denoted by enhanced contribution to R2 (R_{ex}) values, are mapped on the structure of CrkL in red. Almost all residues located at the interface between the SH2 and $SH3^N$ domains show relatively high R_{ex} values, indicating that the binding interface is dynamic.

3.2.3 Dynamic properties of CrkL

To determine the motional properties of CrkL, we used NMR relaxation methodologies[235]. We measured the ^1H-^{15}N NOE, the longitudinal relaxation rate R_1 and the transverse relaxation rate R_2 (Fig. 38c and Fig. 39a). R_1 and R_2 are sensitive to and thus report on changes of the diffusion properties of the protein[202]. If only residues located at rigid parts of the molecule are considered, the ratio R_2/R_1 provides a good estimate of the correlation time (τ_c)[236]. The

relaxation analysis demonstrates that the SH2 and $SH3^N$ domains tumble as a unit, with a τ_c of ~11.0 ns, whereas the $SH3^C$ domain tumbles much faster and in an independent fashion with a τ_c of ~7.8 ns (Fig. 38d). Therefore, the relaxation data are in agreement with the structural data (Fig. 37a), showing that the SH2 and $SH3^N$ domains interact with each other, whereas the $SH3^C$ domain does not interact with any of the other domains. The present results indicate that $SH3^C$ has no role in regulating the activity of the SH2 and $SH3^N$ domains in CrkL, in sharp contrast to CrkII proteins, wherein $SH3^C$ was shown to act as a regulatory structural element by stabilizing the autoinhibitory conformation[139,138]. Further analysis of the relaxation data demonstrates that most of the residues located at the SH2-$SH3^N$ interface show enhanced motions on the μs–ms timescale (Fig. 38e). This observation suggests that the two domains move relatively to each other, giving rise to a dynamic binding interface.

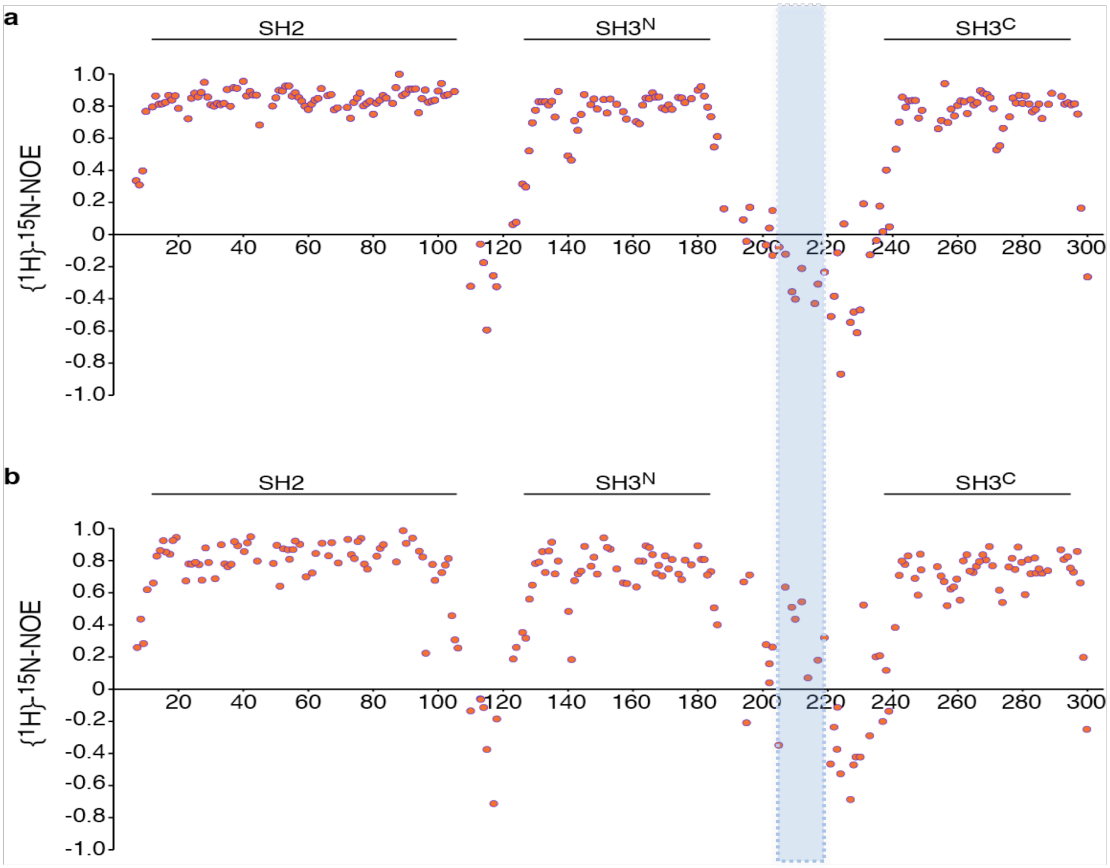

Figure 39. {1H}-15N-NOE values for (a) CrkL (b) pCrkL

3.2.4 The binding site of the SH2 domain in CrkL is occluded

Structural analysis of the SH2-SH3^N interface reveals that although the pTyr-binding site of the SH2 domain is largely accessible, the SH3^N domain masks the binding sites for the residues immediately downstream of the pTyr (Fig. 41). Thus, the structural data raise the possibility that the SH2-SH3^N intramolecular arrangement in CrkL may inhibit the binding activity of SH2 for pTyr ligands. To test this hypothesis, we used isothermal titration calorimetry (ITC) to directly measure the binding energetics of a phosphorylated peptide containing a

consensus sequence for CrkL SH2 (CrkL-pTyr207-peptide) to isolated SH2 and full-length CrkL (Fig. 40).

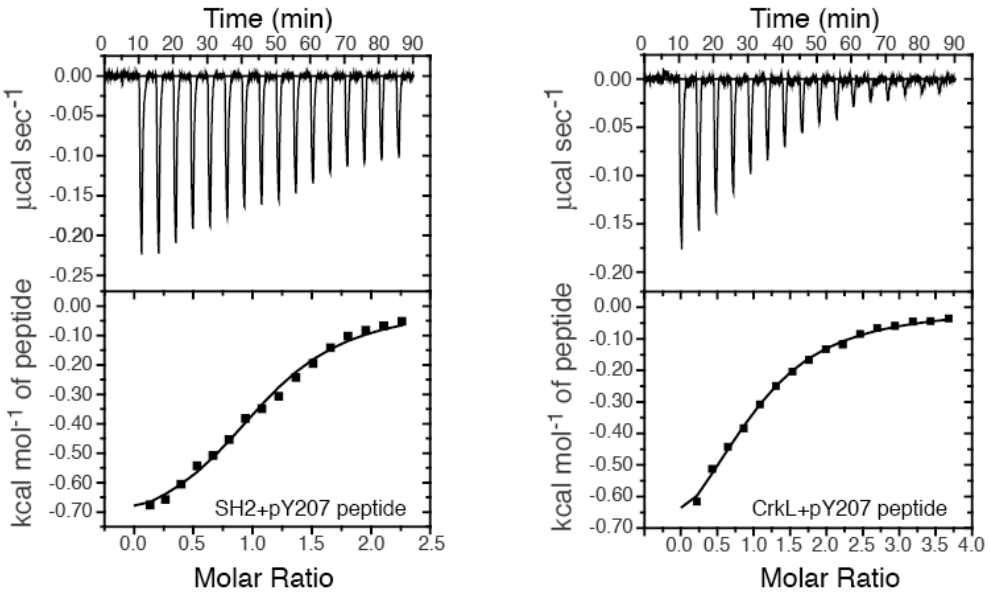

Figure 40. ITC traces and binding isotherms of the calorimetric titration of the CrkL-pTyr207-peptide to isolated SH2 domain and full-length CrkL

The CrkL-pTyr207-peptide binds isolated SH2 with moderate affinity (Kd ~7 µM), whereas the affinity for SH2 in the context of full-length CrkL is weaker by a factor of more than 3 (Kd ~23 µM) (Fig. 42). Thus, the intramolecular arrangement in CrkL gives rise to an autoinhibitory mechanism that partially occludes the SH2 domain, thereby modulating its activity for pTyr ligands.

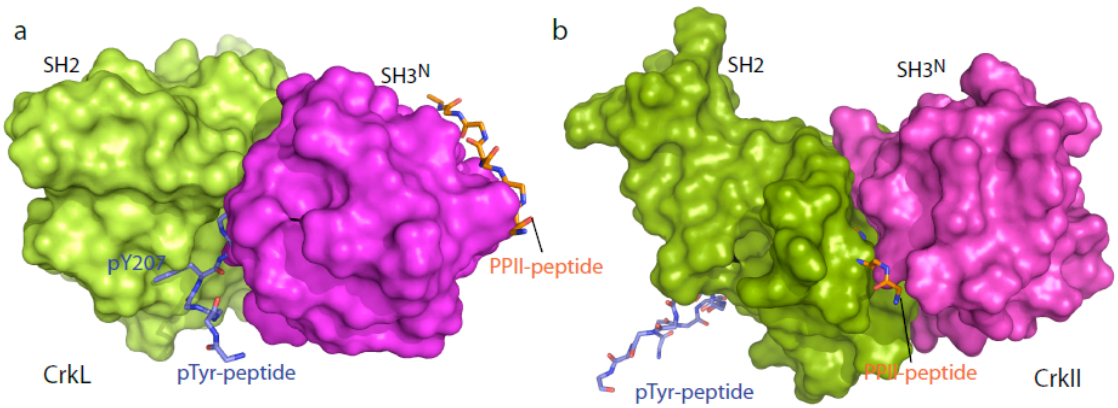

Figure 41. Binding of pTyr- and PPII-peptide ligands to CrkL and CrkII. (a,b) Structure of the SH2-SH3^N module in CrkL (a) and CrkII (b). The pTyr-peptide and PPII-peptide are shown as they have been previously determined to bind the isolated SH2 (PDB code 1JU5) and SH3^N domains (PDB code 1CKA), respectively. The pTyr-peptide binding site in CrkL is partially masked but is completely accessible in CrkII. Conversely, the PPII-peptide binding site in CrkL is completely accessible but is entirely masked in CrkII.

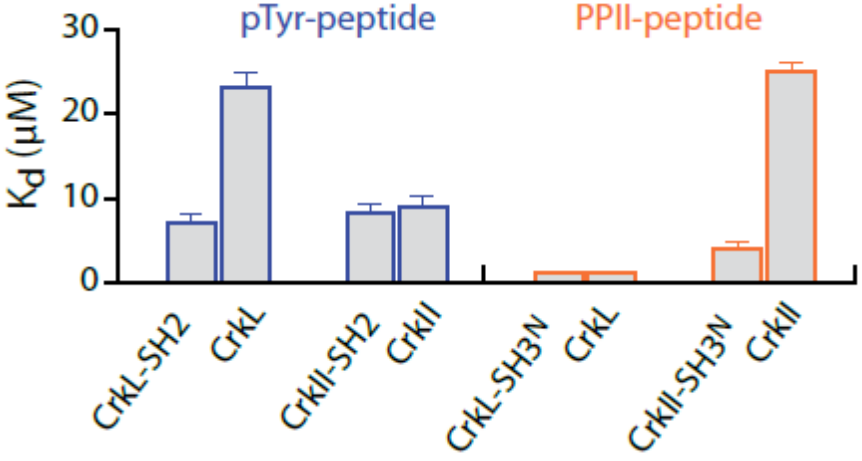

Figure 42. Dissociation constants (K_d) of pTyr-peptide and PPII-peptide complexes with CrkL (Fig. 44) and CrkII. Standard error was determined from three independent experiments. The K_d values of PPII-peptide binding to CrkII were obtained from ref.139.

In contrast to the case with CrkL, CrkII's SH2 domain seems not to be inhibited (Fig. 42b)[139]. Indeed, our ITC experiments show that a phosphorylated peptide containing a consensus sequence for CrkII SH2 (CrkII-pTyr221-peptide) binds with the same affinity to both the isolated SH2 and full-length CrkII (Fig. 42). It was previously proposed that the SH2 domain of CrkII and CrkL have different sequence specificity, with CrkL reported to bind a phosphorylated peptide derived from the fibroblast growth factor receptor (FGFR-pTyr-peptide) with a ~30-fold higher affinity than CrkII . However, ITC data show that the intrinsic affinity of the isolated SH2 domains of CrkII and CrkL for pTyr ligands (Fig. 42), even for the FGFR-pTyr-peptide (Fig. 43a), is very similar and that the two domains have identical specificity properties, as expected on the basis of the structures of the SH2–pTyr-peptide complexes (Fig. 43b). Taken together, the present data show that any pTyr ligand will preferably bind full-length CrkII over full-length CrkL by a factor of ~3–4 (Fig. 42).

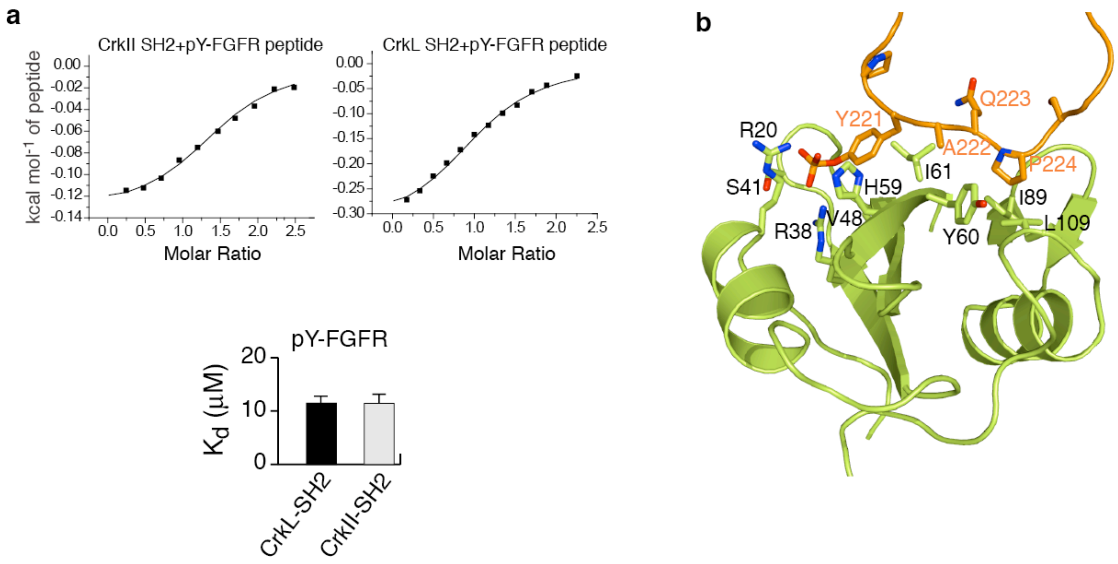

Figure 43. CrkL and CrkII SH2 domains have identical binding preferences. (a) Binding isotherms showing that the FGFR-pTyr-peptide, previously reported to bind with a 30-fold higher affinity to CrkL SH2 than to CrkII SH2[237], in fact binds to both proteins with very similar affinity. (b) Structure of the CrkII SH2 domain in complex with the pY221 peptide2. All of the residues lining the binding pocket are absolutely conserved in CrkL SH2 (Fig. 37a). Only a handful of residues are different in CrkL and CrkII SH2 and these are located remotely to the binding pocket.

3.2.5 The binding site of the $SH3^N$ domain in CrkL is accessible

Structural analyses of CrkII proteins have shown that the canonical PPII-binding site of the $SH3^N$ domain is almost completely occluded (Fig. 41b) [139,138]. As a result, the binding of full-length CrkII by PPII ligands, such as the Abl kinase, is weaker than that of the isolated SH3N by a factor of ~10 (Fig. 41c) [137, 138,237]. The present structural data show that the PPII-binding site of $SH3^N$ in CrkL, in sharp contrast to that of CrkII, is completely accessible (Fig. 41a). Indeed, ITC data (Fig. 44) show that a PPII-peptide containing the consensus sequence for $SH3^N$ binds isolated $SH3^N$ and full-length CrkL with the same affinity (Fig. 41c).

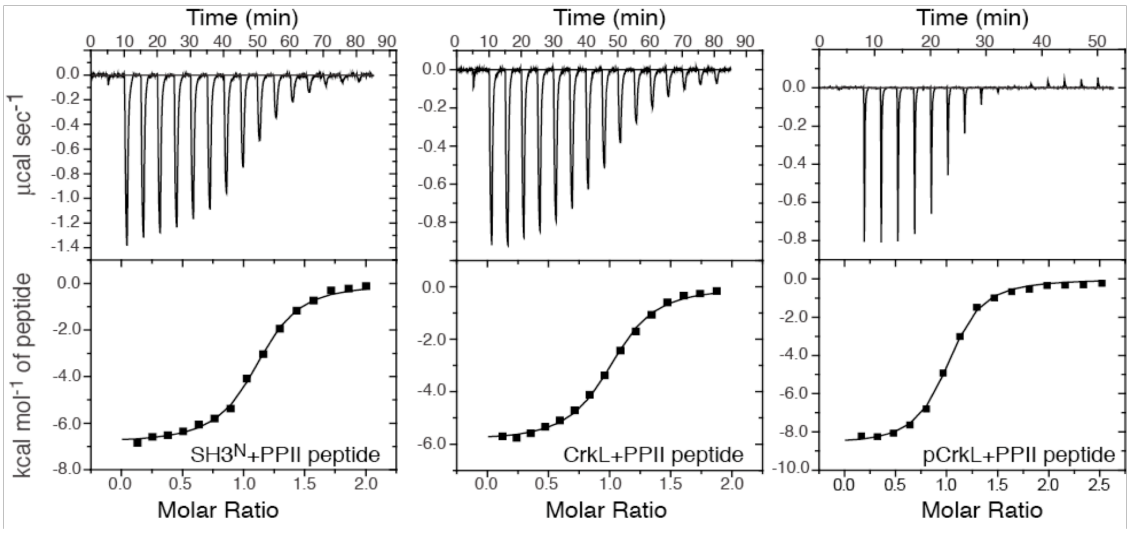

Figure 44. ITC traces and binding isotherms of the calorimetric titration of the PPII peptide to isolated $SH3^N$ domain and full-length CrkL and pCrkL.

NMR analysis of the titration of the PPII-peptide to CrkL shows that no steric clashes occur between the $SH3^N$-bound peptide and the rest of the protein, indicating that PPII-peptide binding to CrkL is unrestricted. In contrast, PPII-peptide titration to CrkII results in drastic conformational rearrangement and relief of the autoinhibitory conformation[137]. Taken together, our results show that a PPII-peptide containing the consensus sequence for $SH3^N$ binds isolated $SH3^N$ and full-length CrkL with the same affinity (Fig. 41c).

3.2.6 CrkL Tyr207 phosphorylation results in SH2 inhibition

Tyr207 in CrkL is phosphorylated by the Abl kinase[218,238]. The pTyr207-Xaa-Xaa-Pro210 region in the CrkL linker is a target sequence for CrkL SH2[219]. NMR

analysis of pCrkL (Fig. 33) shows that the CrkL SH2 domain interacts with pTyr207. MALLS and gel filtration data show that pCrkL remains a monomer in solution (Fig. 31a,b). The data collectively show that, upon phosphorylation of CrkL at Tyr207, the SH2 domain interacts in an intramolecular fashion with pTyr207, similarly to CrkII[139, 141].

We used NMR to determine the solution structure of pCrkL. In agreement with the NMR chemical shift analysis, the structural data show that the phosphorylated linker is bound to the canonical pTyr-binding cleft of the SH2 domain (Fig. 45). The interactions between the linker and the SH2 domain are virtually identical to those reported previously for the structure of CrkII SH2 and a phosphopeptide (Fig. 43b)[234]. Notably, the structural rearrangement elicited in CrkL by the phosphorylation of Tyr207 and the ensuing intramolecular binding to SH2 is minimal. The SH2-$SH3^N$ interface adjusts slightly to accommodate the binding of the pTyr linker region to SH2 (Fig. 45b and Fig. 35c), whereas the $SH3^C$ domain tumbles independently (Fig. 46) as in the unphosphorylated form.

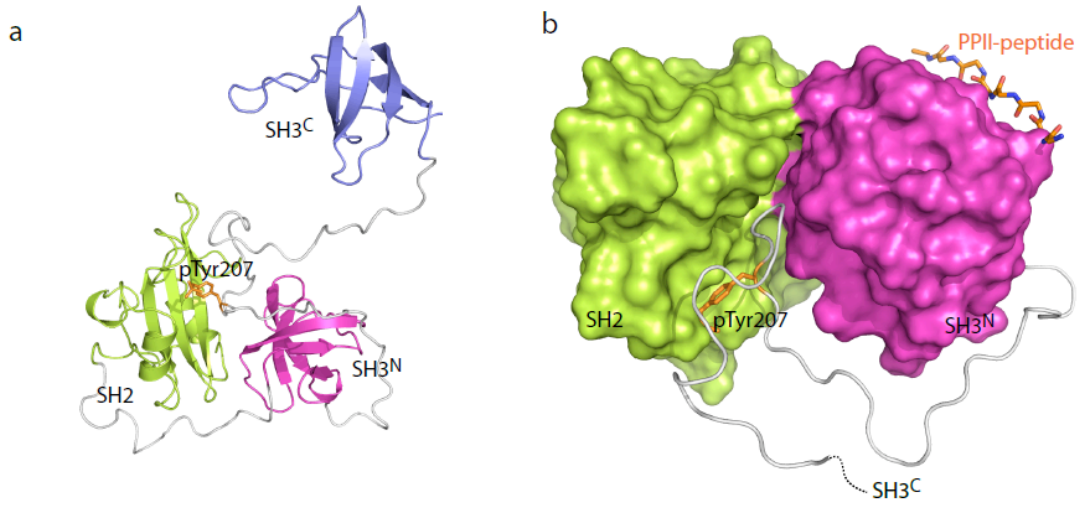

Figure 45. Structural and dynamic properties of pCrkL. (a) Structure of pCrkL. pTyr207 is shown as orange sticks. (b) Close-up view of the pTyr207-binding site. The SH2-SH3^{N} interface adjusts slightly to accommodate the binding of the linker to SH2.

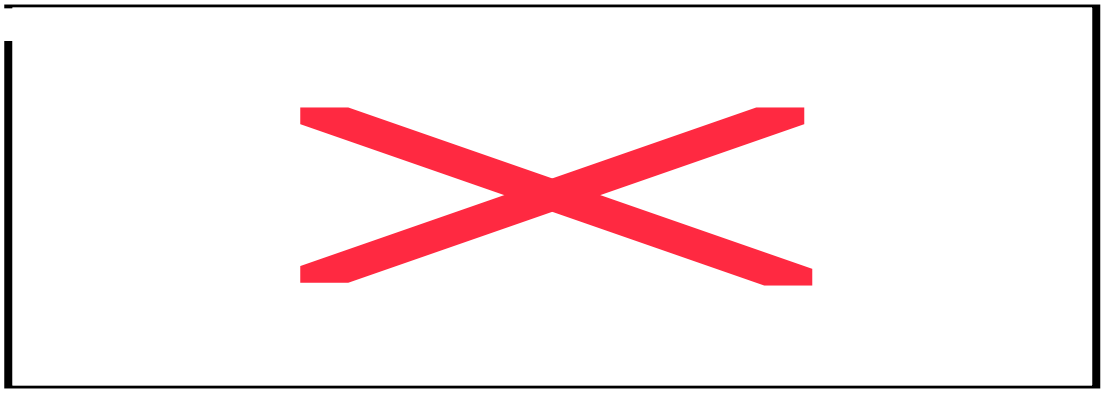

Figure 46. Plot of the R_2/R_1 ratio of pCrkL as a function of residue number. The scheme at right shows that the SH2-SH3^{N} module in pCrkL tumbles as a unit, as in CrkL, whereas the SH3^{C} domain tumbles much faster and independently of the other domains.

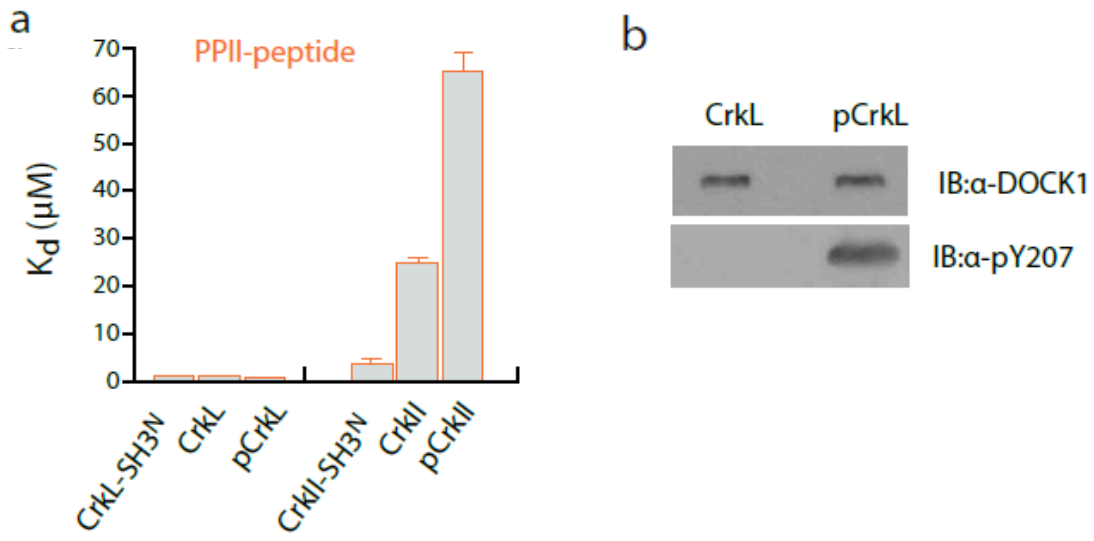

Figure 47. (a) K_d values of PPII-peptide complexes with CrkL (Fig. 40) and CrkII^{139} variants. Standard error was determined from three independent experiments. (b) Pulldown of CrkL and pCrkL with DOCK1, an SH3^{N}-binding physiological partner of CrkL (Fig. 55). IB (immunoblot).

The intramolecular binding of the phosphorylated linker (pTyr207) to the SH2 domain is expected to prevent SH2 from interacting with other phosphorylated ligands. To test this hypothesis, we monitored by NMR the effect of phosphorylating CrkL that is already bound to a phosphorylated peptide encompassing the linker sequence (pTyr-linker; Fig. 48a). Indeed, the NMR data show that, following the addition of catalytic amounts of Abl kinase domain (Abl^{KD}), the phosphorylated linker binds intramolecularly to SH2, and, as a result, the pTyr-linker is displaced (Fig. 48b). Pulldown of CrkL and pCrkL with pTyr-paxillin, a focal adhesion protein that interacts specifically with CrkL $SH2^3$, shows that paxillin forms a complex only with the unphosphorylated CrkL (Fig. 48c). Collectively, the data show that Tyr207 phosphorylation of CrkL by Abl results in intramolecular binding of the linker with the SH2 domain, thereby giving rise to CrkL SH2 inhibition.

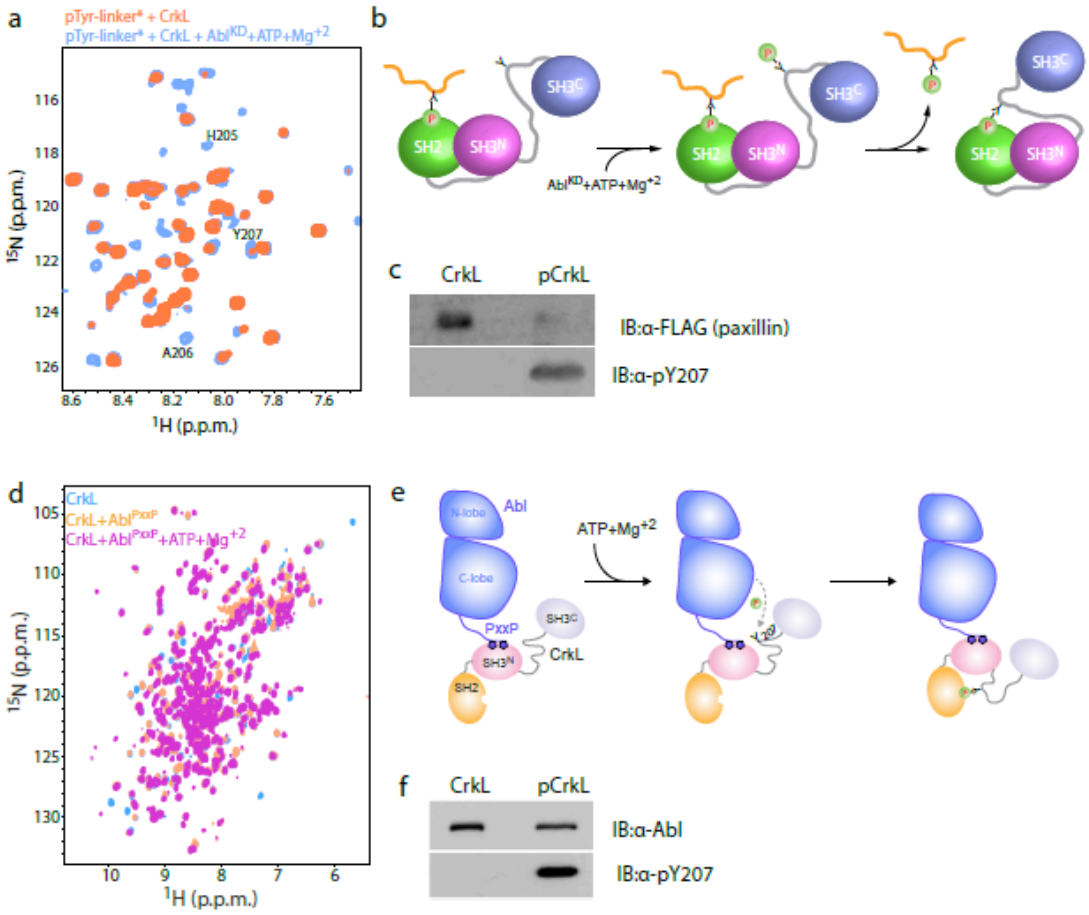

Figure 48. Effect of Tyr207 phosphorylation on CrkL folding and its association with Abl kinase. (a) ^{1}H-^{15}N HSQC NMR spectra of the linker region of CrkL containing the phosphorylated Tyr207 (pTyr-linker) in the presence of CrkL (orange) and after the addition of catalytic amounts of Abl^{KD} and ATP-Mg^{2+} (blue). The pTyr-linker is ^{15}N-labeled, whereas CrkL and Abl^{KD} 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 CrkL induces the intramolecular association of pTyr207 and SH2. As a result, the pTyr-linker is displaced. (c) Pulldown of CrkL and pCrkL with paxillin, an SH2-binding physiological partner of CrkL (Fig. 55). (d) ^{1}H-^{15}N HSQC NMR spectra of free CrkL (blue), in complex with Abl^{PxxP} (orange) and after adding ATP+Mg^{2+} (magenta). Abl^{PxxP} is a construct that encompasses the kinase domain and the first PxxP motif that binds CrkL. (e) Analysis of the NMR

experiments in d shows that CrkL forms a 1:1 complex with Abl^{PxxP} using its $SH3^N$ domain. Phosphorylation of Tyr207 elicits the intramolecular association of pTyr207 and SH2, but the intramolecular folding in CrkL has no effect on the CrkL–Abl^{PxxP} complex, which remains tightly associated. (f) Pulldown of CrkL and pCrkL with full-length Abl (form 1b) (Fig. 55). IB, immunoblot.

3.2.7 pCrkL interacts with signaling partners via $SH3^N$

Upon tyrosine phosphorylation, the SH2 domains in both CrkL (Fig. 48a–c) and CrkII [120,141] are inhibited for pTyr-ligand binding. CrkII has been shown to undergo a major conformational change upon phosphorylation that results in PPII-ligand binding inhibition to $SH3^N$ [139]. However, the structural rearrangement elicited by phosphorylation and the ensuing intramolecular binding in CrkL is minimal (Fig. 45a and b, Fig. 35c, Fig. 49). As a result, the PPII-binding site of the $SH3^N$ domain in phosphorylated CrkL is completely accessible (Fig. 45b). In agreement with the structural data, ITC experiments (Fig. 44) show that a PPII-peptide binds unphosphorylated CrkL (or the isolated $SH3^N$) and pCrkL with very similar affinities (Fig. 47a). Pulldown of CrkL and pCrkL with DOCK1, a guanine exchange factor (GEF) that activates Rac1 and specifically binds the $SH3^N$ domain, shows that DOCK1 associates strongly with both CrkL and pCrkL (Fig. 47b).

Abl binds the $SH3^N$ domain of CrkL and CrkII using a consensus Pro-Xaa-Xaa-Pro (PxxP) motif located C-terminally to its kinase domain[239]. An important implication of $SH3^N$ inhibition in CrkII is that Abl binding to and phosphorylation of CrkII results in pCrkII–Abl complex dissociation[240]. To test the emerging

hypothesis that pCrkL will remain tightly bound to Abl, we titrated Abl^{PxxP}, an Abl construct encompassing the kinase domain and the first PxxP Crk-binding motif, to labeled CrkL (Fig. 48d). NMR analysis indicates that Abl^{PxxP} binds the SH3^{N} domain of CrkL (Fig. 48e). Addition of ATP-Mg^{2+} results in Abl-mediated phosphorylation of Tyr207. NMR analysis (Fig. 4d) shows that pCrkL adopts the intramolecularly folded conformation but forms a tight complex with Abl^{PxxP} (Fig. 48e and Fig44). In agreement with these results, pulldown of CrkL and pCrkL with full-length Abl kinase shows that Abl forms complexes with CrkL that are not dependent on the CrkL phosphorylation state (Fig. 48f and Fig. 49).

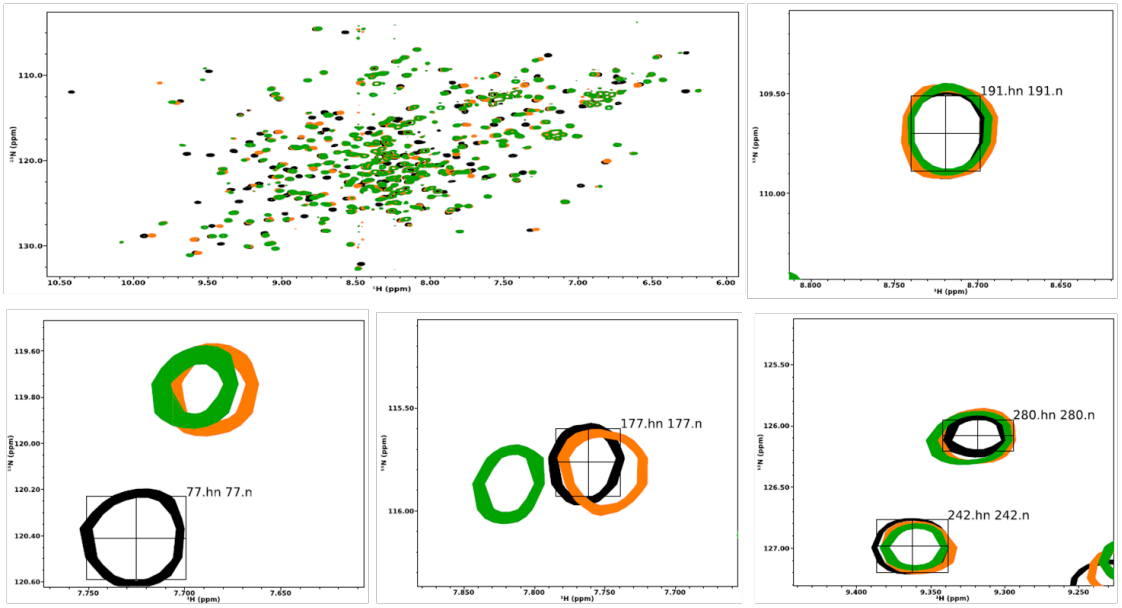

Figure 49. CrkL phosphorylation. ^{1}H-^{15}N HSQC NMR spectra of CrkL (black), pCrkL (orange) and pCrkL bound to C3G preptide (green). Phosphorylation of CrkL was carried out as described in methods using KD of Abl and ATP. Squares represent expanded domains regions: residue 77(SH2), 177(SH3^{N}), 242/280(SH3^{C}) and 191(SH3^{N}-SH3^{C} linker) in three states: apo (black), phosphorylated (orange) and phosphorylated with C3G peptide bound (green). This simple experiment shows that SH3^{N} domain is exposed even in the phosphorylated state.

3.3 Discussion

CrkL is an adaptor protein that regulates important cellular processes ranging from cell adhesion and motility to phagocytosis and apoptosis. *CRKL* has been recently identified as an essential gene in cancer cell proliferation[227] and has been shown to be indispensable for mediating the leukemogenic activity of Bcr-Abl[220,227]. CrkL is constitutively phosphorylated in human CML cells[228,229], and the level of CrkL phosphorylation is being used as a predictor of clinical outcome in patients treated for CML[230]. Despite the important role of CrkL, the lack of information about its structure has impeded a proper understanding of its function.

Because CrkL shares high sequence identity with CrkII, it has been thought that the structures of the two proteins are very similar. The structural data reported here demonstrate that the structural organization of the two proteins is considerably different. The CrkII structure[139] is stabilized by a hydrophobic segment in the linker region, part of which is quite different in CrkL (Fig. 50). Moreover, several contacts exist between the SH2 DE loop and $SH3^{C}$ in CrkII. In contrast, the DE loop is not present in CrkL SH2. The autoinhibited structure of pCrkII is stabilized by a linker region that is totally different in CrkL (Fig. 50b).

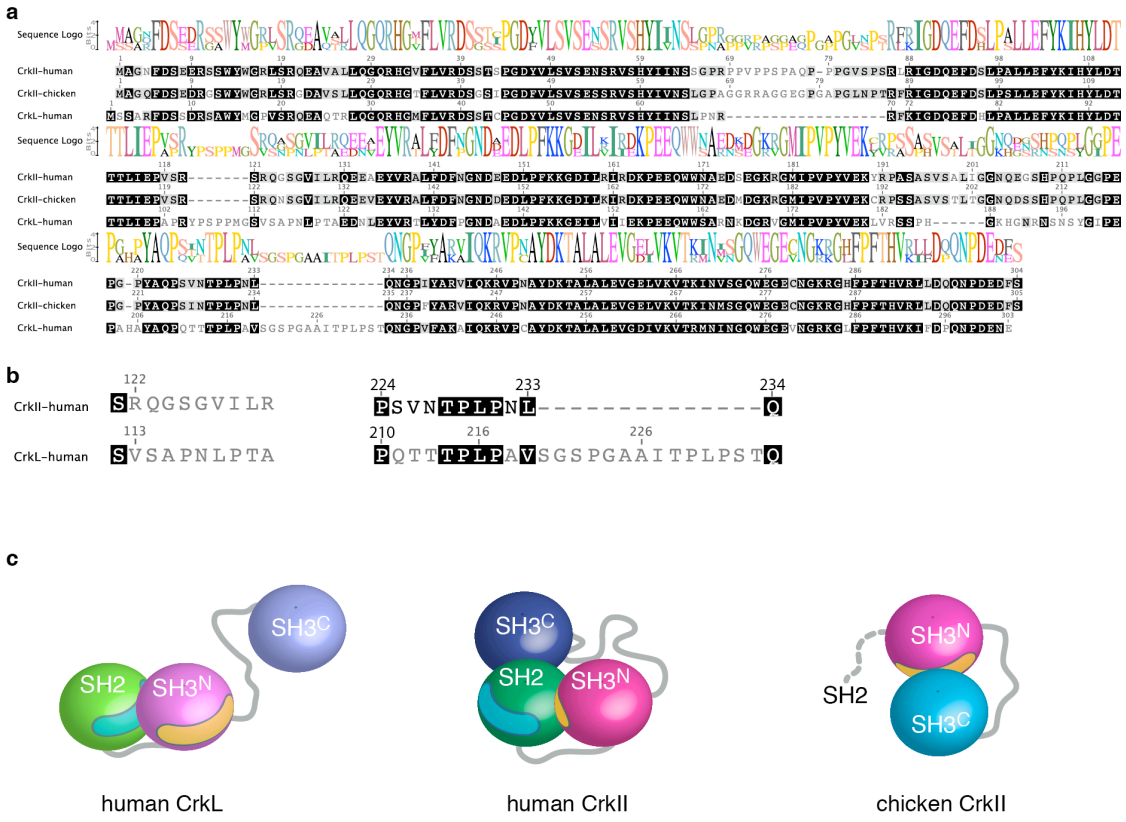

Figure 50. (a) Sequence alignment of human and chicken CrkII and human CrkL. (b) Regions where extensive sequence differences between human CrkII and human CrkL exist and may be responsible for the distinct structure adopted by the two proteins. (c) Structure organization of human CrkL (this work) and human[139] and chicken CrkII[138]. The blue and orange shaded regions in SH2 and $SH3^N$ denote the pY- and PPII-binding sites, respectively. The different structural organization between human CrkL and human CrkII may be explained by specific differences in their sequences. Specifically, residues 224−234 stabilize the structure in CrkII, but the sequence of this region is very different in CrkL. Moreover, the DE-loop in CrkII was seen to interact extensively with the $SH3^C$ domain in CrkII; the DE-loop is not present in CrkL. The structure of pCrkII is stabilized by a region spanning residues 121−131. The sequence of this region is very different in CrkL.

Thus, despite their high sequence identity (up to 72% in the structured regions), a few key sequence differences between CrkL and CrkII seem to modulate the

overall structural organization of the two proteins (Fig. 50). Overall, in the unliganded state, binding of PPII ligands to the $SH3^N$ domain in CrkII proteins is inhibited, whereas it is unrestricted in CrkL. In contrast, pTyr ligand-binding to CrkL SH2 is inhibited, whereas it is unrestricted in CrkII. The distinct structural architectures of CrkL and CrkII determine their signaling input and output, giving rise to distinct physiological functions for the two proteins. These results further highlight the notion that adaptors regulate signaling in a dynamic way and do not simply serve to wire signaling components in a passive manner[13].

In the resting state, the pTyr-binding site of the SH2 in CrkL is inhibited, whereas the one in CrkII is not, and, as a result, the binding of phosphorylated ligands to CrkII will be favored over binding to CrkL (Fig. 41). The differential SH2 binding activity modulation in CrkL versus CrkII can have important implications, for example, in the binding of the p130 Crk-associated substrate (p130CAS), a scaffold protein that mediates integrin-signaling (Fig. 51)[3]. Although p130CAS has multiple phosphorylation sites, it is conceivable that the number of such sites may be limited because of either the action of phosphatases or cell conditions[22]. In this case, competition favors CrkII over CrkL binding to p130CAS (Fig. 51).

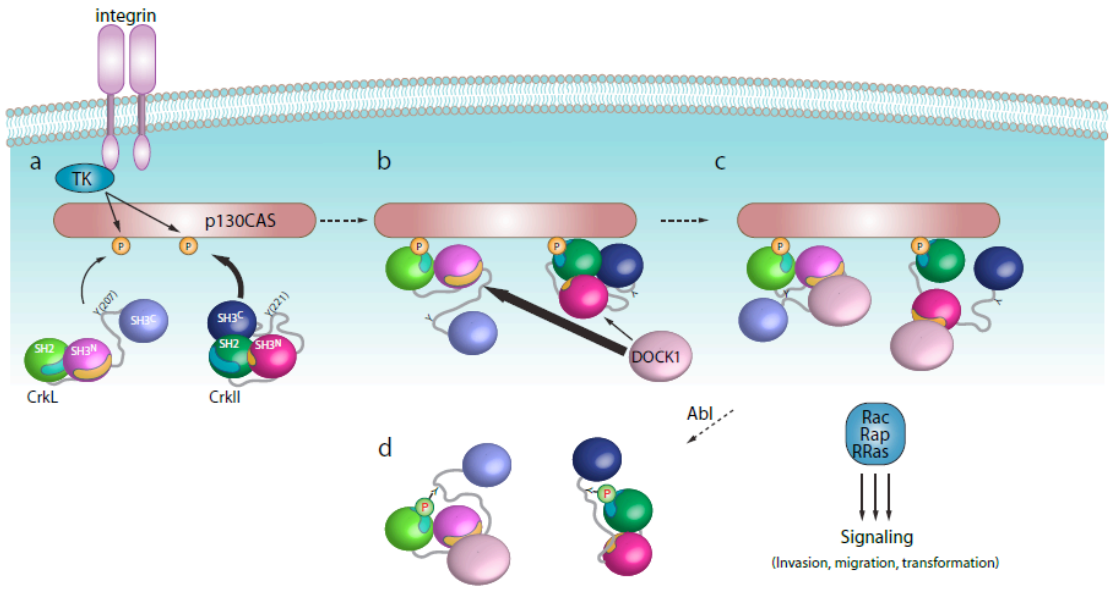

Figure 51. CrkL versus CrkII in integrin signaling. (i) Integrin activation elicits p130CAS phosphorylation by tyrosine kinases (TK), and, as a result, CrkL and CrkII are recruited. (ii) GEFs (for example, DOCK1 and C3G) associate with CrkL and CrkII via their $SH3^N$ domain, giving rise to efficient localized activation (iii) of small GTPases (for example, Rac, Rap and RRas) at the membrane. (iv) Abl-induced phosphorylation of CrkL and CrkII forces their dissociation from p130CAS and thus results in signaling suppression. The distinct structural organization of CrkL and CrkII modulates the interactions with their physiological partners to a different extent. The blue and brown shaded regions in SH2 and $SH3^N$ denote the pTyr- and PPII-binding sites, respectively.

The eventual association of CrkL or CrkII with p130CAS has minimal effect on the overall structure of either CrkL (Figs. 37a and 45a) or CrkII[139]. Therefore, the PPII-binding site of $SH3^N$ in CrkII is inhibited in the p130CAS-bound CrkII[240], whereas the corresponding site in CrkL is accessible. As a result, GEFs, such as DOCK1 and C3G, bind CrkL with a much higher affinity than CrkII (Figs. 42c and

51 (ii)). In this case, the CrkL-mediated complex is expected to more efficiently activate downstream GTPases, giving rise to a stronger signaling outcome than the CrkII-mediated complex (iii in Fig. 51). Enhanced association between CrkL and DOCK1 increases cell migration and invasion[3] and may explain previous observations that CrkL has a much higher oncogenic potential than CrkII in fibroblasts[140].

In both CrkL and CrkII complexes, phosphorylation of Tyr207 (Fig. 45) and Tyr221[120,139,141] respectively, by Abl causes intramolecular folding and displacement of the pCrk proteins from p130CAS (Fig. 48a–c), resulting in negative regulation of cell migration (iv in Fig. 51) [139]. Notably, pCrkL may form a constitutive complex with DOCK1, in contrast to pCrkII, as in pCrkL the $SH3^N$ domain is not inhibited (Fig. 47b). Thus, when phosphatases act to dephosphorylate CrkL and CrkII, the p130CAS−CrkL−DOCK1 complex assembles more readily than the corresponding CrkII complex.

A particularly remarkable result in the present work is that the phosphorylated form of CrkL remains active and may interact via its uninhibited $SH3^N$ domain with various ligands, including the Abl kinase (Fig. 48d–f). The interaction of CrkL with Bcr-Abl substantially stimulates the leukemogenic activity of the oncoprotein[140,241]. The present data demonstrate that CrkL binds and remains tightly bound to Abl even after the kinase has phosphorylated CrkL (Fig. 48e,f), explaining why CrkL is a preferred substrate for Bcr-Abl. This behavior is in sharp contrast to that of pCrkII, which has been shown to be an entirely inactive protein [120]. Therefore, although the intrinsic affinity of the $SH3^N$ domains of CrkL and

CrkII for the PxxP motif of Abl is almost identical (Fig. 42), the overall structural organization of the proteins confers an advantage to CrkL as the binding of CrkII to Abl is suppressed. Although the $SH3^C$ domain has been shown to function as an autoregulatory element in CrkII[137,139,140,237], in CrkL, $SH3^C$ does not interact with any of the other domains, and thus its role remains elusive. Functional data, however, indicate that CrkL $SH3^C$ is indispensable for fibroblast transformation and hematopoietic cell adhesion[238]. It is likely that $SH3^C$ mediates its function by interacting with an as yet unidentified partner or by being tyrosine phosphorylated, as shown recently for CrkII[140].

SH2 and SH3 domains have long been thought to mediate sequence-specific interactions involving PxxP and pTyr motifs, respectively.

However, there is now a growing list of examples indicating that these signaling domains can interact with sequences that do not conform to these general rules[242]. Remarkably, our data demonstrate that the two domains can engage each other in a completely unconventional manner. SH2-SH3 interactions have been previously observed in Itk kinase[243] and in the SAP–Fyn complex[244]. Comparison of the binding mode in these systems and in CrkL demonstrates the great versatility of the SH2 and SH3 domains in mediating interactions in cell signaling (Fig. 52).

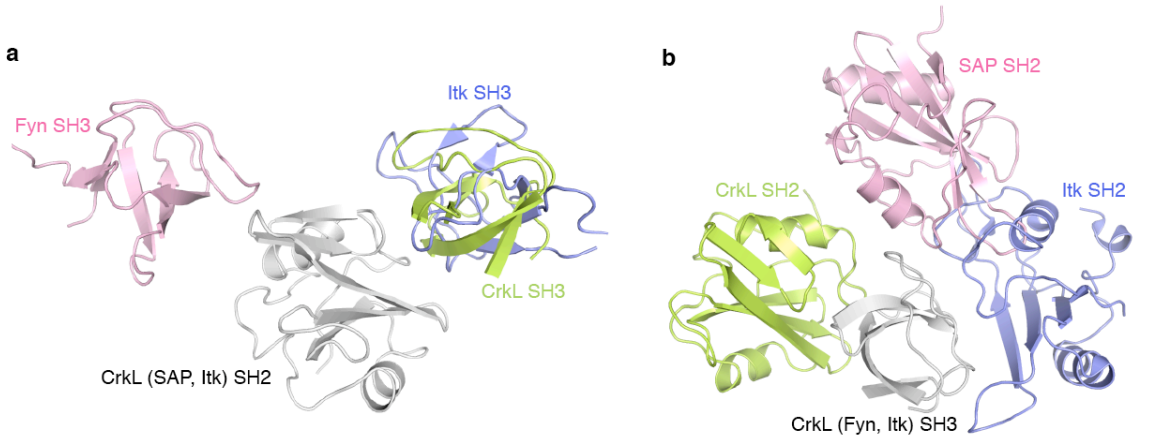

Figure 52. Structural comparison of non-conventional interactions between SH2 and SH3 domains. The SAP SH2−Fyn SH3[244], Itk SH2−SH3[243] and CrkL SH2−SH3 (this work) structures are shown. In (a) the SH2 domains of all three complexes are overlaid (for clarity only the CrkL SH2 domain is visible). The SH3 domains of Itk and CrkL interact with roughly the same region on SH2 whereas the binding site for Fyn SH3 is distinct. In (b) the SH3 domains of all three complexes are overlaid (for clarity only the CrkL SH3 domain is visible). Interestingly, the SH3 domain appears to make use of distinct regions to interact with each one of the SH2 domains.

3.4 Materials and Methods

3.4.1 Protein constructs

The following human CrkL constructs were prepared (Fig. 53): SH2 (residues 1–105), $SH3^N$ (residues 120–188) linker-$SH3^C$ (residues 184–303), SH2-$SH3^N$ (residues 1–188), pTyr-linker (residues 188–230) and full-length CrkL (residues 1–303). 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.

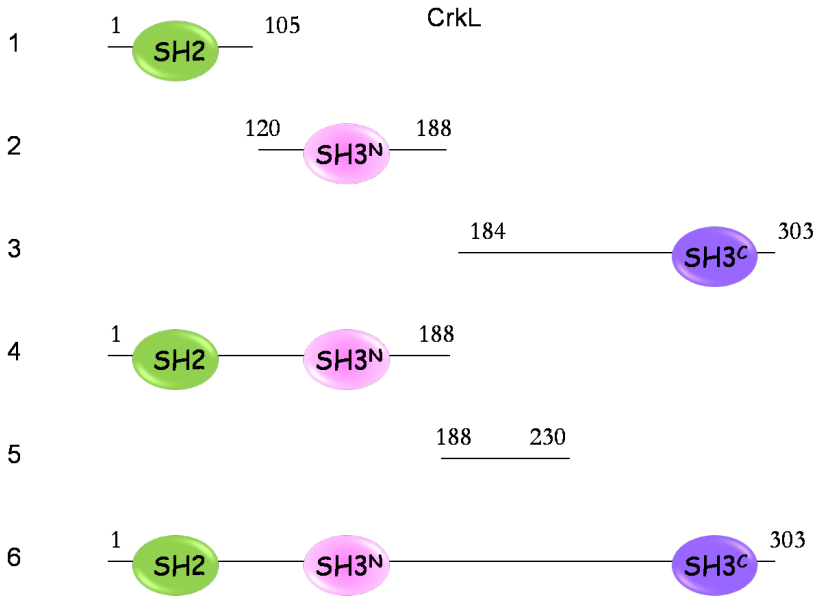

Figure 53.CrkL constructs used in this study.

The following human Abl constructs were prepared: Abl^{KD} (residues 229-534), Abl^{PXXP} (residues 229-556) The constructs were cloned into a modified pet16b vector using the Nde1 and Xho1 restriction sites. A Tev protease cleavage site was introduced between the histidine tag and the protein.

3.4.2 Protein labeling

Isotopically labeled samples for NMR studies were prepared by growing the cells in M9 minimal media supplemented with 1 g l-1 of 15NH4Cl and 2 g l-1 of 13C6-glucose. For isotope labeling, minimal media containing $^{15}NH_4Cl$ and [^{2}H,^{12}C] or [^{2}H,^{13}C]-glucose in 99.9% $^{2}H_2O$ were used.

For the production of U-[^{2}H],Ile-δ1-[$^{13}CH_3$] and Val,Leu-[$^{13}CH_3$,$^{12}CD_3$] samples, 50 mg l^{-1} of alpha-ketobutyric acid (methyl-$^{13}CH_3$) and 100 mg l^{-1} of alpha-

ketoisovaleric acid (dimethyl-$^{13}CH_3$,$^{12}CD_3$) were added to the culture 1 hr prior to addition of IPTG.

3.4.3 Protein preparation and purification

All Abl constructs were transformed in E.coli *BL21*(DE3) that had been already transformed with a vector containing the YopH phosphatase and pREP4-groESL followed by plating on LB plate. A single colony was picked from the plate and transferred into 5 ml of autoclaved LB media. After ~ 8hour of growing, the cells were transferred to 50ml LB. 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 to 1L culture of LB or M9 minimal media (see 3.4.2). Cultures for all CrkL constructs were grown at 37 °C, and protein synthesis was induced by addition of 0.5 mM IPTG at OD600 ~0.4. Cells were harvested after 24 hrs. For Abl constructs at OD600 ~0.4 the temperature was then lowered to 16° C and protein synthesis was also induced with 0.5mM IPTG. Cells were harvested after 24 hrs. Cells were lysed and the cytosolic fraction was separated from the membrane fraction by centrifugation at 70,000 × g. 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 200 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. Phosphorylated CrkL and pYlinker were prepared by the addition of catalytic amounts of AblKD in buffer (50mM KPi (pH 6.8), 150mM NaCl, 1mM β-mercaptoethanol) supplemented with 5 mM ATP and 10 mM MgCl2. The reaction was carried out at room temperature. Phosphorylation efficiency is over 99% as judged by mass spectrometry and NMR.

3.4.4 Peptides

The sequences of the CrkL-pTyr207-, CrkII-pTyr221-, PPII- and FGFR-pTyr-peptides used are EPAHApYAQPTT, PEPGPpYAQPSV, YLQAPELPTKTRTS and AGVSEpYELPEDPR, respectively. The MS-determined molecular masses of the peptides were 1,393.1 Da, 1,221.2 Da, 1,605.3 Da and 1,541.4 Da, respectively. All peptides were >99% pure.

3.4.5 NMR spectroscopy.

NMR experiments were performed on Varian 800- and 600-MHz and Bruker 700- and 600-MHz spectrometers. Complete backbone and side chain assignment for the CrkL proteins studied was achieved using standard triple-resonance pulse sequences, including HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, HCCONH and CCONH. NOEs were measured using two- and three-dimensional 13C and 15N edited NOESY experiments using mixing times of 80 and 100 ms, respectively. All experiments were performed at 32 °C.

3.4.6 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. 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 ^{1}H-^{15}N 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[184.] Ensemble simulated annealing refinement was used as described[245]. 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.

3.4.7 RDC Measurements.

Alignment of the proteins for RDC measurements was achieved using poly(ethylene glycol)/alcohol mixtures[246]. 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. ^{15}N-HSQC (IPAP) (Fig.54) and HNCO based experiments were used to measure one-bond N-H and CαC' RDCs[247].

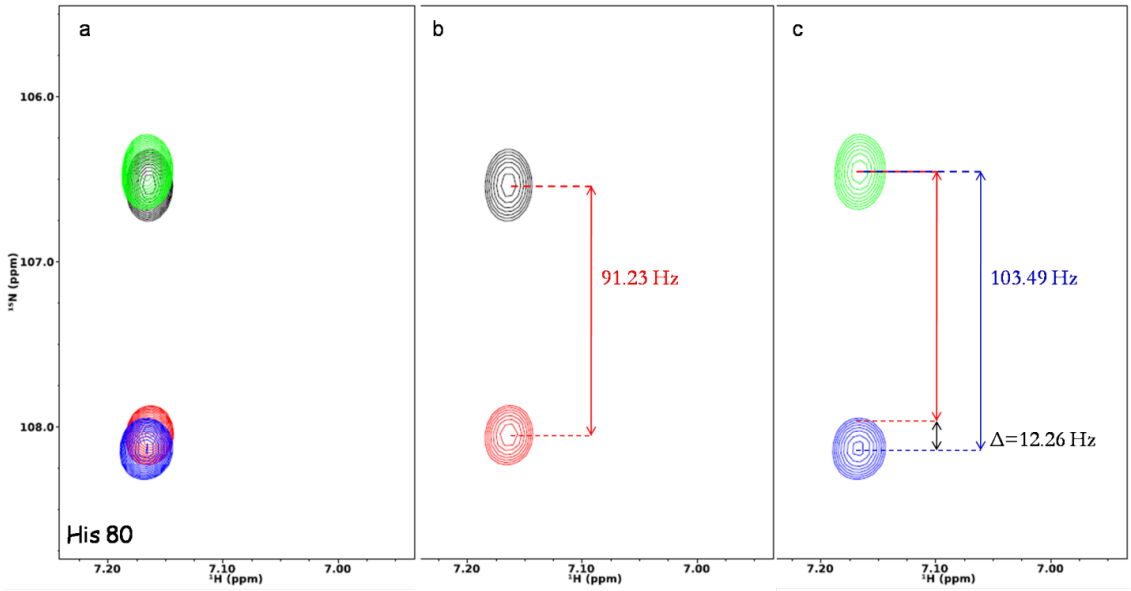

Figure 54. Measurement of ^{15}N-^{1}H residual dipolar couplings for a protein in a 5% C12E5/hexanol (molar ratio=0.96) mixture. (a) overlay of both ^{15}N-^{1}H HSQC residual dipolar couplings, isotropic (black/red) and partly oriented (green/blue). (b) HSQC without decoupling in ^{15}N dimension, isotropic solution ^{15}N-^{1}H splitting observed, equal to ^{15}N-^{1}H one-bond scalar coupling (~92-95 Hz). (c) HSQC without decoupling in ^{15}N dimension -partly oriented ^{15}N-^{1}H splitting observed, equal to ^{15}N-^{1}H one-bond scalar coupling plus RDC

3.4.8 Structure calculation and refinement.

Structure calculations were performed with Xplor-NIH. The ^{13}Cα, ^{13}Cβ, ^{13}C′, Hα, ^{15}N and NH chemical shifts served as input for the TALOS+ program[73] 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%.

3.4.9 Relaxation measurement and analysis.

Three relaxation parameters were measured for all backbone amides of CrkL: the {1H}-15N NOE, the longitudinal relaxation rate R1 and the transverse relaxation rate R2. 15N R1 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)=I0 e-Rt, where I is the intensity of the magnetization. {1H}-15N 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[248].

3.4.10 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, 0.45 mm) 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.

3.4.11 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.

3.4.12 Pull-down assays.

Phosphorylated GST or GST-CrkL was obtained by incubating purified GST or GST-CrkL (~1 μM) with purified AblKD (~0.1 μM) overnight at 25 °C in a buffer containing 50 mM Tris (pH 7.5), 150 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol and 5 mM ATP. Purified GST-CrkL was incubated in the same buffer without Abl^{KD} and ATP overnight for the unphosphorylated sample. Lactose-free glutathione agarose beads (Sigma) were then washed once with PBS containing 0.1% (v/v) Triton-X-100 and incubated with the aforementioned samples for 60

min at 4 °C. Beads were then spun down and washed twice with PBS containing 0.1% (v/v) Triton-X-100. Beads with bound phosphorylated GST or phosphorylated or unphosphorylated GST-CrkL were each incubated with lysates of 293T cells untransfected or transfected with Flag-paxillin or human Abl1b for 2 h at 4 °C. Beads were then spun down and washed twice with PBS containing 0.1% (v/v) Triton-X-100. Samples were then boiled in SDS sample buffer and analyzed by western blotting with antibody to Flag (anti-Flag; Sigma F1804, 1:1,000 dilution), anti-DOCK180 (Santa Cruz SC6043, 1:1,000 dilution), anti-Abl (Sigma SAB4501043, 1:2,000 dilution), anti-CrkL (Cell Signaling 3181, 1:1,000 dilution) and anti-pCrkL (Tyr207; Cell Signaling 3182, 1:1,000 dilution) (Figure 55).

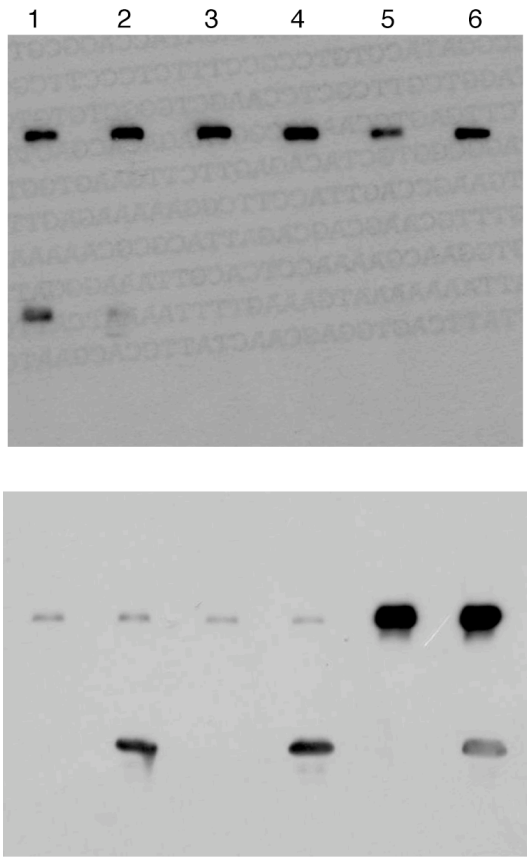

Figure 55. Pull-down of CrkL and pCrkL with DOCK1, paxillin, and full-length Abl (form 1b). Lane 1: Pull down with GST-CrkL from lysates of cells transfected with Flag-Paxillin; lane 2: Pull down with phosphorylated GST-CrkL from lysates of cells transfected with Flag-Paxillin; lane 3: Pull down with GST-CrkL from lysates of untransfected cells: lane 4: Pull down with phosphorylated GST-CrkL from lysates of untransfected cells; lane 5: Pull down with GSTCrkL from lysates of cells transfected with human Abl 1b; lane 6: Pull down with phosphorylated GST-CrkL from lysates of cells transfected with human Abl 1b. The membranes were cut in between 37 and 50 kD and in between 75 and 100 kD. This portion which includes proteins in between 50 and 75 kD was used to probe for Flag-Paxillin and GST-CrkL. The portion of the membrane that included proteins in between 100 and 250 kD was used to probe for DOCK180 and Abl. Both parts of the membranes (50 - 75kD and 100 to 150 kD) were lined up and exposed to film simultaneously which is why we can visualize DOCK180 and Paxillin on the same blot and similarly, Abl and phoshorylated CrkL.

Chapter 4. CrkII as a novel biological substrate for CypA

4.1 Introduction

Cyclophilin A (CypA) is a polyprolylisomerase that is ubiquitously expressed in all human cells. It is known to catalyze the cis-trans isomerization of X-P motifs. Recent studies have shown that CypA is overexpressed in many human cancers[249]. The role of CypA in this area is not well understood. We have identified a novel binding site for CypA on CrkII. Although several proteins have been identified to date to interact with CypA, the underlying mechanism of CypA action and the physiological implications of these interactions have remained unknown in most cases. Despite the fact that CypA exhibits peptidyl−prolylcis−trans isomerase (PPIase) activity, it is unclear whether CypA acts as an enzyme or a binding partner in mediating the biological processes in which it is involved. Here we show that CypA binds specifically to, and modulates phosphorylation of, CrkII by Abl kinase. We show that the conserved proline residue in CrkII next to the negative-regulatory tyrosine-phosphorylation site undergoes cis−trans isomerization. CypA is recruited to the site and delays phosphorylation 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.

4.2 PPIase family

In proteins, proline residues can exist in two unique cis and trans peptide bond conformations[250]. The mechanism of proline cis-trans isomerization is widely used to control a number of important biological processes including gene regulation[251], cell signaling[137],[252], enzyme function[253] and virus infection[254]. However, this is an intrinsically slow conversion, but can be catalyzed by the enzymes called peptidylprolyl*cis-trans*isomerases (PPIases). The PPIases contain four distinct enzyme families: cyclophilins (Cyps)[255], FK506-binding proteins (FKBPs)[256], parvulins[257] and Ser/Thr phosphatase 2A (PP2A) activator PTPA[258].Among all PPIases, Cypsand FKBPs have received the most of the attention due to the fact that they are cellular targets for immunosuppresive drugs Cyclosprin A (CsA) and FK506.

4.2.1 Cyclophilins

Cyclophilins (Cyps) are found in all cells of all studied organisms. In humans there are at least 16 cyclophilin proteins[259], including CypA, CypB, CypC, CypD, CypE, Cyp40, and CypNK. CypA was the first to be identified as a drug target (for the immunosuppressive drug CsA[260]) and makes up ~0.1-0.6% of the total cytosolic protein[261]. Cyps play an important function in stabilizing the cis-trans transition state by accelerating the isomerization[262]. This process has been shown to be important in protein folding and multi-domain protein assembly[263].

4.2.2 Cyps in human diseases

Cyclophilins are found in various pathological conditions including HIV[264], hepatitis B and C[265] , and cardiovascular diseases including atherosclerosis[266]. Cyps were shown to be overexpressed in certain forms of human cancers. In particular, CypA shows an unusually high expression level in several cancer types including lung cancer[267], liver cancer[268], pancreatic cancer[269], glioblastoma[270] and hematologic malignancies[271]. In these cases it promotes cancer cell proliferation, cell migration/invasion and drug resistance. The mechanism by which CypA contributes to cancer progression is not fully understood. However CypA appears to be involved in multiple pathways including binding to the membrane receptor CD147[272], a plasma membrane protein shown to interact with CypA on the membrane of the immune cells[273].

4.2.3 Structure and mechanism of action

CypA is the most studied member of the Cyp family of PPIase. CypA folds as a β-barrel with eight β-strands and three α-helices flanking the β-barell, connected by flexible loops (Fig. 56b). The active site of CypA consists of four β-sheets and four loops which together form a narrow cleft[274] where substrates binds (Fig. 56a). It was shown that during the isomerization process there is very little structural change on the CypA active site[275]. Residues of CypA involved in the interactions with the substrate include Arg55, Ile57, Phe60, Met61, Gln63, Ala101, Asn102, Gln111, Phe113,Trp121, Leu122, and His126. Among them,

Arg55 was shown to be crucial for catalytic activity, and Arg55 to Ala mutant retain less than 1% of the wild-type catalytic efficiency[276].

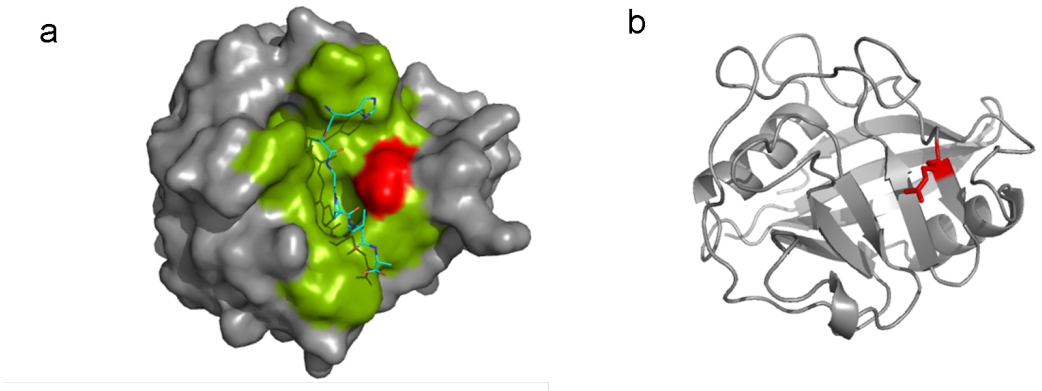

Figure 56. Structure of CypA-peptide complex (PDB code 1AWQ)[277]. (a) residues that are involved in the interactions with the substrate are highlighted in green (R55 in red), substrate (HAGPIA) in yellow. (b) cartoon representation of uncomplexed CypA structure (R55 in red).

The crystal structures analysis of CypA, complexed with HIV-1 protein in both cis and trans conformations, provided mechanistic insight to the CypA catalytic mechanism. The HIV-1 CA (capsid) protein is composed of two helices connected by a loop. The loop contains a G89-P90 peptide bond which is a CypA target. Relative to the P90, the C-terminal segment of HIV-1 CA (P90-P93) adopts the same conformation in all X-Ray structures regardless of cis or trans conformation. The N-terminal region to HIV-1CA P90 is markedly different, but only for residues 88 and 89 next to the substrate proline. The proposed model shows that the proline remains fixed relative to the enzyme as Arg55 of CypA anchors the P90 oxygen and activates the proline amide of the isomeric peptide bond. Howard et. al, explain that the steric clash of the sidechain in the residue

preceding P90 prevents optimal binding of the substrate in the trans conformation. By the mutation G89A in HIV-1 CA, they have shown that Ala would place the C^{β} methyl group unacceptably close to the Arg55.

4.3 Results

4.3.1 Human-CrkII, a new substrate for CypA

4.3.1.1 Conserved GP motif in Ckr protein family

Despite the high sequence homology between CrkII and CrkL (Fig. 53) and the identical binding preferences of their SH2 and SH3 domains, these two proteins contain slight sequential differences (Fig. 57) and have been shown to play non-overlapping roles in the cells[1].

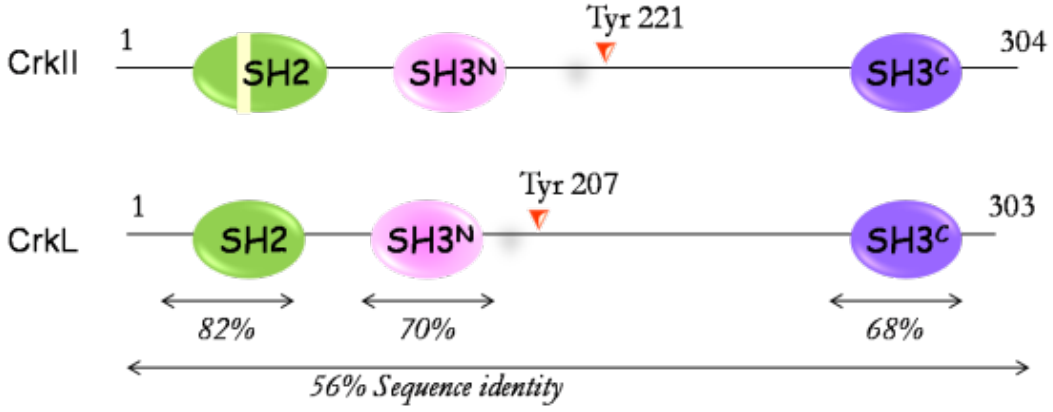

Figure 57. Comparison of sequence identity between human-CrkII and CrkL.

Both proteins contain a ~50 residue linker that connects the two SH3 domains which is crucial for regulation[137,139]. However the Crk protein possesses high degree of similarity across species. From the primary amino acid sequence

analysis of human and chicken CrkII, two identical motifs can be identified, namely G219-P220/G236-P237 in human and G220-P221/G237-P238 in chicken (Fig. 58). The G237-P238 motif in chicken CrkII regulates autoinhibition of the $SH3^N$ domain via a cis-trans isomerization mechanism[251].However, study of full-length human-CrkII did not reveal any cis-trans isomerization in the analogous region[137].

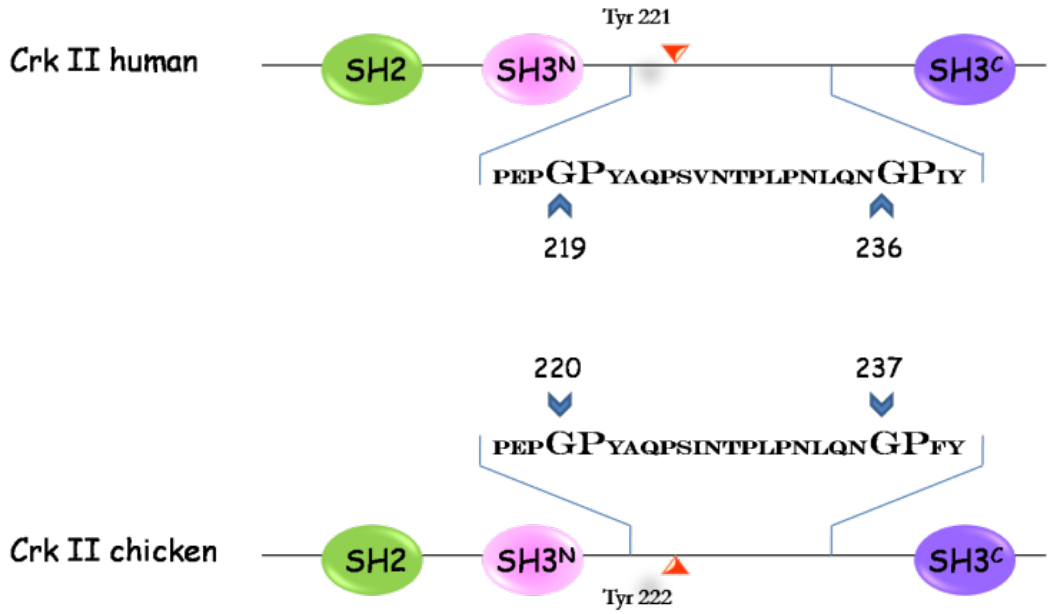

Figure 58. Figure 2b.CrkII domain organization of human and chicken protein. Regions comprising “GP”motifs are expanded.

A possible reason for this is that cis-trans isomerization could be very subtle or absent in the case of the full length protein used in this study (buffer and temperature specific)[137]. Also, in the human-CrkII protein, the isomerization of G219-P220 and G236-P237 could have little impact on the overall structure because its autoinhibitory mechanism is not controlled by cis-trans isomerization.

NMR titration of human-CrkII revealed that CypA interacts exclusively with the G219-P220 motif, revealing another possible mode of human-CrkII regulation. On the other hand, CrkL does not contain a GP motif and is therefore not susceptible to CypA regulation.

It has been shown previously that CypA preferentially binds Gly-Pro consensus sequences[278] and catalyzes the cis-trans isomerization of peptide bonds preceding the proline residue. Sequence alignment revealed that the residue preceding Tyr207 in CrkL is Ala, as opposed to Pro in human-CrkII (Fig. 59),.

Figure 59. Sequence alignment of CrkL and CrkII proteins from different spiecies. The CrkL is lacking conserved GP motif presented in all CrkII proteins.

4.3.1.2 Evidence of cis-trans isomerization at G220-P221 in CrkII

NMR experiments show that CrkII undergoes *cis-trans* isomerization at G220-P221. 2D ^{15}N HSQC specra analyses of linker-$SH3^{C}$ (residues 190-297) reveal,

that both G220-P221 and G237-P238 prolyl bonds exist in the cis and trans conformations (Fig. 60c,d).

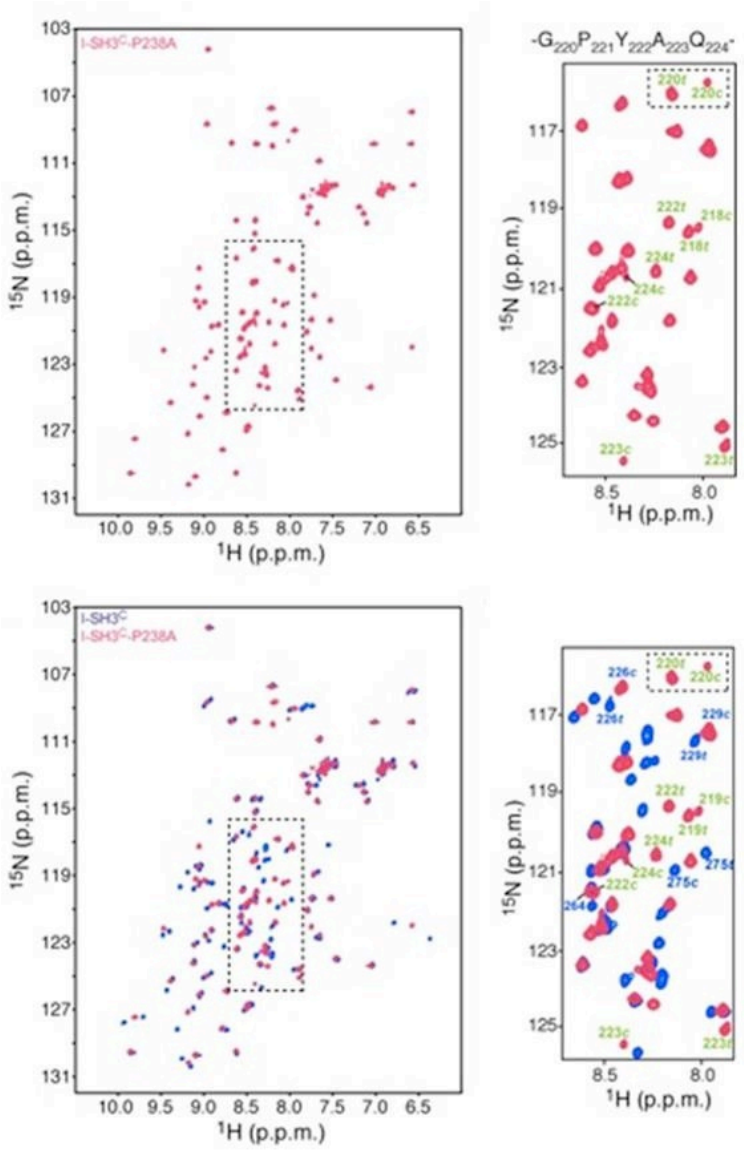

Figure 60. GP motif in chicken CrkII controls its cis-trans isomerization. (a) 1H-^{15}N HSQC NMR spectra of chicken-CrkII_P238A mutant, (b) magnified region of (a), (c) overlay of Chicken-CrkII (blue) and chicken-CrkII_P238A mutant (red), (d) magnified region of (c). (a),(b),(c) and (d) clearly demonstrate that the mutaation in the G237-P238 region disrupt the cis-trans isomerization, whereas cis-trans in the G220-P221 region remains unaffected.

NMR spectra of linker-$SH3^{C}$_P238A mutant of chicken-CrkII (Fig. 60a) show that cis-trans isomerization is not present at G237-P238, thus the mutant adopt one specific conformation. However, G220-P221 remains unaffected, and two sets of peaks in slow exchange can be observed (Fig. 60b). This indicates that these two isomerization sites are completely independent.

By using the ZZ exchange experiment we can characterize slow conformational exchange processes in proteins[279]. These experiments work by monitoring the exchange of longitudinal magnetization between the major and minor peaks as a function of time. Here, in the case where CypA is absent, no exchange peak is observed (Fig. 61, black). However, in the presence of a catalytic amount of CypA, exchange peaks appear on the spectrum (Fig. 61, orange and blue), providing direct evidence that CypA catalyzes the Gly219−Pro220cis−trans isomerization process.

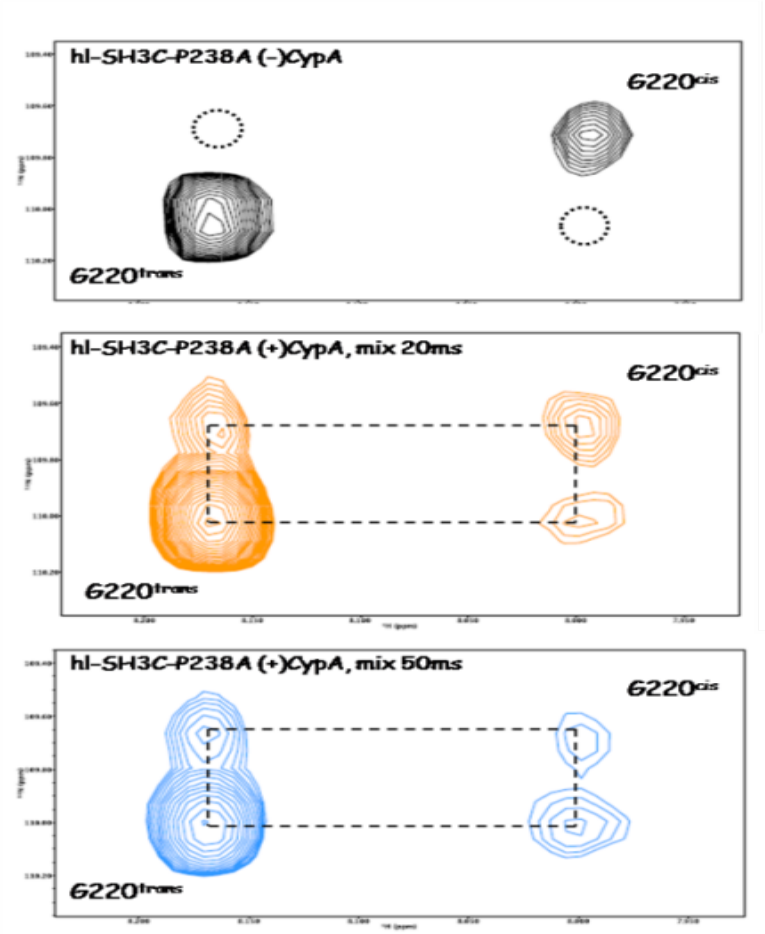

Figure 61. 2D ^{1}H-^{15}N ZZ NMR exchange experiment.CrkII fragment without the cyclophilin (black), with CypA and the mixing time of 20ms (orange) and with CypA and the mixing time of 50ms (blue). CypA was used in catalytic amount (~5%).

Using a corresponding human-CrkII construct, we observe that Pro220 (P221 in chicken) also shows the cis-trans effect (Fig. 62a,b), and could be a potential binding site for CypA. However, the Pro238 isomerization was not observed, and appears to be specific to chicken.

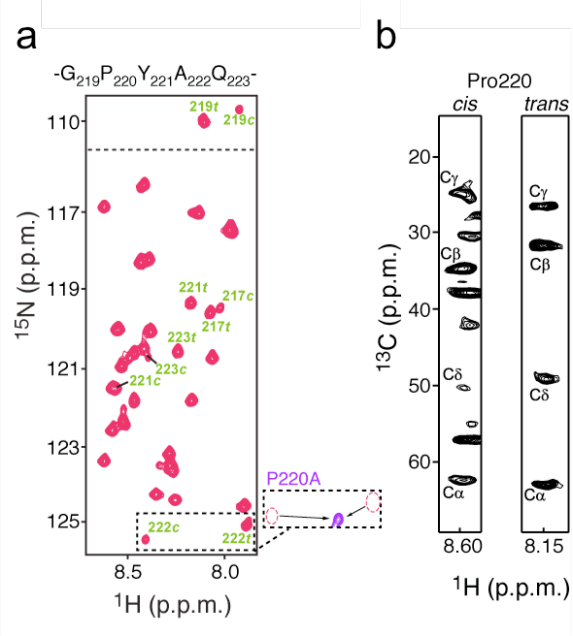

Figure 62. Human-CrkII undergoes cis-trans isomerization. (a) ^{1}H-^{15}N HSQC NMR spectra of human-CrkII_(188-303), (b) strips from CCO-NH spectra showing ^{13}C chemical shift of Pro220 for both cis and trans of the G220-P221 bond.

4.3.1.3 CrkII binds to the catalytic site of CypA

We decided to investigate the possible interaction between CrkII and CypA with CrkL serving as a control. NMR titration of CypA was performed for both CrkL and human-CrkII. The changes in NMR spectra were observed only when human-CrkII was titrated with CypA (Fig. 63a), whereas no changes were observed for CrkL (Fig. 63b). This supports our initial assessment from inspection of sequence analysis data. Moreover, chemical shift analysis shows that the human-CrkII G219-P220 motif is directly involved in the binding. (human-CrkII assignment, courtesy of Dr F. Inagaki)

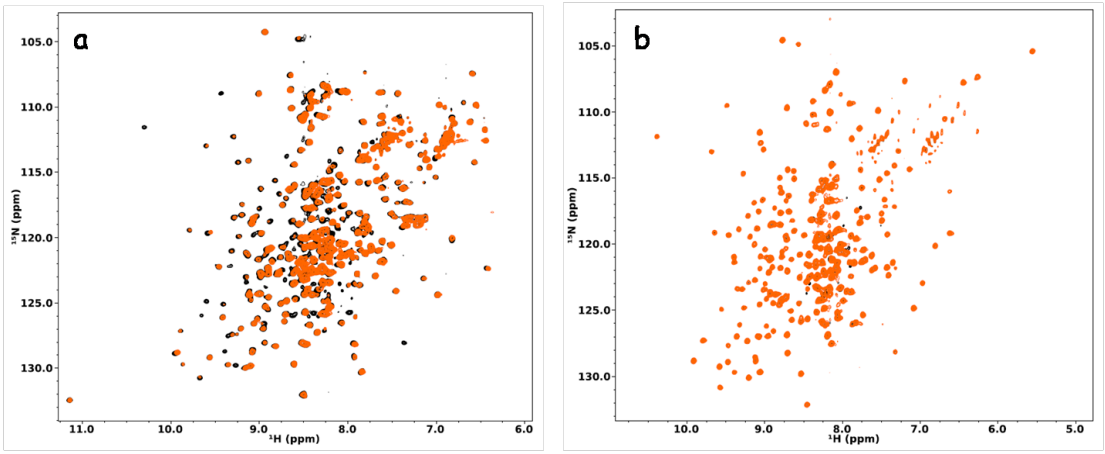

Figure 63. Figure 7.Binding of CypA to hCrkII and CrkL. ^{1}H-^{15}N HSQC NMR spectra of (a) free hCrkII (black), in complex with CypA (orange) and (b) free CrkL (black), in complex with CypA (orange). The hCrkII and CrkL are ^{15}N-labeled, whereas CypA is unlabeled.

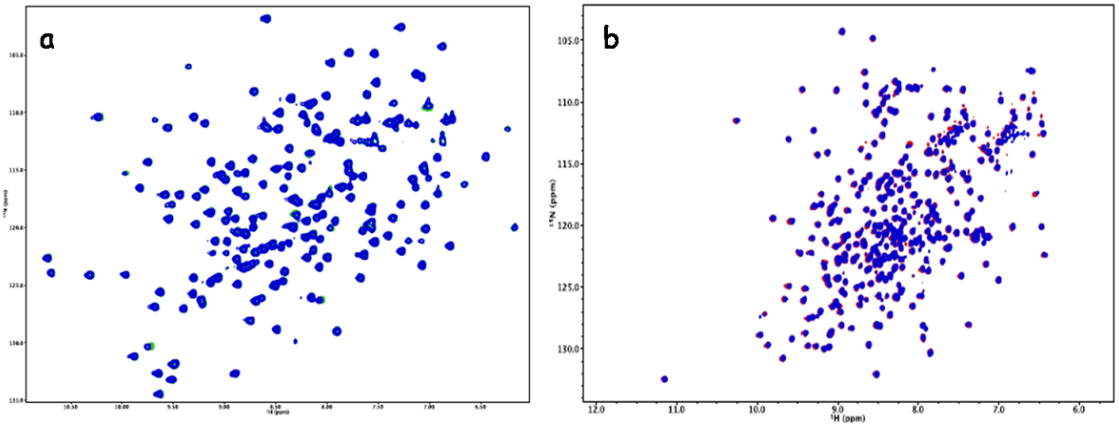

Figure 64. ^{1}H-^{15}N HSQC NMR spectra of (a) CypA (green), in complex with hCrkII_P220A and (b) hCrkII (red), hCrkII in presence of CsA. In (a) CypA is 15N-labeled and hCrkII_P220A mutant is unlabeled, in (b) hCrkII is 15N-labeled in free hCrkII and in bound to Csa form.

The mutational studies also agree with the above findings. Mutation P220A in CrkII completely abrogates the interaction between the CypA and CrkII (Fig. 64a). Also, the CrkII titration with CypA in the presence of Cyclosphorin A (CsA -

an inhibitor of CypA with nanomolar affinity)[276] displaces CypA from CrkII (Fig. 64b).

Reverse titration experiments were performed where ^{15}N- labeled CypA was titrated to unlabeled human-CrkII (Fig. 65). Chemical shift analysis showed that CrkII binds to the catalytic site of CypA.

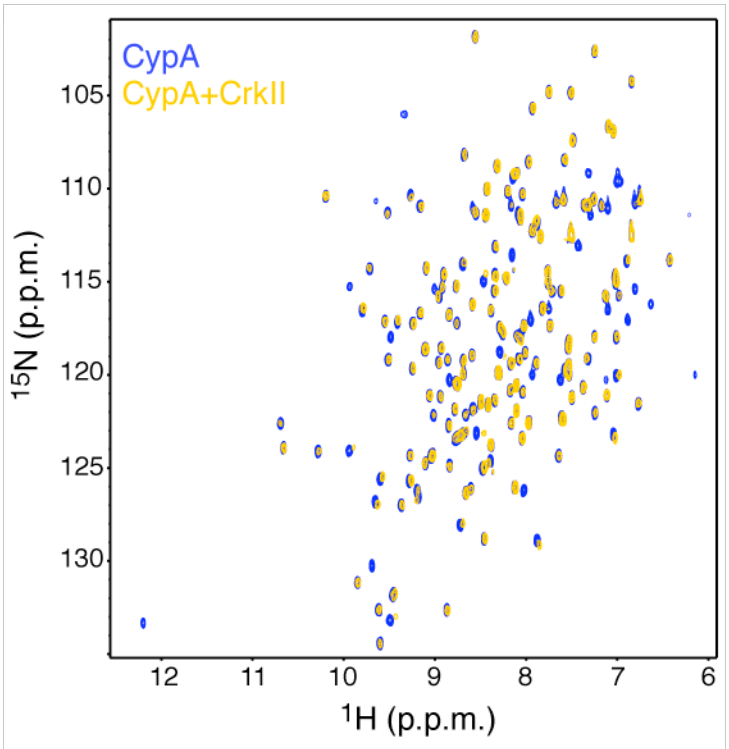

Figure 65. NMR titration of CypA with human-CrkII. ^{1}H-^{15}N HSQC NMR titration spectra between CypA (15N-labeled) and human-CrkII (unlabeled)

A titration experiment was also carried out to measure the binding affinity of CypA for the Gly219-Pro220 site. The chemical shift changes at each titration point were measured and the dissociation constant extracted (Fig. 66). The affinity was calculated to be ~18uM.

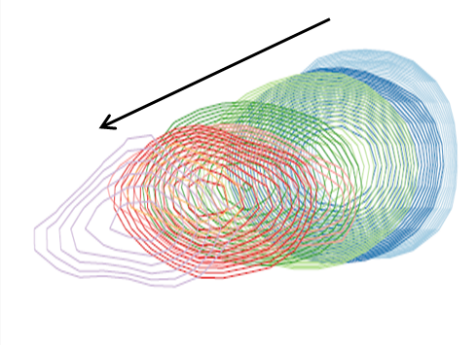

Figure 66. ^{15}N-labeled CypA was titrated with human-CrkII. The black arrow below demonstrate signal migration of CypA (W121) HE1 proton upon human-CrII binding.

4.3.1.4 Structural basis for the recognition ohuman-CrkII by CypA

To gain further insight into the structural basis of this interaction, we decided to solve the structure of CypA in complex with human-CrkII. Analysis of chemical shift in both human-CrkII and synthetic peptide (PEFGPYAQP - the region interacting with CypA) from human-CrkII shows that the effects induced either by the peptide or by full-length CrkII on the CypA site are essentially identical. Therefore, to simplify the NMR experiments, the synthetic peptide was used. C13-HSCQ NMR spectra of CypA with and without the peptide (Fig. 67) were used as a reference for the NOE assignment.

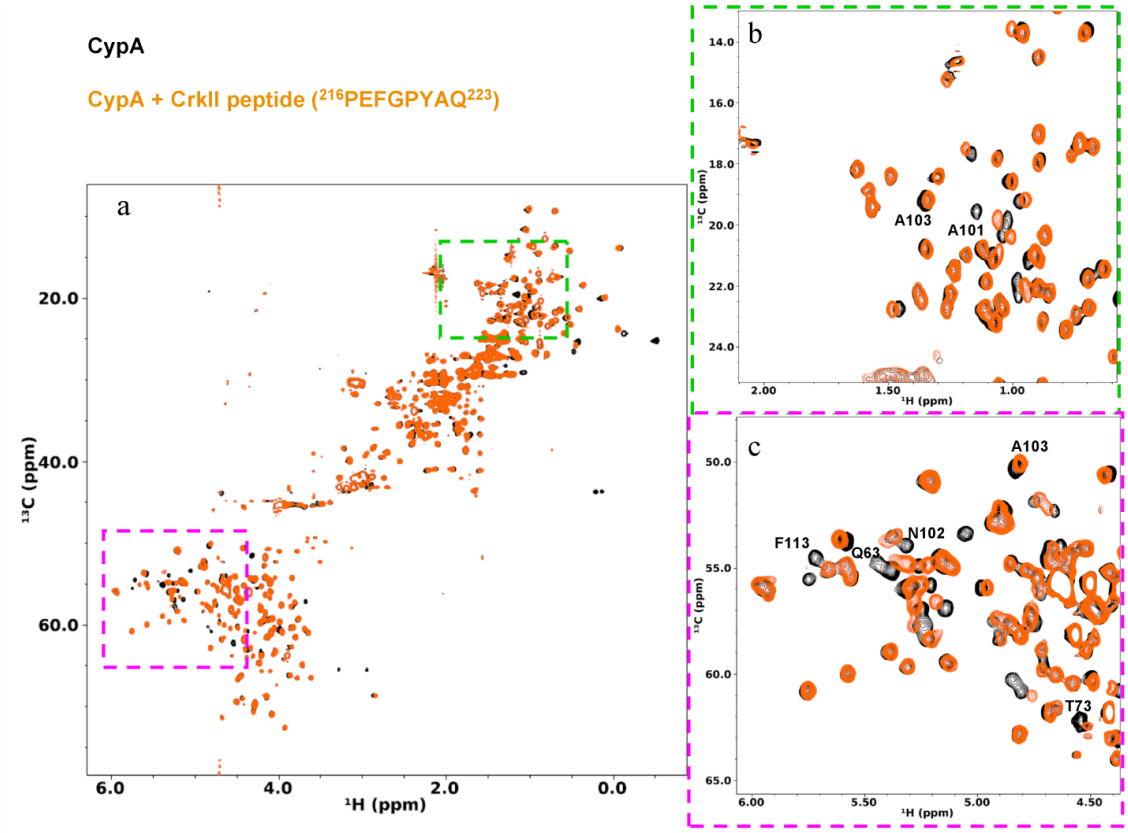

Figure 67. 13C-HSQC spectra of CypA (black) and in complex with CrkII peptide (orange). (a) Full spectral window, (b) methyl region of the spoectra, (c) C_α region, labeled residues belong to CypA catalytic site.

The NOE pattern observed revealed that the mode of interaction between CypA and CrkII is comparable to the interactions previously observed in CypA-ligand structures. Structures show that the Gly219−Pro220 motif of human-CrkII binds deep into the active site cleft of CypA (Fig. 68a and 68b).

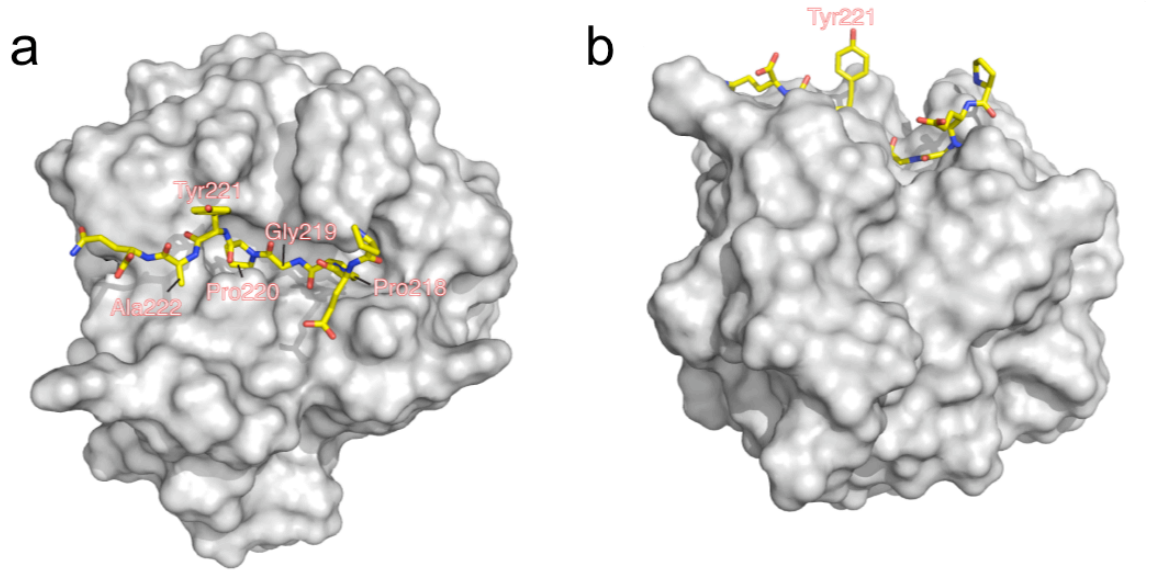

Figure 68. Structure of CypA bound to human-CrkII peptide. (a) top view, (b) side view. The Gly219−Pro220 region of CrkII binds deep into the active site cleft of CypA

Our observations confirm a previous study, where the GP motif of the HIV-capsid binds to CypA in the same fashion[280]. 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. 68a). The backbone carbonyl of Tyr221, the residue that becomes phosphorylated by theAbl kinase, forms a hydrogen bond with Trp121. In contrast, the side chain of Tyr221 is mostly exposed to the solvent (Fig. 68b) 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 human-CrkII linker region that encompasses the tyrosine-phosphorylation site. When

CypA binds to human-CrkII its Y221 is not accesible, therefore it can not be phosphorylated by Abl.

In order to further investigate the effect of CypA on human-CrkII activity, Western Blot (WB) experiments were performed (Fig. 69). Experiments show that the presence of CypA significantly decreases the level of human-CrkII Tyr221 phosphorylation by up to factor of ~12 (Fig. 69a). These findings are in agreement with structural data, where we show that Tyr221 sits in the CypA cleft and therefore cannot be sufficiently phosphorylated. Similar experiment with CrkL and CypA showed no inhibition of CrkL Tyr207. (Fig. 69b)

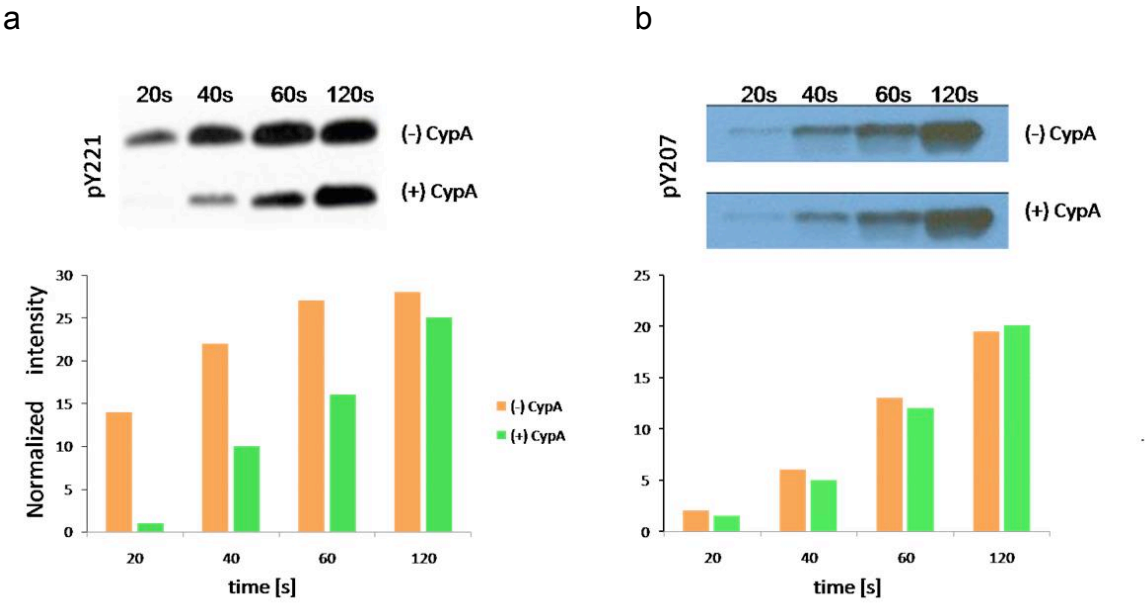

Figure 69. Immunoblotting analysis of kinase assay using human-CrkII and CkrL as a substrate for Abl in the presence and absence of CypA. (a) Tyr221 phosphorylation monitoring at different incubation times and normalized intensity graph to quantify the CypA effect on human-CrkII. (b) Tyr207 phosphorylation monitoring at different incubation times and normalized intensity graph to quantify the CypA effect on CrkL

To further test the notion of CypA mediation inhibition of CrkII Tyr221, we compared the effect of cyclosporine in EGF treated MDA-MB-468 human breast cancer cells. We observe a 3-fold increase in Tyr221 phosphorylation upon CsA treatment (Fig 70).

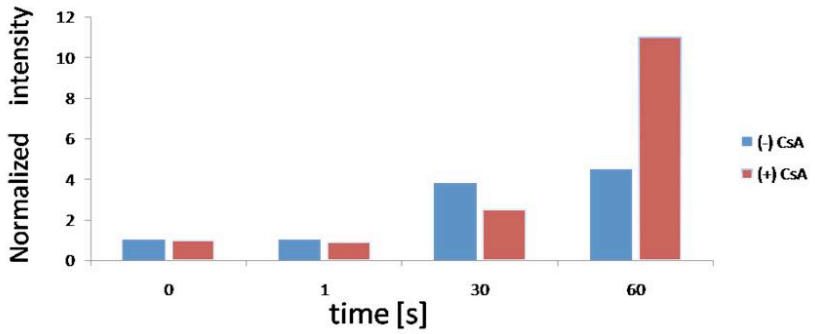

Figure 70. Effect of CsA on Tyr221 phosphorylation in MDA-MB-468 cells.

4.3.1.5 CypA and human-CrkII colocalization in living cells

Another important challenge is to confirm that CypA and human-CrkII can colocalize in living cells. FRET experiments were designed and performed. Using a pair of fluorescent proteins as donors or acceptors of FRET attached to the protein of interest we can directly monitor protein-protein interaction in living cells. By using CypA-CFP and human-CrkII-YFP constructs, we were able to see that CypA forms a constitutive complex with human-CrkII. The co-localized proteins show FRET effect with 50% efficiency (Fig. 71a). Conversely, when CsA is used, the complex dissociates and FRET effect is not perceived (Fig. 71b). This is in agreement with our previous findings.

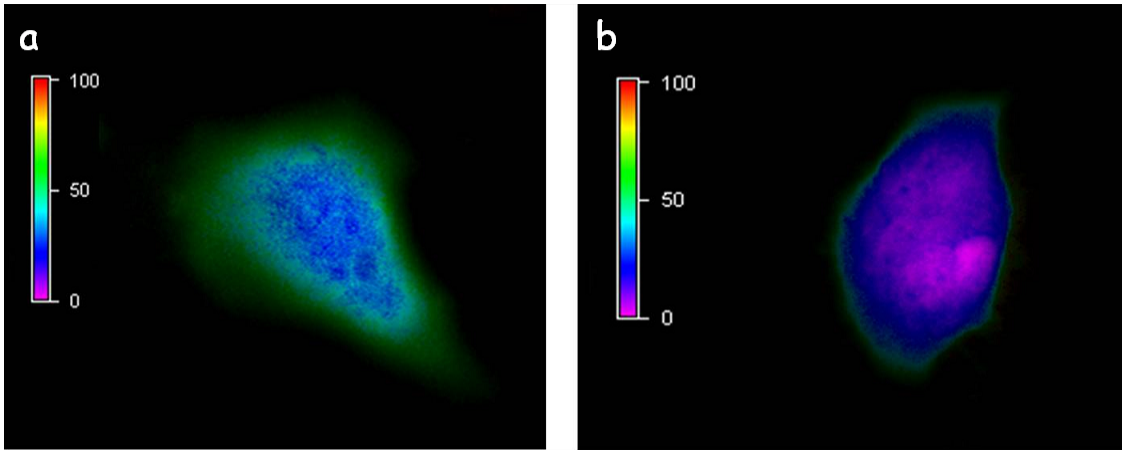

Figure 71. Fluorescence resonance energy transfer images of HeLa cellstransfected with (a) CypA-CFP and hCrKII-YFP, (b) CypA-CFP and hCrkII-YFP + CsA

4.4 Discussion

Here we investigated the effect of CypA on human-CrkII activity. By using a multi-technique approach, we have shown that human-CrkII recruits CypA not only as an enzyme but as a stoichiometric binding partner. The binding of CypA to the G219-P220 motif clearly decrease the phosphorylation efficiency of CrkII by Abl. This additional level of human-CrkII regulation, by capturing and delaying Tyr221 phosphorylation, could work as a switch between the autoinhibited and active conformations of CrkII. The dissociation constant for the CypA-CrkII complex is ~18uM. Although this binding affinity is moderate, increase in cellular concentration of either or both proteins will lead to an increased tendency for this complex to form. An important factor to remember here is that CrkII can be phosphorylated in multiple regions, including Try221. Recently published results from *Sriram et. al* show that Tyr251 can be phosphorylated in CrkII and that

region can activate Abl kinase[140]. It is unclear whether CrkII can participate in two separate pathways depending on its phosphorylation level. In the case of Tyr221 inhibition of phosphorylation by CypA would mean that a greater fraction of the cytosolic pool of CrkII can participate in signaling pathways that require Tyr221 in the unphosphorylated form.

Interestingly, Abl-mediated CrkL phosphorylation is not affected by CypA, even though CrkL and CrkII are homologs, with ~60% identity between the two proteins. How this might relate to specific Crk function will require further investigation. Since many signalling processes require phosphorylated human-CrkII, our results show that CypA may play an important role in regulation of the signaling pathway carried out by CrkII.

4.5 Materials and methods

4.5.1 Protein preparation, labeling and purification

4.5.1.1 Preparation and purification of CypA

The plasmid contained 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 37 °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, pH 6.5, 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 purity was checked by SDS-PAGE and concentration was determined spectrophotometrically at 280 nm using an extinction coefficient of 8470 M^{-1} cm^{-1}.

4.5.1.2 Preparation and purification of human-CrkII

The preparation of the various constructs used in this study were made using standard cloning protocols from a cDNA of human CrkII (courtesy of Dr. Birge, UMDNJ). The coding region of the full-length protein and the fragments were inserted into pet42a vector (EMDChemicals) using the Nco1 and Xho1 restriction

sites. The pet42a vector allows expression of proteins of interest with a GST fusion tag at the N-terminal along with a His6-tag. This allowed us to use a more efficient purification protocol using fast flow Ni-Sepharose 6 Fast Flow resin (GE healthcare) instead of the GSH Sepharose resin. The SH2(1-120), $SH3^{N}$(130-190), $SH3^{C}$(238-304), linker-SH3C(188-304) $SH3^{N}$-linker-$SH3^{C}$(130-304) and the full length (1-304) of CrkII were successfully cloned and expressed. We also included a TEV protease cleavage site to separate the tag from the protein of interest. All constructs were sequenced and expression in BL 21(DE3) confirmed for expression of protein of expected size and solubility. For LB, N15 and C13 samples a common protocol was used where a single colony was used to inoculate a starter culture of 5 ml LB and after ~6h at 37°C cells are centrifuge at 4000rpm and resuspended in 50 ml M9 minimal medium containing 2g/L 13C6-glucose and 1g/L 15NH4Cl. After overnight at 37°C the culture is centrifuge at low speed and transferred to freshly prepared 1L minimal medium and grown at 37°C. After OD 600 of 0.4 is reached protein synthesis is induced with 0.5mM IPTG. Cells are harvested 3 to 4 hours after induction and resuspended in lysis buffer (50mMTris-HCL, 1M NaCL, 30mM imidazole, 5mM beta-mercaptoethanol) The protein prepared in such a way is uniformly 15N,13C labeled. The cells are lysed and the cytosolic fraction is separated from the membrane fraction by centrifugation at 20,000 x g. Protein is purified on Ni-NTA resin and after cleavage with TEV protease to separate the GST-His6 tag from the protein, the protein was passed over Ni-NTA again to remove the tag and protease. The protein is found in the flow-through. The protein was further purified on a

Superdex-75 gel-filtration column. Fractions are checked by SDS-Page and dialyzed against NMR buffer (50mM Phosphate, 150mM NaCl, 0.5mM EDTA, 3mM β-mercaptoethanol) using Amicon cell (Millipore).

4.5.1.3 Preparation and purification of CrkL1-303

As described in Chapter 3 (see 3.4.3).

4.5.1.4 Preparation and purification of Abl

As described in Chapter 3 (see 3.4.3).

4.5.2 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 6.5),

140 mM potassium phosphate, 1mM BME and 1g L-1 NaN3. Concentration of all protein NMR sample was 0.6-0.8mM. Spectra were recorded at 25 °C. 2D 1H-15N 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 pH 6.5 and temperature 25°C.

4.5.3 FRET Study

4.5.3.1 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 subcloned 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.

4.5.3.2 Cell Culture

HELA cells were cultured 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-treme GENE 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.

4.5.3.3 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.

4.5.4 Western Blot

We performed western blot by using purified protein. For detection of phophorylated Tyr222 or Tyr207 we use primary antibody, phospho-CrkII (Tyr222) or phospho-CrkL (Tyr207) antibody from Cell Signaling. We make ~6-7 reaction mixture depending on our need. 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 we aliquot human-CrkII and CrkL into the eppendorfs and then in excess (~6-7 times) add CypA for complete saturation of both proteins. In another eppendorf we make c-Abl reaction mixture. First we add c-Abl and then to it we add 1mM MgCl2, 0.6mM ATP and immediately add it to the samples, followed by collection of each sample at different interval of time and stopping the reaction by using SDS-loading buffer. All the samples were run on a 10% SDS gel. In the next step we transfer proteins

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]. Then we place the nitrocellulose membrane into blocking solution (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. Then we wash for 2-3 times with 1X TBST buffer. Then we incubate the membrane in primar antibody solution (substrate to antibody ratio; 1:1000) for overnight. Next morning we wash out exces of primary antibody by using 1X TBST buffer for 3-4 times. We then incubate it with secondary antibody in blocking solution for further one more hour. Secondary 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.

4.5.5 MDA-MB-468 cells Western-Botting

MDA-MB-468 human breast cancer cells (EGFR overexpressing) were serum-starved overnight and then pre-treated with DMSO or CsA (15 uM) for 30 minutes. Cells were then stimulated with EGF (100 ng/ml) for 1, 30 or 60 seconds and lysates were analyzed by western-blotting (see 4.5.4) with anti-pY221 Crk, anti-Crk or anti-Actin antibodies.

References

[1] Birge RB, Kalodimos C, Inagaki F, Tanaka S. Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. Cell Commun Signal. 2009 May 10;7:13.

[2] I Isakov N. A new twist to adaptor proteins contributes to regulation of lymphocyte cell signaling. Trends Immunol. 2008 Aug;29(8):388-96. Epub 2008 Jul 1.

[3] Cabodi S, del Pilar Camacho-Leal M, Di Stefano P, Defilippi P. Integrin signalling adaptors: not only figurants in the cancer story. Nat Rev Cancer. 2010 Dec;10(12):858-70.

[4] Mayer BJ, Hamaguchi M, Hanafusa H. A novel viral oncogene with structural similarity to phospholipase C. Nature. 1988 Mar 17;332(6161):272-5.

[5] Mayer BJ, Hamaguchi M, Hanafusa H: Characterization of p47gag-crk, a novel oncogene product with sequence similarity to a putative modulatory domain of protein-tyrosine kinases and phospholipase C. Cold Spring Harb Symp Quant Biol 1988, 53(Pt 2):907-914

[6] Matsuda M, Mayer BJ, Hanafusa H. Identification of domains of the v-crk oncogene product sufficient for association with phosphotyrosine-containing proteins. Mol Cell Biol. 1991 Mar;11(3):1607-13.

[7] Matsuda M, Mayer BJ, Fukui Y, Hanafusa H. Binding of transforming protein, P47gag-crk, to a broad range of phosphotyrosine-containing proteins. Science. 1990 Jun 22;248(4962):1537-9.

[8] Tsuchie H, Chang CH, Yoshida M, Vogt PK. A newly isolated avian sarcoma virus, ASV-1, carries the crk oncogene. Oncogene. 1989 Nov;4(11):1281-4.

[9] Reichman CT, Mayer BJ, Keshav S, Hanafusa H. The product of the cellular crk gene consists primarily of SH2 and SH3 regions. Cell Growth Differ. 1992 Jul;3(7):451-60.

[10] Matsuda M, Tanaka S, Nagata S, Kojima A, Kurata T, Shibuya M. Two species of human CRK cDNA encode proteins with distinct biological activities. Mol Cell Biol. 1992 Aug;12(8):3482-9.

[11] ten Hoeve J, Morris C, Poustka A, Groffen J, Heisterkamp N. Isolation of NotI sites from chromosome 22q11. Genomics. 1993 Dec;18(3):588-97.

---

[12] Pawson T, Scott JD. Signaling through scaffold, anchoring, and adaptor proteins. Science. 1997 Dec 19;278(5346):2075-80.

[13] Pawson T. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. Cell. 2004 Jan 23;116(2):191-203.

[14] Feller SM, Posern G, Voss J, Kardinal C, Sakkab D, Zheng J, Knudsen BS. Physiological signals and oncogenesis mediated through Crk family adapter proteins. J Cell Physiol. 1998 Dec;177(4):535-52.

[15] Smith JJ, Richardson DA, Kopf J, Yoshida M, Hollingsworth RE, Kornbluth S. Apoptotic regulation by the Crk adapter protein mediated by interactions with Wee1 and Crm1/exportin. Mol Cell Biol. 2002 Mar;22(5):1412-23.

[16] Ogawa S, Toyoshima H, Kozutsumi H, Hagiwara K, Sakai R, Tanaka T, Hirano N, Mano H, Yazaki Y, Hirai H. The C-terminal SH3 domain of the mouse c-Crk protein negatively regulates tyrosine-phosphorylation of Crk associated p130 in rat 3Y1 cells. Oncogene. 1994 Jun;9(6):1669-78.

[17] Akakura S, Kar B, Singh S, Cho L, Tibrewal N, Sanokawa-Akakura R, Reichman C, Ravichandran KS, Birge RB. C-terminal SH3 domain of CrkII regulates the assembly and function of the DOCK180/ELMO Rac-GEF. J Cell Physiol. 2005 Jul;204(1):344-51.

[18] Zvara A, Fajardo JE, Escalante M, Cotton G, Muir T, Kirsch KH, Birge RB. Activation of the focal adhesion kinase signaling pathway by structural alterations in the carboxyl-terminal region of c-Crk II. Oncogene. 2001 Feb 22;20(8):951-61.

[19] Muralidharan V, Dutta K, Cho J, Vila-Perello M, Raleigh DP, Cowburn D, Muir TW. Solution structure and folding characteristics of the C-terminal SH3 domain of c-Crk-II. Biochemistry. 2006 Jul 25;45(29):8874-84.

[20] de Jong R, ten Hoeve J, Heisterkamp N, Groffen J. Tyrosine 207 in CRKL is the BCR/ABL phosphorylation site. Oncogene. 1997 Feb 6;14(5):507-13.

[21] Feller SM. Crk family adaptors-signaling complex formation and biological roles. Oncogene. 2001 Oct 1;20(44):6348-71.

[22] Sawada Y, Tamada M, Dubin-Thaler BJ, Cherniavskaya O, Sakai R, Tanaka S, Sheetz MP. Force sensing by mechanical extension of the Src family kinase substrate p130Cas. Cell. 2006 Dec 1;127(5):1015-26.

[23] Defilippi P, Di Stefano P, Cabodi S. p130Cas: a versatile scaffold in signaling networks. Trends Cell Biol. 2006 May;16(5):257-63.

[24] Noren NK, Foos G, Hauser CA, Pasquale EB. The EphB4 receptor suppresses breast cancer cell tumorigenicity through an Abl-Crk pathway. Nat Cell Biol. 2006 Aug;8(8):815-25.

[25] Liu BA, Shah E, Jablonowski K, Stergachis A, Engelmann B, Nash PD. The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. Sci Signal. 2011 Dec 6;4(202):ra83.

[26] Waksman G, Kumaran S, Lubman O. SH2 domains: role, structure and implications for molecular medicine. Expert Rev Mol Med. 2004 Jan 30;6(3):1-18.

[27] Waksman G, Kumaran S, Lubman O. SH2 domains: role, structure and implications for molecular medicine. Expert Rev Mol Med. 2004 Jan 30;6(3):1-18.

[28] Sadowski I, Stone JC, Pawson T. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps. Mol Cell Biol. 1986 Dec;6(12):4396-408.

[29] Waksman G, Kominos D, Robertson SC, Pant N, Baltimore D, Birge RB, Cowburn D, Hanafusa H, Mayer BJ, Overduin M, Resh MD, Rios CB, Silverman L, Kuriyan J. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. Nature. 1992 Aug 20;358(6388):646-53.

[30] Bullock AN, Rodriguez MC, Debreczeni JE, Songyang Z, Knapp S. Structure of the SOCS4-ElonginB/C complex reveals a distinct SOCS box interface and the molecular basis for SOCS-dependent EGFR degradation. Structure. 2007 Nov;15(11):1493-504.

[31] Liu BA, Jablonowski K, Shah EE, Engelmann BW, Jones RB, Nash PD. SH2 domains recognize contextual peptide sequence information to determine selectivity. Mol Cell Proteomics. 2010 Nov;9(11):2391-404.

[32] Lander, E.S et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860-921.

[33] Zarrinpar A, Bhattacharyya RP, Lim WA. The structure and function of prolinerecognition domains. Sci STKE. 2003 Apr 22;2003(179):RE8.

[34] Mayer BJ. SH3 domains: complexity in moderation. J Cell Sci. 2001

Apr;114(Pt 7):1253-63. (2001)

[35] Li SS. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. Biochem J. 2005 Sep 15;390(Pt 3):641-53.

[36] Yu H, Rosen MK, Shin TB, Seidel-Dugan C, Brugge JS, Schreiber SL. Solution structure of the SH3 domain of Src and identification of its ligand-binding site.Science. 1992 Dec 4;258(5088):1665-8.

[37] Musacchio A, Noble M, Pauptit R, Wierenga R, Saraste M. Crystal structure of a Src-homology 3 (SH3) domain. Nature. 1992 Oct 29;359(6398):851-5.

[38] Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. Cicchetti P, Mayer BJ, Thiel G, Baltimore D. Science. 1992 Aug 7;257(5071):803-6.

[39] Ren R, Mayer BJ, Cicchetti P, Baltimore D. Identification of a ten-amino acid proline-rich SH3 binding site. Science. 1993 Feb 19;259(5098):1157-61.

[40] Chen, J. K., W. S. Lane, A. W. Brauer, A. Tanaka, S. L. Schreiber. Biased combinatorial libraries: novel ligands for the SH3 domain of phosphatidylinositol 3-kinase. J Am Chem Soc 1993, 115, 12591-12592

[41] Kaneko T, Li L, Li SS. The SH3 domain--a family of versatile peptide- and protein-recognition module. Front Biosci. 2008 May 1;13:4938-52.

[42] Wu C, Ma MH, Brown KR, Geisler M, Li L, Tzeng E, Jia CY, Jurisica I, Li SS. Systematic identification of SH3 domain-mediated human protein-protein interactions by peptide array target screening. Proteomics. 2007 Jun;7(11):1775-85.

[43] Kaneko T, Li L, Li SS. The SH3 domain--a family of versatile peptide- and protein-recognition module. Front Biosci. 2008 May 1;13:4938-52.

[44] Tanaka S, Morishita T, Hashimoto Y, Hattori S, Nakamura S, Shibuya M, Matuoka K, Takenawa T, Kurata T, Nagashima K, et al. C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. Proc Natl Acad Sci U S A. 1994 Apr 12;91(8):3443-7.

[45] Reynolds AB, Kanner SB, Wang HC, Parsons JT. Stable association of activated pp60src with two tyrosine-phosphorylated cellular proteins. Mol Cell Biol. 1989 Sep;9(9):3951-8.

[46] Bouton AH, Riggins RB, Bruce-Staskal PJ. Functions of the adapter protein Cas: signal convergence and the determination of cellular responses. Oncogene 2001; 20: 6448–6458.

[47] Tornillo G, Bisaro B, Camacho-Leal Mdel P, Galiè M, Provero P, Di Stefano P, Turco E, Defilippi P, Cabodi S. p130Cas promotes invasiveness of three-dimensional ErbB2-transformed mammary acinar structures by enhanced activation of mTOR/p70S6K and Rac1. Eur J Cell Biol. 2011 Feb-Mar;90(2-3):237-48.

[48] Honda H, Oda H, Nakamoto T, Honda Z, Sakai R, Suzuki T, Saito T, Nakamura K, Nakao K, Ishikawa T, Katsuki M, Yazaki Y, Hirai H. Cardiovascular anomaly, impaired actin bundling and resistance to Src-induced transformation in mice lacking p130Cas. Nat Genet. 1998 Aug;19(4):361-5.

[49] Cabodi S, Defilippi P. The essence of integrin signal transduction. In: Danen E, ed. Georgetown, TX: Landes/Eurekah, PubMed. 2005.

[50] Gustavsson A, Yuan M, Fällman M. Temporal dissection of beta1-integrin signaling indicates a role for p130Cas-Crk in filopodia formation. J Biol Chem. 2004 May 28;279(22):22893-901.

[51] Webb DJ, Donais K, Whitmore LA, Thomas SM, Turner CE, Parsons JT, Horwitz AF. FAK-Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly. Nat Cell Biol. 2004 Feb;6(2):154-61.

[52] Kiyokawa E, Hashimoto Y, Kobayashi S, Sugimura H, Kurata T, Matsuda M. Activation of Rac1 by a Crk SH3-binding protein, DOCK180. Genes Dev. 1998 Nov 1;12(21):3331-6.

[53] Cheresh DA, Leng J, Klemke RL. Regulation of cell contraction and membrane ruffling by distinct signals in migratory cells. J Cell Biol. 1999 Sep 6;146(5):1107-16.

[54] Cho SY, Klemke RL. Extracellular-regulated kinase activation and CAS/Crk coupling regulate cell migration and suppress apoptosis during invasion of the extracellular matrix. J Cell Biol. 2000 Apr 3;149(1):223-36.

[55] Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE. Reversible unfolding of individual titin immunoglobulin domains by AFM. Science. 1997 May 16;276(5315):1109-12.

[56] Patwardhan P, Shen Y, Goldberg GS, Miller WT: Individual Cas phosphorylation sites are dispensable for processive phosphorylation by Src and anchorage-independent cell growth. J Biol Chem 2006, 281:20689-20697.

[57] Kain KH, Klemke RL. Inhibition of cell migration by Abl family tyrosine kinases through uncoupling of Crk-CAS complexes. J Biol Chem. 2001 May 11;276(19):16185-92.

[58] Takino T, Tamura M, Miyamori H, Araki M, Matsumoto K, Sato H, Yamada KM. Tyrosine phosphorylation of the CrkII adaptor protein modulates cell migration. J Cell Sci. 2003 Aug 1;116(Pt 15):3145-55.

[59] Glenney JR Jr and Zokas L. Novel tyrosine kinase substrates from Rous sarcoma virus-transformed cells are present in the membrane skeleton. *J Cell Biol* 108: 2401–2408, 1989.

[60] Turner CE, Glenney JR Jr, and Burridge K. Paxillin: a new vinculin-binding protein present in focal adhesions. *J Cell Biol* 111:1059–1068, 1990.

[61] Turner CE and Miller JT. Primary sequence of paxillin contains putative SH2 and SH3 domain binding motifs and multiple LIM domains: identification of a vinculin and pp125Fak-binding region. *J Cell Sci* 107: 1583–1591, 1994.

[62] Brown MC, Perrotta JA, and Turner CE. Identification of LIM3 as the principal determinant of paxillin focal adhesion localization and characterization of a novel motif on paxillin directing vinculin and focal adhesion kinase binding. *J Cell Biol* 135: 1109–1123, 1996.

[63] Schaller, M. D. & Parsons, J. T. pp125FAK-dependent tyrosine phosphorylation of paxillin creates a high-affinity binding site for Crk. *Mol. Cell. Biol.* 15, 2635–2645 (1995).

[64] Bellis, S. L., Miller, J. T. & Turner, C. E. Characterization of tyrosine phosphorylation of paxillin *in vitro* by focal adhesion kinase. *J. Biol. Chem.* 270, 17437–17441 (1995).

[65] Turner, C. E. Paxillin. *Int. J. Biochem. Cell Biol.* 30, 955–959 (1998).

[66] Feller, S. M. *et al*. Physiological signals and oncogenesis mediated through Crk family adaptor proteins. *J. Cell Physiol.* 177, 535–552 (1998).

[67] Petit V, Boyer B, Lentz D, Turner CE, Thiery JP, and Valles AM. Phosphorylation of tyrosine residues 31 and 118 on paxillin regulates cell migration through an association with CRK in NBT-II cells. *J Cell Biol* 148: 957–970, 2000.

[68] Klemke, R. L. *et al*. CAS/Crk coupling serves as a “molecular switch” for induction of cell migration. *J. Cell Biol.* 140, 961–972 (1998).

[69] Dawid IB, Breen JJ, and Toyama R. LIM domains: multiple roles as adapters and functional modifiers in protein interactions. *Trends Genet* 14: 156–162, 1998.

[70] Khurana T, Khurana B, and Noegel AA. LIM proteins: association with the actin cytoskeleton. *Protoplasma* 219: 1–12, 2002.

[71] Tumbarello DA, Brown MC, and Turner CE. The paxillin LD motifs. *FEBS Lett* 513: 114–118, 2002.

[72] Coté, J. F., Turner, C. E. & Tremblay, M. L. 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.* 274, 20550–20560 (1999).

[73] Shen, Y. *et al*. The noncatalytic domain of protein-tyrosine phosphatase-PEST targets paxillin for dephosphorylation *in vivo*. *J. Biol. Chem.* 275, 1405–1413 (2000).

[74] Lewis, J. M., Baskaran, R., Taagepera, S., Schwartz, M. A. & Wang, J. Y. Integrin regulation of c-Abl tyrosine kinase activity and cytoplasmic- nuclear transport. *Proc. Natl Acad. Sci. USA* 93, 15174–15179 (1996).

[75] Sattler, M., Pisick, E., Morrison, P. T. & Salgia, R. Role of the cytoskeletal protein paxillin in oncogenesis. *Crit. Rev. Onc.* 11, 63–76 (2000).

[76] Pannekoek WJ, Kooistra MR, Zwartkruis FJ, Bos JL (2009) Cell-cell junction formation: the role of Rap1 and Rap1 guanine nucleotide exchange factors. Biochim Biophys Acta 1788: 790–796.

[77] Knudsen BS, Feller SM, Hanafusa H (1994) Four proline-rich sequences of the guanine-nucleotide exchange factor C3G bind with unique specificity to the first Src homology 3 domain of Crk. J Biol Chem 32781–32787.

[78] Kirsch K.H., Georgescu,M.M. and Hanafusa,H. (1998) Direct binding of p130(Cas) to the guanine nucleotide exchange factor C3G. J. Biol. Chem., 273, 25673–25679.

[79] Uemura, N. and Griffin, J. D. (1999). The adapter protein Crkl links Cbl to C3G after integrin ligation and enhances cell migration. *J. Biol. Chem.* 274, 37525-37532.

[80] Ballif, B. A., Arnaud, L., Arthur, W. T., Guris, D., Imamoto, A. and Cooper, J. A. (2004). Activation of a Dab1/CrkL/C3G/Rap1 pathway in Reelin-stimulated neurons. *Curr. Biol.* 14, 606-610

[81] Tanaka S., and Hanafusa H. (1998) Guanine-nucleotide exchange protein C3G activates JNK1 by a ras- independent mechanism. JNK1 activation inhibited by kinase negative forms of MLK3 and DLK mixed lineage kinases. J. Biol. Chem., *273*:1281–1284.

[82] Mochizuki, N. *et al*. Crk activation of JNK via C3G and R-Ras. J. Biol. Chem. 275, 12667–12671 (2000).

[83] Guerrero C, Martín-Encabo S, Fernández-Medarde A, Santos E. C3G-mediated suppression of oncogene-induced focus formation in fibroblasts involves inhibition of ERK activation, cyclin A expression and alterations of anchorage-independent growth. Oncogene. 2004 Jun 17;23(28):4885-93.

[84] Ichiba, T., Kuraishi, Y., Sakai, O., Nagata, S., Groffen, J., Kurata, T., Hattori, S. and Matsuda, M. (1997). Enhancement of guanine-nucleotide exchange activity of C3G for Rap1 by the expression of Crk, CrkL, and Grb2. *J. Biol. Chem.* 272, 22215-22220.

[85] Ohba, Y., Ikuta, K., Ogura, A., Matsuda, J., Mochizuki, N., Nagashima, K., Kurokawa, K., Mayer, B. J., Maki, K., Miyazaki, J. et al. (2001). Requirement for C3G-dependent Rap1 activation for cell adhesion and embryogenesis. *EMBO J.* 20, 3333-3341.

[86] Sakkab, D. *et al*. Signaling of hepatocyte growth factor/scatter factor (HGF) to the small GTPase Rap1 via the large docking protein Gab1 and the adapter protein CRKL. J. Biol. Chem. 275, 10772– 10778 (2000).

[87] Bos JL, de Rooij J, Reedquist KA. Rap1 signaling: adhering to new models. Nat Rev Mol Cell Biol. 2001 May;2(5):369-77.

[88] York RD, Yao H, Dillon T, Ellig CL, Eckert SP, McCleskey EW, Stork PJ: Rap1 mediates sustained MAP kinase activation induced by nerve growth factor. *Nature* 1998, 392:622-626.

[89] Bonfini L, Karlovich CA, Dasgupta C, Banerjee U. The Son of sevenless gene product: a putative activator of Ras. Science 1992;255:603–6.

[90] Horvitz HR, Sternberg PW. Multiple intercellular signaling systems control the development of the Caenorhabditis elegans vulva. Nature 1991;351:535–41.

[91] Bowtell D, Fu P, Simon M, Senior P. Identification of marine homologues of the Drosophila son of sevenless gene: potential activators of ras. Proc Natl Acad Sci USA 1992;89:6511–5.

[92] Fath I, Apiou F, Schweighoffer F, Chevallier-Multon MC, Ciora T, Dutrillaux B, et al. Identification of two human homologues to Drosophila SOS (son of sevenless) localized on two different chromosomes. Nucleic Acids Res 1993;21:4398.

[93] Webb GC, Jenkins NA, Largaespada DA, Copeland NG, Fernandez CS, Bowtell DD. Mammalian homologues of the Drosophila Son of sevenless gene map to marine chromosomes 17 and 12 and to human chromosomes 2 and 14, respectively. Genomics 1993;18:14–9.

[94] Chardin P, Mattei MG. Chromosomal localization of two genes encoding human ras exchange factors: SOS1 maps to the 2p22 ! p16 region and SOS2 to the 14q21 ! q22 region of the human genome. Cytogenet Cell Genet 1994;66:68–9.

[95] Jorge R, Zarich N, Oliva JL, Azanedo M, Martinez N, de la Cruz X, et al. HSos1 contains a new amino-terminal regulatory motif with specific binding affinity for its Pleckstrin homology domain. J Biol Chem 2002;277:44171–9.

[96] Sondermann H, Soisson SM, Boykevisch S, Yang SS, Bar-Sagi D, Kuriyan J. Structural analysis of autoinhibition in the Ras activator Son of Sevenless. Cell. 2004;119:393-405.

[97] Feller SM, Knudsen B, Hanafusa H. Cellular proteins binding to the first Src homology 3 (SH3) domain of the proto-oncogene product c-Crk indicate Crk-specific signaling pathways. Oncogene. 1995 Apr 20;10(8):1465-73.

[98] Matsuda M, Hashimoto Y, Muroya K, Hasegawa H, Kurata T, Tanaka S, Nakamura S, Hattori S. CRK protein binds to two guanine nucleotide-releasing proteins for the Ras family and modulates nerve growth factor-induced activation of Ras in PC12 cells. Mol Cell Biol. 1994 Aug;14(8):5495-500.

[99] Feller SM, Knudsen B, Wong TW, Hanafusa H: Detection of SH3- binding proteins in total cell lysates with glutathione S-transferase- SH3 fusion proteins: SH3 blot assay. Methods Enzymol 1995, 255:369-378.

[100] Hasegawa H, Kiyokawa E, Tanaka S, Nagashima K, Gotoh N, Shibuya M, Kurata T, Matsuda M: DOCK180, a major CRK-binding protein, alters cell morphology upon translocation to the cell membrane. Mol Cell Biol 1996, 16:1770-1776.

[101] Gumienny TL, Brugnera E, Tosello-Trampont AC, Kinchen JM, Haney LB, Nishiwaki K, Walk SF, Nemergut ME, Macara IG, Francis R, Schedl T, Qin Y, Van Aelst L, Hengartner MO, Ravichandran KS. CED-12/ELMO, a novel member of the CrkII/Dock180/Rac pathway, is required for phagocytosis and cell migration. Cell. 2001 Oct 5;107(1):27-41.

[102] Zhou Z, Caron E, Hartwieg E, Hall A, Horvitz HR. The C. elegans PH domain protein CED-12 regulates cytoskeletal reorganization via a Rho/Rac GTPase signaling pathway. Dev Cell. 2001 Oct;1(4):477-89.

[103] Dolfi F, Garcia-Guzman M, Ojaniemi M, Nakamura H, Matsuda M, Vuori K. The adaptor protein Crk connects multiple cellular stimuli to the JNK signaling pathway. Proc Natl Acad Sci U S A. 1998 Dec 22;95(26):15394-9.

[104] Schmidt A, Hall A. Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. Genes Dev. 2002 Jul 1;16(13):1587-609.

[105] Vetter IR, Wittinghofer A. The guanine nucleotide-binding switch in three dimensions. Science. 2001 Nov 9;294(5545):1299-304.

[106] Cote JF, Vuori K: Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. J Cell Sci 2002, 115:4901-4913

[107] Brugnera E, Haney L, Grimsley C, Lu M, Walk SF, Tosello-TrampontbAC, Macara IG, Madhani H, Fink GR, Ravichandran KS: Unconventional Rac-GEF activity is mediated through the Dock180-bELMO complex. Nat Cell Biol 2002, 4:574-582.

[108] Lu M, Kinchen JM, Rossman KL, Grimsley C, deBakker C, Brugnera E, Tosello-Trampont AC, Haney LB, Klingele D, Sondek J, Hengartner MO, Ravichandran KS: PH domain of ELMO functions in trans to regulate Rac activation via Dock180. Nat Struct Mol Biol 2004, 11:756-762

[109] Grimsley CM, Kinchen JM, Tosello-Trampont AC, Brugnera E, Haney LB, Lu M, Chen Q, Klingele D, Hengartner MO, Ravichandran KS. Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration. J Biol Chem. 2004 Feb 13;279(7):6087-97.

[110] Côté JF, Motoyama AB, Bush JA, Vuori K. A novel and evolutionarily conserved PtdIns(3,4,5)P3-binding domain is necessary for DOCK180 signalling. Nat Cell Biol. 2005 Aug;7(8):797-807.

[111] Hasegawa H, Kiyokawa E, Tanaka S, Nagashima K, Gotoh N, Shibuya M, Kurata T, Matsuda M. DOCK180, a major CRK-binding protein, alters cell

morphology upon translocation to the cell membrane. Mol Cell Biol. 1996 Apr;16(4):1770-6.

[112] Matsuda M, Ota S, Tanimura R, Nakamura H, Matuoka K, Takenawa T, Nagashima K, Kurata T. Interaction between the amino-terminal SH3 domain of CRK and its natural target proteins. J Biol Chem. 1996 Jun 14;271(24):14468-72.

[113] Kiyokawa E, Hashimoto Y, Kurata T, Sugimura H, Matsuda M. Evidence that DOCK180 up-regulates signals from the CrkII-p130(Cas) complex. J Biol Chem. 1998 Sep 18;273(38):24479-84.

[114]Tosello-Trampont AC, Kinchen JM, Brugnera E, Haney LB, Hengartner MO, Ravichandran KS: Identification of two signaling submodules within the CrkII/ELMO/Dock180 pathway regulating engulfment of apoptotic cells. Cell Death Differ 2007, 14:963-972.

[115] Nishihara H, Maeda M, Oda A, Tsuda M, Sawa H, Nagashima K, Tanaka S: DOCK2 associates with CrkL and regulates Rac1 in human leukemia cell lines. Blood 2002, 100:3968-3974.

[116] Wang, J. Y. & Baltimore, D. Cellular RNA homologous to the Abelson marine leukemia virus transforming gene: expression and relationship to the viral sequence. Mol. Cell. Biol. 3, 773–779 (1983).

[117] Hanks, S. K. Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. Genome Biol. 4, 111 (2003).

[118] Hantschel O, Superti-Furga G. Regulation of the c-Abl and Bcr-Abl tyrosine kinases. Nat Rev Mol Cell Biol. 2004 Jan;5(1):33-44.

[119] Hantschel O, Nagar B, Guettler S, Kretzschmar J, Dorey K, Kuriyan J, Superti-Furga G. A myristoyl/phosphotyrosine switch regulates c-Abl. Cell. 2003 Mar 21;112(6):845-57.

[120] Feller, S. M., Knudsen, B. & Hanafusa, H. c-Abl kinase regulates the protein binding activity of c-Crk. EMBO J. 13, 2341–2351 (1994).

[121] Smith, J. M., Katz, S. & Mayer, B. J. Activation of the abl tyrosine kinase in vivo by src homology 3 domains from the src homology 2/Src homology 3 adaptor Nck. J. Biol. Chem. 274, 27956–27962 (1999).

[122] Harrison, S. C. Variation on an Src-like theme. Cell 112, 737–740 (2003).

[123] Courtneidge, S. A. Cancer: escape from inhibition. Nature 422, 827–828 (2003)

[124] Harrison SC. Variation on an Src-like theme. Cell. 2003 Mar 21;112(6):737-40.

[125] Colicelli J. ABL tyrosine kinases: evolution of function, regulation, andspecificity. Sci Signal. 2010 Sep 14;3(139):re6. Review. Erratum in: Sci Signal. 2011 Aug 30;4(188):er4.

[126] S. Antoku, K. Saksela, G. M. Rivera, B. J. Mayer, A crucial role in cell spreading for the interaction of Abl PxxP motifs with Crk and Nck adaptors. *J. Cell Sci.* 121, 3071–3082 (2008).

[127] Shishido T, Akagi T, Chalmers A, Maeda M, Terada T, Georgescu MM, Hanafusa H: Crk family adaptor proteins trans-activate c-Abl kinase. Genes Cells 2001, 6:431-440.

[128] K. Q. Tanis, D. Veach, H. S. Duewel, W. G. Bornmann, A. J. Koleske, Two distinct phosphorylation pathways have additive effects on Abl family kinase activation. *Mol. Cell. Biol.* 23, 3884–3896 (2003).

[129] Brasher BB, Van Etten RA. c-Abl has high intrinsic tyrosine kinase activity that is stimulated by mutation of the Src homology 3 domain and by autophosphorylation at two distinct regulatory tyrosines. J Biol Chem. 2000 Nov 10;275(45):35631-7.

[130] Reichman C, Singh K, Liu Y, Singh S, Li H, Fajardo JE, Fiser A, Birge RB: Transactivation of Abl by the Crk II adapter protein requires a PNAY sequence in the Crk C-terminal SH3 domain. Oncogene 2005, 24:8187-8199.

[131] Ben-Neriah Y, Daley GQ, Mes-Masson AM, Witte ON, Baltimore D. The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. Science. 1986 Jul 11;233(4760):212-4.

[132] Sawyers CL. Disabling Abl-perspectives on Abl kinase regulation and cancer therapeutics. Cancer Cell. 2002 Feb;1(1):13-5.

[133] Oda T, Heaney C, Hagopian JR, Okuda K, Griffin JD, Druker BJ. Crkl is the major tyrosine-phosphorylated protein in neutrophils from patients with chronic myelogenous leukemia. J Biol Chem. 1994 Sep 16;269(37):22925-8.

[134] ten Hoeve J, Arlinghaus RB, Guo JQ, Heisterkamp N, Groffen J. Tyrosine phosphorylation of CRKL in Philadelphia+ leukemia. Blood. 1994 Sep 15;84(6):1731-6.

[135] Lucas CM, Harris RJ, Giannoudis A, Knight K, Watmough SJ, Clark RE. BCR-ABL1 tyrosine kinase activity at diagnosis, as determined via the pCrkL/CrkL

ratio, is predictive of clinical outcome in chronic myeloid leukemia. Br J Haematol. 2010 May;149(3):458-60.

[136] Harkiolaki M, Gilbert RJ, Jones EY, Feller SM. The C-terminal SH3 domain of CRKL as a dynamic dimerization module transiently exposing a nuclear export signal. Structure. 2006 Dec;14(12):1741-53.

[137] Sarkar P, Reichman C, Saleh T, Birge RB, Kalodimos CG. Proline cis-trans isomerization controls autoinhibition of a signaling protein. Mol Cell. 2007 Feb 9;25(3):413-26.

[138] Sarkar P, Saleh T, Tzeng SR, Birge RB, Kalodimos CG. Structural basis for regulation of the Crk signaling protein by a proline switch. Nat Chem Biol. 2011 Jan;7(1):51-7.

[139] Kobashigawa Y, Sakai M, Naito M, Yokochi M, Kumeta H, Makino Y, Ogura K, Tanaka S, Inagaki F. Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. Nat Struct Mol Biol. 2007 Jun;14(6):503-10.

[140] Sriram G, Reichman C, Tunceroglu A, Kaushal N, Saleh T, Machida K, Mayer B, Ge Q, Li J, Hornbeck P, Kalodimos CG, Birge RB. Phosphorylation of Crk on tyrosine 251 in the RT loop of the SH3C domain promotes Abl kinase transactivation. Oncogene. 2011 Nov 17;30(46):4645-55. doi: 10.1038/onc.2011.170.

[141] Rosen MK, Yamazaki T, Gish GD, Kay CM, Pawson T, Kay LE. Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. Nature. 1995 Mar 30;374(6521):477-9.

[142] Park TJ, Curran T. Crk and Crk-like play essential overlapping roles downstream of disabled-1 in the Reelin pathway. J Neurosci. 2008 Dec 10;28(50):13551-62.

[143] Matsuki T, Pramatarova A, Howell BW. Reduction of Crk and CrkL expression blocks reelin-induced dendritogenesis. J Cell Sci. 2008 Jun 1;121(Pt 11):1869-75. Epub 2008 May 13.

[144] Peterson AC, Marks RE, Fields PE, Imamoto A, Gajewski TF. T cell development and function in CrkL-deficient mice. Eur J Immunol. 2003 Oct;33(10):2687-95.

[145] Park TJ, Boyd K, Curran T. Cardiovascular and craniofacial defects in Crk-null mice. Mol Cell Biol. 2006 Aug;26(16):6272-82.

[146] Guris DL, Fantes J, Tara D, Druker BJ, Imamoto A. Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome.Nat Genet. 2001 Mar;27(3):293-8.

[147] Yanagi H, Wang L, Nishihara H, Kimura T, Tanino M, Yanagi T, Fukuda S, Tanaka S. CRKL plays a pivotal role in tumorigenesis of head and neck squamous cell carcinoma through the regulation of cell adhesion. Biochem Biophys Res Commun.2012 Feb 3;418(1):104-9.

[148] Miller CT, Chen G, Gharib TG, Wang H, Thomas DG, Misek DE, Giordano TJ, Yee J,Orringer MB, Hanash SM, Beer DG. Increased C-CRK proto-oncogene expression is associated with an aggressive phenotype in lung adenocarcinomas. Oncogene. 2003 Sep 11;22(39):7950-7.

[149] Kim YH, Kwei KA, Girard L, Salari K, Kao J, Pacyna-Gengelbach M, Wang P, Hernandez-Boussard T, Gazdar AF, Petersen I, Minna JD, Pollack JR. Genomic and functional analysis identifies CRKL as an oncogene amplified in lung cancer. Oncogene. 2010 Mar 11;29(10):1421-30.

[150] Rodrigues SP, Fathers KE, Chan G, Zuo D, Halwani F, Meterissian S, Park M. CrkI and CrkII function as key signaling integrators for migration and invasion of cancer cells. Mol Cancer Res. 2005 Apr;3(4):183-94.

[151] Wang L, Tabu K, Kimura T, Tsuda M, Linghu H, Tanino M, Kaneko S, Nishihara H, Tanaka S. Signaling adaptor protein Crk is indispensable for malignant feature of glioblastoma cell line KMG4. Biochem Biophys Res Commun. 2007 Nov 3;362(4):976-81.

[152] Birchmeier W, Behrens J, Weidner KM, Hülsken J, Birchmeier C. Epithelial differentiation and the control of metastasis in carcinomas. Curr Top Microbiol Immunol. 1996;213 (Pt 2):117-35.

[153] Linghu H, Tsuda M, Makino Y, Sakai M, Watanabe T, Ichihara S, Sawa H, Nagashima K, Mochizuki N, Tanaka S. Involvement of adaptor protein Crk in malignant feature of human ovarian cancer cell line MCAS. Oncogene. 2006 Jun 15;25(25):3547-56.

[154] Oda T, Heaney C, Hagopian JR, Okuda K, Griffin JD, Druker BJ. CrkI is the major tyrosine-phosphorylated protein in neutrophils from patients with chronic myelogenous leukemia. J Biol Chem. 1994 Sep 16;269(37):22925-8.

[155] Weidow CL, Black DS, Bliska JB, Bouton AH. CAS/Crk signalling mediates uptake of Yersinia into human epithelial cells. Cell Microbiol. 2000 Dec;2(6):549-60.

[156] Burton EA, Plattner R, Pendergast AM. Abl tyrosine kinases are required for infection by Shigella flexneri. EMBO J. 2003 Oct 15;22(20):5471-9. Erratum in: EMBO J. 2003 Nov 3;22(21):5962.

[157] Pielage JF, Powell KR, Kalman D, Engel JN. RNAi screen reveals an Abl kinase-dependent host cell pathway involved in Pseudomonas aeruginosa internalization. PLoS Pathog. 2008 Mar 21;4(3):e1000031.

[158] Suzuki M, Mimuro H, Suzuki T, Park M, Yamamoto T, Sasakawa C. Interaction of CagA with Crk plays an important role in Helicobacter pylori-induced loss of gastric epithelial cell adhesion. J Exp Med. 2005 Nov 7;202(9):1235-47.

[159] Yee AA, Savchenko A, Ignachenko A, Lukin J, Xu X, Skarina T, Evdokimova E, Liu CS, Semesi A, Guido V, Edwards AM, Arrowsmith CH. NMR and X-ray crystallography,complementary tools in structural proteomics of small proteins. J Am Chem Soc.2005 Nov 30;127(47):16512-7.

[160] Ferentz AE, Wagner G. NMR spectroscopy: a multifaceted approach to macromolecular structure. Q Rev Biophys. 2000 Feb;33(1):29-65.

[161] Kay LE. NMR studies of protein structure and dynamics. J Magn Reson. 2005 Apr;173(2):193-207.

[162] McTigue MA, Wickersham JA, Pinko C, Showalter RE, Parast CV, Tempczyk-Russell A, Gehring MR, Mroczkowski B, Kan CC, Villafranca JE, Appelt K. Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: a key enzyme in angiogenesis. Structure. 1999 Mar 15;7(3):319-30.

[163] Dyson HJ, Wright PE. Equilibrium NMR studies of unfolded and partially folded proteins. Nat Struct Biol. 1998 Jul;5 Suppl:499-503.

[164] Williamson MP, Havel TF, Wüthrich K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry. J Mol Biol. 1985 Mar 20;182(2):295-315.

[165] Sprangers R, Velyvis A, Kay LE. Solution NMR of supramolecular complexes: providing new insights into function. Nat Methods. 2007 Sep;4(9):697-703.

[166] Gordon S. Rule, T.K.H., *Fundamentals of Protein NMR Spectroscopy*. 1st ed, ed. R. Kaptein. Vol. 5. 2006, Dordrecht: Springer. 530.

[167] Pellecchia M, Sem DS, Wüthrich K. NMR in drug discovery. Nat Rev Drug Discov. 2002 Mar;1(3):211-9

[168] Sattler, M., J. Schleucher, and C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progress in Nuclear Magnetic Resonance Spectroscopy, 1999. 34(2): p. 93-158.

[169] Lin Y, Wagner G. Efficient side-chain and backbone assignment in large proteins: application to tGCN5. J Biomol NMR. 1999 Nov;15(3):227-39.

170 An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins Gaetano T. Montelione, Barbara A. Lyons, S. Donald Emerson, and Mitsuru Tashiro Journal of the American Chemical Society 1992 114 (27), 10974-10975

[171] Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in $^{13}C/^{15}N$-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization. S. Grzesiek, J. Anglister, A. Bax

[172] Kanelis V, Forman-Kay JD, Kay LE. Multidimensional NMR methods for protein structure determination. IUBMB Life. 2001 Dec;52(6):291-302.

[173] H. P. Mo and T. C. Pochapsky, Intermolecular interactions characterized by nuclear Overhauser effects. Prog. Nucl. Magn. Reson. Spectrosc., 1997, 30, 1.

[174] Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM.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 Jul 25;28(15):6150-6.

[175] Zuiderweg ER, Fesik SW. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. Biochemistry. 1989 Mar 21;28(6):2387-91.

[176] Yi Q, Scalley-Kim ML, Alm EJ, Baker D. NMR characterization of residual structure in the denatured state of protein L. J Mol Biol. 2000 Jun 23;299(5):1341-51.

[177] Wider G, Wüthrich K. NMR spectroscopy of large molecules and multimolecular assemblies in solution. Curr Opin Struct Biol. 1999 Oct;9(5):594-601.

[178] Sattler, M., Introduction to biomolecular NMR spectroscopy. EMBL Heidelberg, 2004.

[179] Gardner KH, Kay LE. The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct.1998;27:357-406.

[180] Rajesh S, Nietlispach D, Nakayama H, Takio K, Laue ED, Shibata T, Ito Y. A novel method for the biosynthesis of deuterated proteins with selective protonation at the aromatic rings of Phe, Tyr and Trp. J Biomol NMR. 2003 Sep;27(1):81-6.

[181] Isaacson RL, Simpson PJ, Liu M, Cota E, Zhang X, Freemont P, Matthews S. A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. J Am Chem Soc. 2007 Dec 19;129(50):15428-9. Epub 2007 Nov 28.

[182] Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc. 2006;1(2):749-54.

[183] Kay LE, Gardner KH. Solution NMR spectroscopy beyond 25 kDa. Curr Opin Struct Biol. 1997 Oct;7(5):722-31.

[184] Battiste JL, Wagner G. 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 May 9;39(18):5355-65.

[185] Clore GM, Tang C, Iwahara J. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. Curr Opin Struct Biol. 2007 Oct;17(5):603-16. Epub 2007 Oct 29.

[186] Kosen PA. Spin labeling of proteins. Methods Enzymol. 1989;177:86-121.

[187] Bertoncini CW, Jung YS, Fernandez CO, Hoyer W, Griesinger C, Jovin TM, Zweckstetter M. Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. Proc Natl Acad Sci U S A. 2005 Feb 1;102(5):1430-5.

[188] Fawzi NL, Doucleff M, Suh JY, Clore GM. Mechanistic details of a protein-protein association pathway revealed by paramagnetic relaxation enhancement titration measurements. Proc Natl Acad Sci U S A. 2010 Jan 26;107(4):1379-84.

[189] Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S. NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. Science. 2005 Feb 25;307(5713):1317-21.

[190] Iwahara J, Clore GM. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. Nature. 2006 Apr 27;440(7088):1227-30.

[191] Tjandra N, Bax A. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science. 1997 Nov 7;278(5340):1111-4.

[192] MacDonald D, Lu P. Residual dipolar couplings in nucleic acid structure determination. Curr Opin Struct Biol. 2002 Jun;12(3):337-43.

[193] Jain NU. Use of residual dipolar couplings in structural analysis of protein-ligand complexes by solution NMR spectroscopy. Methods Mol Biol. 2009;544:231-52.

[194] Salmon L, Bouvignies G, Markwick P, Lakomek N, Showalter S, Li DW, Walter K,Griesinger C, Brüschweiler R, Blackledge M. Protein conformational flexibility from structure-free analysis of NMR dipolar couplings: quantitative and absolute determination of backbone motion in ubiquitin. Angew Chem Int Ed Engl. 2009;48(23):4154-7.

[195] Ottiger M, Bax A. Characterization of magnetically oriented phospholipids micelles for measurement of dipolar couplings in macromolecules. J Biomol NMR. 1998 Oct;12(3):361-72.

[196] Hansen MR, Mueller L, Pardi A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat Struct Biol. 1998 Dec;5(12):1065-74.

[197] Markus Rückert and Gottfried Otting Alignment of Biological Macromolecules in Novel Nonionic Liquid Crystalline Media for NMR Experiments J*. Am. Chem. Soc.*, 2000, 122 (32), pp 7793–7797

[198] Ottiger M, Delaglio F, Bax A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson. 1998 Apr;131(2):373-8.

[199] Clore GM, Gronenborn AM, Bax A. A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. J Magn Reson. 1998 Jul;133(1):216-21.

[200] Lipsitz RS, Tjandra N. Residual dipolar couplings in NMR structure analysis. Annu Rev Biophys Biomol Struct. 2004;33:387-413.

[201] Henzler-Wildman K, Kern D. Dynamic personalities of proteins. Nature. 2007 Dec 13;450(7172):964-72.

[202] Mittermaier A, Kay LE. New tools provide new insights in NMR studies of protein dynamics. Science. 2006 Apr 14;312(5771):224-8.

[203] Ishima R, Torchia DA. Protein dynamics from NMR. Nat Struct Biol. 2000 Sep;7(9):740-3.

[204] Akke M. NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis. Curr Opin Struct Biol. 2002 Oct;12(5):642-7.

[205] P. R. Selvin The Renaissance of fluorescence resonance energy transfer *Nature Structural Biology*, 2000, *7*, 730-734.

[206] Philippe I.H. Bastiaens and Rainer Pepperkok; Observing Proteins in Their Natural Habitat: the Living Cell, *TIBS,* 2000, 25; 631-637

[207] Brian Herman, Victoria E. Centonze Frohlich, Joseph R. Lakowicz, Thomas J. Fellers and Michael W. Davidson; Fluorescence Resonance Energy Transfer (FRET) Microscopy, Molecular Expression^{TM}

[208] Joseph R. Lakowicz; Principles of Fluorescence Spectroscopy*, Springer*, 2006.

[209] Gerd Ulrich Nienhaus; Exploring Protein Structure and Dynamics under Denaturing Conditions by Single-Molecule FRET Analysis, *Macromolecular Bioscience*. 2006, 6: 907–922.

[210] Achillefs N. Kapanidisa and Shimon Weiss; Fluorescent Probes and Bioconjugation Chemistries for Single-Molecule Fluorescence Analysis of Biomolecules, *Journal of Chemical Physics*, 2002, 117: 10953-10964

[211] Elza V. Kuzmenkina, Colin D. Heyes and G. Ulrich Nienhaus; Single-molecule Förster Resonance Energy Transfer Study of Protein Dynamics under Denaturing Conditions, *PNAS*, 2005,102: 15471–15476

[212] Kim, Y.H. *et al.* Genomic and functional analysis identifies CRKL as an oncogene amplified in lung cancer. *Oncogene* 29, 1421–1430 (2010).

[213] 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. USA* 106, 2182–2187 (2009).

[214] Wang, H. *et al.* The role of Crk/Dock180/Rac1 pathway in the malignant behavior of human ovarian cancer cell SKOV3. *Tumour Biol.* 31, 59–67 (2010).

[215] Feng, R. *et al.* miR-126 functions as a tumour suppressor in human gastric cancer. *Cancer Lett.* 298, 50–63 (2010).

[216] Matsuda, M. *et al.* Two species of human CRK cDNA encode proteins with distinct biological activities. *Mol. Cell. Biol.* 12, 3482–3489 (1992).

[217] ten Hoeve, J., Morris, C., Heisterkamp, N. & Groffen, J. Isolation and chromosomal localization of CRKL, a human crk-like gene. *Oncogene* 8, 2469–2474 (1993).

[218] de Jong, R., ten Hoeve, J., Heisterkamp, N. & Groffen, J. Tyrosine 207 in CRKL is the BCR/ABL phosphorylation site. *Oncogene* 14, 507–513 (1997).

[219] Songyang, Z. *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell* 72, 767–778 (1993).

[220] 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* 3, 215–226 (1995).

[221] Cheerathodi, M. & Ballif, B.A. Identification of CrkL-SH3 binding proteins from embryonic murine brain: implications for reelin signaling during brain development. *J. Proteome Res.* 10, 4453–4462 (2011).

[222] Antoku, S. & Mayer, B.J. Distinct roles for Crk adaptor isoforms in actin reorganization induced by extracellular signals. *J. Cell Sci.* 122, 4228–4238 (2009).

[223] Senechal, K., Halpern, J. & Sawyers, C. The CRKL adaptor protein transforms fibroblasts and functions in transformation by the BCR-ABL oncogene. *J. Biol. Chem.* 271, 23255–23261 (1996).

[224] Sattler, M. & Salgia, R. Role of the adapter protein CRKL in signal transduction of normal hematopoietic and BCR/ABL-transformed cells. *Leukemia* 12, 637–644 (1998).

[225] Colicelli, J. ABL tyrosine kinases: evolution of function, regulation, and specificity. *Sci. Signal.* 3, re6 (2010).

[226] Seo, J.H. *et al.* A specific need for CRKL in p210BCR-ABL–induced transformation of mouse hematopoietic progenitors. *Cancer Res.* 70, 7325–7335 (2010).

[227] Luo, B. *et al.* Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci. USA* 105, 20380–20385 (2008).

[228] Nichols, G.L. *et al.* Identification of CRKL as the constitutively phosphorylated 39-kD tyrosine phosphoprotein in chronic myelogenous leukemia cells. *Blood* 84, 2912–2918 (1994).

[229] ten Hoeve, J., Arlinghaus, R.B., Guo, J.Q., Heisterkamp, N. & Groffen, J. Tyrosine phosphorylation of CRKL in Philadelphia+ leukemia. *Blood* 84, 1731–1736 (1994).

[230] Lucas, C.M. *et al.* BCR-ABL1 tyrosine kinase activity at diagnosis, as determined via the pCrkL/CrkL ratio, is predictive of clinical outcome in chronic myeloid leukaemia. *Br. J. Haematol.* 149, 458–460 (2010).

[231] Guris, D.L., Fantes, J., Tara, D., Druker, B.J. & Imamoto, A. Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome. *Nat. Genet.* 27, 293–298 (2001).

[232] Wang, J. *et al.* Crk and CrkL present with different expression and significance in epithelial ovarian carcinoma. *Mol. Carcinog.* 50, 506–515 (2011).

[233] Mihrshahi, R. & Brown, M.H. Downstream of tyrosine kinase 1 and 2 play opposing roles in CD200 receptor signaling. *J. Immunol.* 185, 7216–7222 (2010).

[234] . Donaldson, L.W., Gish, G., Pawson, T., Kay, L.E. & Forman-Kay, J.D. Structure of a regulatory complex involving the Abl SH3 domain, the Crk SH2 domain, and a Crk-derived phosphopeptide. *Proc. Natl. Acad. Sci. USA* 99, 14053–14058 (2002).

[235] Mittermaier, A.K. & Kay, L.E. Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.* 34, 601–611 (2009).

[236] Tjandra, N., Feller, S., Pastor, R. & Bax, A. Rotational diffusion anisotropy of human ubiquitin from N-15 NMR relaxation. *J. Am. Chem. Soc.* 117, 12562–12566 (1995).

[237] Seo, J.-H., Suenaga, A., Hatakeyama, M., Taiji, M. & Imamoto, A. Structural and Functional Basis of a Role for CRKL in a Fibroblast Growth Factor 8-Induced Feed-Forward Loop. *Mol Cell Biol* 29, 3076–3087 (2009).

[238] Senechal, K., Heaney, C., Druker, B. & Sawyers, C. Structural requirements for function of the Crkl adapter protein in fibroblasts and hematopoietic cells. *Mol. Cell. Biol.* 18, 5082–5090 (1998).

[239] Ren, R., Ye, Z.S. & Baltimore, D. Abl protein-tyrosine kinase selects the Crk adapter as a substrate using SH3-binding sites. *Genes Dev.* 8, 783–795 (1994).

[240] . Huang, X., Wu, D., Jin, H., Stupack, D. & Wang, J.Y.J. Induction of cell retraction by the combined actions of Abl-CrkII and Rho-ROCK1 signaling. *J. Cell Biol.* 183, 711–723 (2008).

[241] Kardinal, C. *et al.* Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients. *FASEB J.* 14, 1529–1538 (2000).

[242] Li, S.S. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem. J.* 390, 641–653 (2005).

[243] Severin, A., Joseph, R.E., Boyken, S., Fulton, D.B. & Andreotti, A.H. Proline isomerization preorganizes the Itk SH2 domain for binding to the Itk SH3 domain. *J. Mol. Biol.* 387, 726–743 (2009).

[244] Chan, B. *et al.* SAP couples Fyn to SLAM immune receptors. *Nat. Cell Biol.* 5, 155–160 (2003).

[245] Tang, C., Schwieters, C. & Clore, G. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* 449, 1078–1082 (2007).

[246] Ruckert, M. & Otting, G. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J Am Chem Soc* 122, 7793–7797 (2000).

[247] Yao L, Ying J, Bax A. Improved accuracy of 15N-1H scalar and residual dipolarcouplings from gradient-enhanced IPAP-HSQC experiments on protonated proteins. J Biomol NMR. 2009 Mar;43(3):161-70.

[248]Mandel AM, Akke M, Palmer AG 3rd. Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. J Mol Biol. 1995 Feb 10;246(1):144-63

[249] Lee J, An overview ofcyclophilins in human diseases. *J. Int. Med.Resech*2010. 38:1561-1574.

[250] Nicholson LK, Lu KP. Prolylcis-trans Isomerization as a molecular timer in Crk signaling.Mol Cell. 2007 Feb 23;25(4):483-5.

[251] Nelson CJ, Santos-Rosa H, Kouzarides T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell. 2006 Sep 8;126(5):905-16.

[252]Brazin KN, Mallis RJ, Fulton DB, Andreotti AH. Regulation of the tyrosine kinaseItk by the peptidyl-prolylisomerasecyclophilin A. ProcNatlAcadSci U S A. 2002 Feb 19;99(4):1899-904.

[253]OuYang B, Pochapsky SS, Dang M, Pochapsky TC. A functional proline switch in cytochrome P450cam. Structure. 2008 Jun;16(6):916-23.

[254]Franke EK, Yuan HE, Luban J. Specific incorporation of cyclophilin A into HIV-1 virions.Nature. 1994 Nov 24;372(6504):359-62.

[255] Wang P, Heitman J. The cyclophilins. Genome Biol. 2005;6(7):226.

[256]Gollan PJ, Ziemann M, Bhave M. PPIase activities and interaction partners of FK506-binding proteins in the wheat thylakoid.Physiol Plant. 2011 Dec;143(4):385-95.

[257]Rulten S, Thorpe J, Kay J. Identification of eukaryotic parvulin homologues: a new subfamily of peptidylprolylcis-trans isomerases. Biochem Biophys Res Commun. 1999 Jun 16;259(3):557-62.

[258]Janssens V, Goris J. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. Biochem J. 2001 Feb 1;353(Pt 3):417-39.

[259]Galat A. Peptidylprolylcis/trans isomerases (immunophilins): biological diversity--targets--functions. Curr Top Med Chem. 2003;3(12):1315-47.

[260]Handschumacher RE, Harding MW, Rice J, Drugge RJ, Speicher DW. Cyclophilin: a specific cytosolic binding protein for cyclosporin A. Science. 1984 Nov 2;226(4674):544-7.

[261]Obchoei S, Wongkhan S, Wongkham C, Li M, Yao Q, Chen C. Cyclophilin A: potential functions and therapeutic target for human cancer. Med SciMonit. 2009 Nov;15(11):RA221-32.

[262]Fanghänel J, Fischer G. Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. Front Biosci. 2004 Sep 1;9:3453-78.

[263]Göthel SF, Marahiel MA. Peptidyl-prolylcis-trans isomerases, a superfamily ofubiquitous folding catalysts. Cell Mol Life Sci. 1999 Mar;55(3):423-36.

[264] Towers GJ, Hatziioannou T, Cowan S, Goff SP, Luban J, Bieniasz PD. Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. Nat Med. 2003 Sep;9(9):1138-43.

[265]Wohlfarth C, Efferth T. Natural products as promising drug candidates for the treatment of hepatitis B and C. ActaPharmacol Sin. 2009 Jan;30(1):25-30.

[266] Satoh K, Shimokawa H, Berk BC. Cyclophilin A: promising new target in cardiovascular therapy. Circ J. 2010 Nov;74(11):2249-56.

[267] Howard BA, Zheng Z, Campa MJ, Wang MZ, Sharma A, Haura E, Herndon JE 2nd, Fitzgerald MC, Bepler G, Patz EF Jr. 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 Dec;46(3):313-23.

[268] Chen S, Zhang M, Ma H, Saiyin H, Shen S, Xi J, Wan B, Yu L. Oligo-microarray analysis reveals the role of cyclophilin A in drug resistance. Cancer ChemotherPharmacol. 2008 Mar;61(3):459-69.

[269]Mikuriya K, Kuramitsu Y, Ryozawa S, Fujimoto M, Mori S, Oka M, Hamano K, Okita K, Sakaida I, Nakamura K. 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 Apr;30(4):849-55.

[270] Sun S, Wang Q, Giang A, Cheng C, Soo C, Wang CY, Liau LM, Chiu R. Knockdown of CypA inhibits interleukin-8 (IL-8) and IL-8-mediated proliferation and tumor growth of glioblastoma cells through down-regulated NF-κB. J Neurooncol. 2011 Jan;101(1):1-14.

[271] Bane FT, Bannon JH, Pennington SR, Campiani G, Williams DC, Zisterer DM, Mc Gee MM. The microtubule-targeting agents, PBOX-6 [pyrrolobenzoxazepine 7-[(dimethylcarbamoyl)oxy]-6-(2-naphthyl)pyrrolo-[2,1-d] (1,5)-benzoxazepine] and paclitaxel, induce nucleocytoplasmic redistribution of the peptidyl-prolylisomerases, cyclophilin A and pin1, in malignant hematopoietic cells. J PharmacolExpTher. 2009 Apr;329(1):38-47.

[272]Pushkarsky T, Zybarth G, Dubrovsky L, Yurchenko V, Tang H, Guo H, Toole B, Sherry B, Bukrinsky M. CD147 facilitates HIV-1 infection by interacting with virus-associated cyclophilin A. ProcNatlAcadSci U S A. 2001 May 22;98(11):6360-5.

[273]Damsker JM, Bukrinsky MI, Constant SL. Preferential chemotaxis of activated human CD4+ T cells by extracellular cyclophilin A. J Leukoc Biol. 2007 Sep;82(3):613-8.

[274]Hur S, Bruice TC. The mechanism of cis-trans isomerization of prolyl peptides by cyclophilin. J Am Chem Soc. 2002 Jun 26;124(25):7303-13.

[275] Li G, Cui Q. What is so special about Arg 55 in the catalysis of cyclophilin A? insights from hybrid QM/MM simulations. J Am Chem Soc. 2003 Dec 10;125(49):15028-38.

[276]Zydowsky LD, Etzkorn FA, Chang HY, Ferguson SB, Stolz LA, Ho SI, Walsh CT. Active site mutants of human cyclophilin A separate peptidyl-prolylisomerase activity from cyclosporin A binding and calcineurin inhibition. Protein Sci. 1992 Sep;1(9):1092-9.

[277]Vajdos FF, Yoo S, Houseweart M, Sundquist WI, Hill CP. Crystal structure of cyclophilin A complexed with a binding site peptide from the HIV-1 capsid protein. Protein Sci. 1997 Nov;6(11):2297-307.

[278]Piotukh K, Gu W, Kofler M, Labudde D, Helms V, Freund C. Cyclophilin A binds to linear peptide motifs containing a consensus that is present in many human proteins. J Biol Chem. 2005 Jun 24;280(25):23668-74.

[279] Kern D, Eisenmesser EZ, Wolf-Watz M. Enzyme dynamics during catalysis measured by NMR spectroscopy. Methods Enzymol. 2005;394:507-24.

[280] Howard BR, Vajdos FF, Li S, Sundquist WI, Hill CP. Structural insights into the catalytic mechanism of cyclophilin A. Nat Struct Biol. 2003 Jun;10(6):475-81.

Curriculum Vitae

Wojciech Jankowski

Education

Doctor of Philosophy, Chemistry
2006-2012 Rutgers University, New Brunswick, NJ
Department of Chemistry & Chemical Biology

Master of Science, Biotechnology
1999.2005 Warsaw University of Technology, Warsaw, Poland
Faculty of Chemical and Process Engineering

Research Experience

2006-2012 **Graduate Research Assistant**, Department of Chemistry & Chemical Biology Rutgers University, New Brunswick, NJ
Advisor: Prof. Charalampos G. Kalodimos

2003-2005 **Research Staff**, Laboratory of Virology
National Institute of Public Health, Warsaw, Poland
Advisor: Andrzej Piasek, PhD

Publications

1. 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.