Studies on protein-protein interaction and its applications in ALIX/SFKs-SH3 complexes

Xiaoli Shi

Director: **Prof. Xavier Morelli**; Institut de Microbiologie de la Méditérannée (IMM), CNRS, France

Prof. Mingdong Huang; Division of Chemical Biology

Fujian Institute of Research on the Structure of Matter

(FJIRSM), Chinese Academy of Sciences(CAS)

Major: Molecular biology and Biochemistry

Degree: Doctor of Philosophy

Date of Defense: Feb 12th, 2011

Abstract

Src homology (SH) 3 domains is one of the most wide-spreaded protein modules found in nature. They mediate both inter- and intra-molecular protein-protein interactions (PPIs) through the formation and disassociation of multi-protein complexes. These SH3-mediated interactions are responsible for signal transduction, cytoskeleton organization and other cellular processes. The *nef* gene of Human immunodeficiency virus (HIV-1) encodes the HIV-1 Nef protein, which is important for optimal virus replication and development of AIDS (acquired immunize deficiency syndrome) in HIV-1 infected persons. Previous studies show that the HIV-1 Nef protein uses a "tertiary" binding mode to achieve high affinity and selectivity toward SH3 domains of Src-family kinases (SFKs). Two tertiary adjacent hot-spots on a SH3 binding surface seem to be important for the HIV-1 Nef to execute its cellular function that is required for AIDS progression. Whether this strategy of 'tertiary' binding mode of SH3 domains can be found in human cellular proteins, besides HIV-1 Nef, is an important question in the specificity of the HIV-1 Nef protein as an anti-HIV target. To search for and characterize novel downstream and/or regulatory binding partners, we carried out a screen by yeast-two hybrid using SH3 domains of SFKs as the bait. We identified Alix (ALG-2 [apoptosis-linked gene 2]-interacting protein X) as a novel protein interacting with Hemopoietic cell kinase (Hck) SH3 domain. Alix has similar selectivity towards SH3 domains of SFKs as the HIV-1 Nef. We have combined biophysical and structural biology analysis, including ITC (isothermal titration calorimetry), SPR (surface Plasmon resonance), GST (glutathione S-transferase) pull-down, interferometry, HSQC (heteronuclear single quantum coherence) and SAXS (small-angle X-ray scattering) to explore the characteristics of Alix-SH3 recognition mode. This study shows that Alix as a unique cellular protein, whcih is structurally different but functionally similar in recognizing HIV-1 Nef. The structural information of the Alix-Hck association facilitates the understanding of how Hck and Alix assist viral budding and cell surface receptor regulation.

Keyword: Alix, Hck, Src family kinases, SH3 domains, PPIs, HIV-1 Nef**,** RT-loop

Abbreviation

Src homology domain	SH
protein-protein interactions	PPIs
Acquired Imunume Deficiency Syndrome	AIDS
src-family nonreceptor kinases	SFKs
hemopoietic cell kinase	Hck
isothermal titration calorimetry	ITC
surface plasmon resonance	SPR
heteronuclear single quantum coherence	HSQC
small-angle X-ray scattering	SAXS
apoptosis-linked gene-2	ALG-2
ALG-2-interacting protein X	ALIX
the SH3 domain-containing expressed in tumorigenic astrocytes	SETA
charomatin-modifying protein; charged multivesicular body protein, CHMP4	CHMP4
tumor susceptibility gene101	Tsg101
vacuolar protein sorting	Vps
endosomal sorting complexes required for transport	ESCRTs
multivesicular body	MVB

proline rich region	PRR
glutathione S-transferase	GST
Protein Data Bank	PDB
polyproline type II	PPII
small ubiquitin-related modifier	SUMO
solvent accessible surface area	ASA
epidermal growth factor	EGF
insulin and insulin-like growth factor-1	IGF-1
vascularendothelial growth factor	VEGF
phosphotyrosine binding	PTB
receptor protein tyrosine kinases	RTKs
non-receptor protein tyrosine kinases	nRTKs
focal association kinases	Fak

CONTENTS

Part I Introduction

1. Principles of protein-protein interactions (PPIs)

PPIs (protein-protein interactions) play key roles in cellular biological processes, such as the assembly of signaling protein complexes and large protein networks that regulate cellular behaviors including cell growth, differentiation, motility, polarity and apoptosis. Disorder or disruption of PPIs in normal cellular processes leads to diseases. Therefore, PPIs have became an important class of targets for developing cellular probes and therapeutic agents [1].

Many biological processes involve formation of various protein-protein complexes. According to Nooren and Thornton, these complexes can be divided into two groups with three different ways [2]. First, according to the nature of identical or non-identical subunits, they are classified into two types: homo and hetero-oligomers. Homo-oligomers have identical or homologous protein units. Such complexes generally associate by the same interface on each protomer, but in some case homo-dimers further oligomerize by using different interfaces, for example actin. In contrast to homo-oligomers, hetero-oligomers have non-identical protein subunits and can form large complexes by using different interfaces. Second, protein-protein complexes can be classified into obligate or non-obligate complexes whether their components exist independently or not. The promoters in an obligate complex could not exist independently and are often expressed simultaneously and thus co-localized upon synthesis (Figure 1). Non-obligate oligomers consist of stable monomers that exist independently but each subunit can initially reside in different compartments (e.g. the enzyme-inhibitor, antibody-antigen, and receptor-ligand) or be localized within a compartment (e.g. intracellular signaling complexes) (Figure 1). Previous studies have demonstrated that obligate PPIs are often homo-oligomers, and mostly non-obligate complexes are hetero-oligomers. However, some obligate-complexes are also hetero-oligomers, for example human cathepsin D. Furthermore homo-oligomers can also form non-obligate oligomers (e.g. sperm lysine). Finally, PPIs can also be classified based on the lifetime of the complexes: transient and permanent complexes.

Transient complexes can be sub-classified into weak or strong transient complexes, depending on the association intensity between subunit. Weak transient interactions are sensitive to physiologic conditions and environment such as pH variation or temperature, and present a dynamic equilibrium in solution. Strong transient associations require a molecular trigger (stronge exterior factors) to break the oligomeric equilibrium. For example, the Gα-Gβγ subunits are assembled into hetero-trimeric G protein, regulating by GTP/GDP exchange [2, 3]. In general, obligate complexes are permanent, whereas non-obligate oligomers may be transient (e.g. G proteins and mostly signaling pathway proteins) or permanent (e.g. thrombin in complex with rodniin inhibitor and antibody-antigen).

The formation of either non-obligate or transient complexes plays a key role in cellular regulation and is responsible for their corresponding biological functions. The control mechanism of PPIs mainly depends on the binding affinity (Kd) between their components and the binding energy (ΔG) contributing to the complex formation. For example, the complexes formed by antibody and antigen present a high affinity and ΔG, and the association is often permanent and irreversible. In contrast, weak transient complexes are regulated physicochemically and changed by the affinity between subunit, sometimes at an order of magnitude, such as the association and dissociation of G protein complexes. A Gα-Gβγ trimer bound with GDP has a 1000-fold higher affinity than Gα or Gβγ subunit upon GTP binding. In other cases, many proteins have a 100 uM affinity for ADP/ATP, but since ADP/ATP exists in so high concentrations in the cell, these complexes are always formed. In summary, it is important to distinguish between different types of complexes when analyzing the intermolecular interfaces that occur within them.

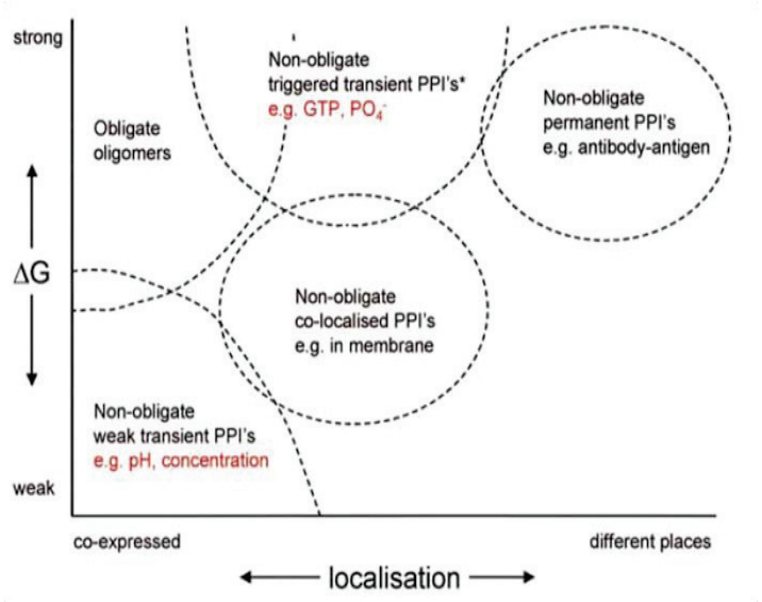


Figure 1. Relationship between different types of protein-protein interactions (PPIs), their binding energy (ΔG) and localization of their subunit. Triggers that control the transient oligomerization are given in red. * Large conformational changes are usually associated with these transient PPIs [2].

2. Analysis of protein-protein interfaces

Structural data are available for various proteins or protein complexes to explore the nature of the structural interfaces. The structure of interfaces can be analyzed by parameters such as the size of the contact area, propensity of amino acids to be present at the interface, the polarity of the interface, planarity to characterize a protein-protein interface, *et al* [2-4].

2.1 Size and shape of protein-protein interfaces

PPIs are primarily defined based on the changes of solvent accessible surface area (ΔASA). The interface residues were defined as those having ASAs that markedly decreased upon complexation [3]. Structural analysis shows that the interfaces in obligate complexes (e.g. homo-oligomers) are generally larger than those of non-obligate complexes (e.g. hetero-oligomers) [2, 3]. The average ΔASA of homo-dimer and hetero-dimers are 1462.1 $Å^2$ and 1334.6 $Å^2$ respectively on the ProtorP server, which are both larger than their corresponding standard deviation of

1302.9 $Å^2$ (homo-dimer) and 1221.4 $Å^2$ (hetero-dimers) [5]. It has also been reported that complexes with interfaces larger than ~ 1000 $Å^2$ are likely to undergo conformational changes upon complex formation [2]. Besides ΔASA, planarity is another parameter to distinguish protein-protein interfaces. Two proteins may interact and form a protein-protein interface with relatively flat surfaces or form a convoluted interface [3]. Planarity is used to measure how far the interface residues deviated from the plane by calculating the RMSD (root mean square deviation) of all the interface atoms from the least-squares plane through the atoms. Previous studies demonstrated the heterodimers have interfaces that are more planar than homodimers [2, 3]. The homodimers have higher mean RMSD [3]. Non-obligate weak transient PPIs show more planar interfaces than non-obligate strong transient PPIs. The latter resembles homo-dimers, and both of their interfaces in subunits are more closely packed but less planar [2]. Another parameter denoted as the "circularity" of the interface permits to define the shape of the protein binding sites [3]. The circularity is the ratio of the lengths of the principal axes of the least-squares plane through the atoms in the interface. A ratio near 1.0 indicates that the interface is approximately circular meaning that the atoms that form the buried interface are all within a circular plane. Usually, the interface of hetero-complexes are not perfectly circular but are rather oblong [6].

2.2 Characteristics of protein-protein interfaces

The physical natures of protein-protein interfaces are quite diverse. Of the involved characteristics, complementary and hydrophobicity play most important roles. In general, higher complementary means better binding affinity and tendency to form a more stable complex. Compared to the electrostatic complementary of interfaces between homodimers and heterodimers, permanent homodimers are most complementary and non-obligate heterodimers are the least complementary [1-3, 7]. It is often assumed that proteins associate through hydrophobic patches on their interfaces and their interior is more hydrophobic to stabilize globular protein structures. Most homodimers tend to be tightly bound and the interfaces are more

hydrophobic than the rest of the exterior regions. In contrast, the heterodimers are generally less strongly bound and the interfaces are more hydrophilic compared with the interfaces in homodimers, presumably reflecting that the monomeric proteins are normally independently folded in solution.

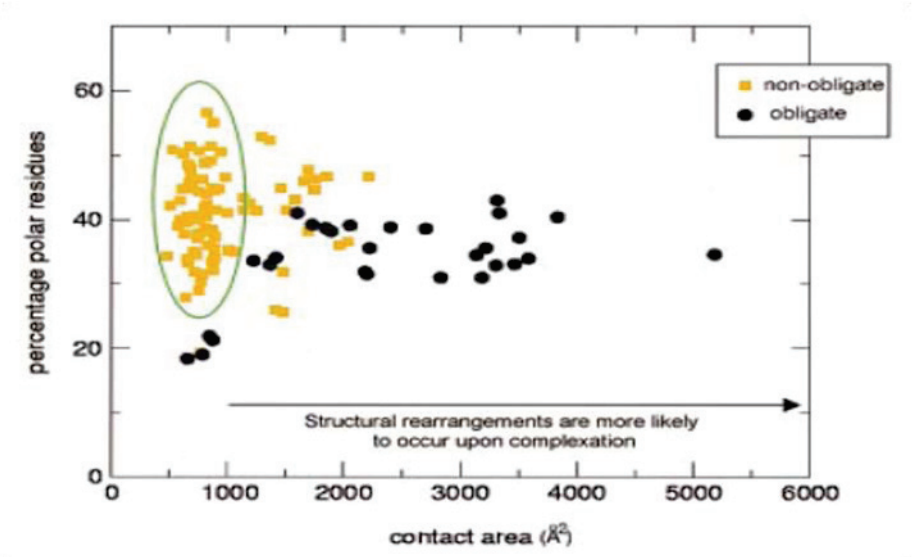


Figure 2. Contact area and polarity of the interfaces of various non-obligate and obligate complexes [2].

Statistic analysis of amino acids on the PPIs surfaces shows that, with the exception of methionine, hydrophobic residues show a greater preference at the interfaces of homo-oligomers than those of hetero-oligomers. The lower propensities for hydrophobic residues in interfaces of hetero-oligomers are balanced by an increased propensity for polar residues (Figure 2) [2, 3]. Based on a variety of alanine scanning experiments, it was suggested that protein-protein interfaces present one or more small regions called “hot spots”. Bogan and Thorn found that “hot spots” were comprised of relatively polar core surrounded by a ring of hydrophobic residues, which may act to exclude waters from the binding site and stabilize the polar interactions. Most binding energy is localized in this region [8]. Later, X-ray crystallography showed that “hot spots” are compact and centralized region on protein-protein interfaces, and contain important residues contributing to binding

energy for the complex formation [9]. Disruption of the interaction in “hot spots” from binding partners leads to break up of the protein complex. Therefore, “hot spots” are targeted by drug design. Alanine scanning mutagenesis is widely performed to explore key amino acid propensity in “hot spots”. Though there is no agreement with it yet, analysis data revealed that some amino acids are highly responsible for the binding energy in most cases. For instance, some researchers found that the three amino acids of Trp, Arg and Tyr are the most important residues on the “hot spot” and contribute most of the binding energy, and residues of Ile, Asp and His are suggested to contribute somewhat less [4, 10]. Another group predicted that Asn, Ser and Glu produce still significant contributions to binding energy. In addition, His, Ser and Glu are also found as the significant contributors for binding energy in “hot spots” [3, 4]. Therefore, this lack of a clear pattern increases the challenge of drug design to block PPIs.

Crystal structures elucidate that a “hot spot” do not cover the entire protein-protein binding surface, and that the residues involved in a “hot-spot” largely contribute to the total binding energy of the PPIs. Hot spots look highly complementary to each other, with buried charged residues forming salt bridges and hydrophobic residues from one surface fitting into small nooks on the opposite surface. Inhibitors of a size suitable for protein-protein interactions do not in principle need to cover the whole interfaces; it must just occupy the “hot-spot” [9]. So, identification of “hot spots” facilitates the design of inhibitors. However, functional and structural adaptivity is another hot spots feature that additionally increases the difficulty for drug design. That is to say, one protein will use the same conserved “hot spot” to associate with different binding partners. In this case, it is challenging to develop specific and potent inhibitors that only disrupt the targeted PPIs because of the shared “conserved” region in PPI interfaces. For example, viruses tend to mimic cellular protein-protein interactions and compete with the cellular proteins to elicit cellular reactions favorable for their host infectivity and replication in primary cells. Therefore, targeting such viral PPIs interface has a high likelihood to cross-react with, and to inhibit, cellular protein interactions as well. For PPIs inhibition to be successful,

it is thus critical to understand the molecular and energetic principles that govern interactions of pathogen proteins with cellular ones and to distinguish them from cellular PPIs [11].

3. Protein-protein interactions mediated by Src homology 3 domains

Src is an oncogene initially found in Rous sarcoma virus [12]. This protein tyrosine kinase contains a Src homology (SH) 2 and a SH3 domain in addition to its kinase domain. SH2 and SH3 domains are wide-spreaded protein modules identified in many cellular biological processes. SH2 domains contain about 100 amino acids, and recognize phosphotyrosine-containing ligands. For example, SH2 domains participate in cellular signaling pathways by binding to the autophosphotyrosine site in receptor tyrosine kinases (PDGR and EGFR). SH3 domains were initially discovered in the N-terminal non-catalytic part of the *src* tyrosine kinases [13]. Subsequently, SH3 domain was identified as a conserved sequence in many other important cellular proteins, present from yeast to human genomes [14, 15], including kinases, lipases, GTPases, adaptor proteins, structure proteins and others. In these proteins, SH3 domains act as modules mediating specific protein-protein interactions, and are thus involved in diverse cellular processes, such as cytoskeleton organization, signal transduction, cell cycle regulation, actin organization etc. [16, 17]. They mediate both inter- and intra-molecular protein-protein interactions by recognizing proline-rich containing ligands. SH3 domains are found in many signaling proteins, including cytoplasmic tyrosine kinases, the Crk adaptor protein, phospholipase C-γ and so on. Many proteins such as Src family kinases, contain both SH2 and SH3 domains. SH3 domains collaborate with SH2 domains to regulate biological processes. Despite the unrelenting attention to SH3 domain, there is still much left to uncover.

3.1 Topological characteristics of SH3 domains

The SH3 domain is one of the most prevalent protein-protein interaction modules found in nature. It can be recognized by versatile peptides and proteins. SH3 domains are small molecules comprising approximately 60 amino acid residues, which share a

significant sequence similarity between divergent signaling proteins and other cellular proteins. For example, there are 14 conserved amino acids, (Ala89, Leu90, Tyr91, Asp92, Tyr93, Asp100, Gly106, Glu107, Trp119, Trp120, Gly131, Tyr132, Pro134 and Tyr137), in Src family kinase (SFKs) SH3 domains (Figure 3). Crystallography and multi-dimensional nuclear magnetic resonance (NMR) showed SH3 domains share a common compact *β*–barrel structure made of five anti-parallel *β*–strands, which are arranged into two sheets packed at 90˚ angles (Figure 4) [18-21]. The first sheet is formed by *β*–strands a, e, and b1, while the second is formed by *β*-strands b2, c, and d. The NH2- and COOH-termini are located in adjacent strands at the end of first *β*–sheet, and their proximity allows the SH3 domain to exist as an independent and globular entity in solution. There are three loops in the SH3 domain structure, called the RT-loop, nSrc-loop and distal loop, respectively. The RT-loop is located between the *β*–strands a and b. The nSrc and distal loops are shorter and are located between *β*–strands b and c, and c and d, respectively. Besides the three loop components, there is a 3_{10} helix, situated between residues 133 to 137. Crystal and NMR structures demonstrate that a SH3 domain associates with various peptides or proteins using the same “hot spot” on its surface. This “hot spot” consists of conserved aromatic residues, flanked by the two loops of RT-Src loop and nSrc loop, to form a conserved patch on SH3 surface which contributes to the protein-ligand interaction [21] (Figure 5). The aromatic patch consists predominantly of tyrosine and tryptophan. For Src SH3, a conserved surface patch consists of the central four aromatic residues of Tyr91 + Tyr137 (top pair) and Trp119 + Tyr 132 (bottom pair) (numbered according to Fyn SH3) (Figure 5). Besides the conserved patch, there is a hydrophobic core which does not contribute to the molecular surface. In Fyn SH3, this core comprises the residues Ala89, Phe103, Phe109, Ala122, Ile133 and Val138. In contrast to the aromatic patch, the hydrophobic core is not conserved. Because the hydrophobic core involves close packing in the interior of the protein, all the groups of hydrophobic side chains are predominantly inaccessible to solvent. The hydrophobic core is critical for protein stability.

Analysis of sequence alignments between Src family kinase SH3 domains,

suggested there are both highly conserved and non-conserved amino acids in the RT-Src and n-Src loops. It have been demonstrated by mutagenesis and crystallography that all these residues contribute to specificity and affinity of ligand binding. For instance, the conserved amino acid of Asp100 on the RT loop makes strong ionic interactions with the specific amino acid of Arg or Lys flanking the core motif of PxxP (where P is a proline residue, and x is any amino acid). While the key amino acid 96 located in the RT loop is less conserved. Meanwhile, the RT loop flexibility induced by non-conserved Ile96 has been proved to enhance the specificity of SH3 domains of Src-family for its binding ligands, such as HIV-1 Nef protein [22].

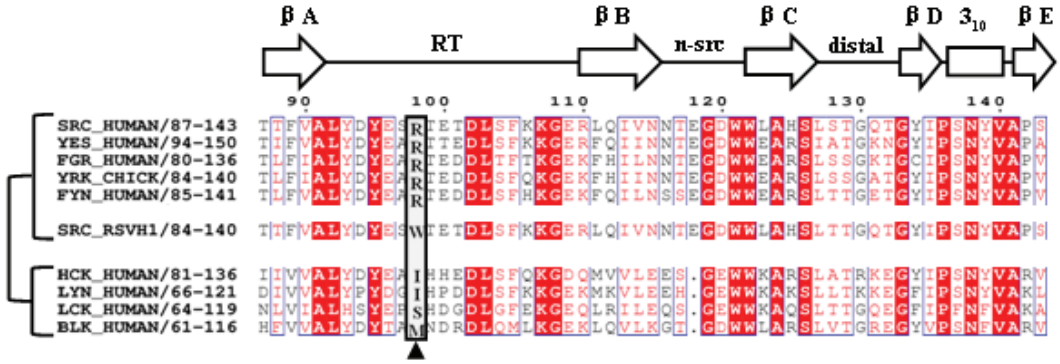


Figure 3.Sequence alignment of Src family SH3 domains from Hck, Fyn, Lck and Src with secondary structural elements common to Src family SH3 domain. The residues involved in hydrophobic pocket are highlighted in red

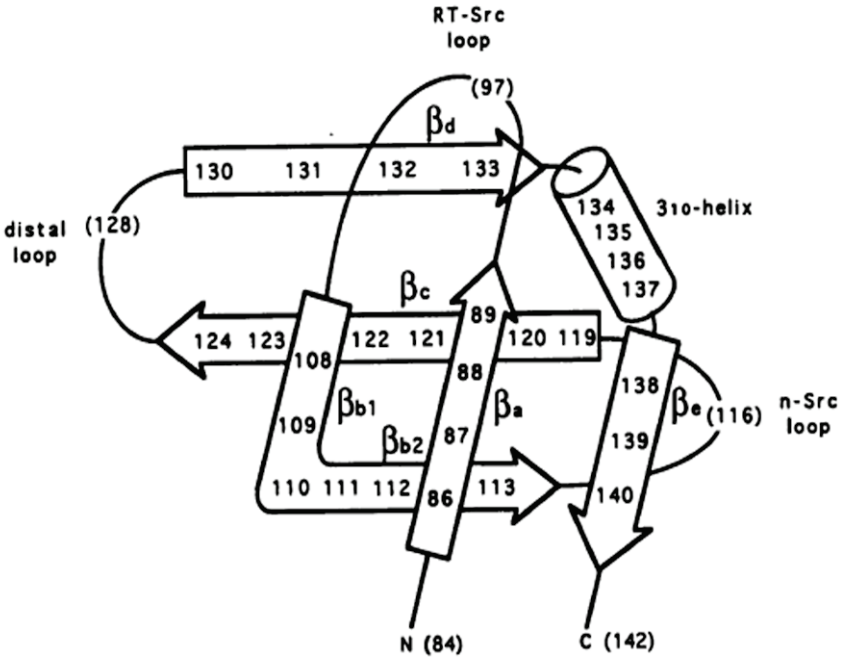


Figure 4.Topology diagram of Fyn SH3 [21]. The amino acid residues are numbered according to the sequence reported by Kawakami (1986) *et al* and Semba *et a l* (1986).

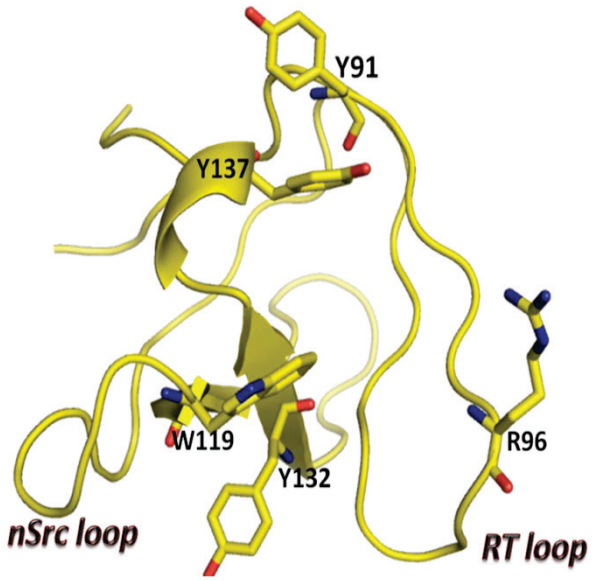


Figure 5. Ribbon diagram illustrating the structure of Fyn-SH3 domain

3.2 SH3 domains: versatile peptides and proteins recognition module

3.2.1 Canonical peptide motifs recognized by the SH3 domains

In 1992, Hongtao, Yu and his colleagues determined the solution structure of Src SH3 domain, and identified two polyproline-containing sequences derived from 3BP-1 bound to the conserved aromatic patch on the surface of Src-SH3 domains [20]. Shortly after, Ruibao Ren *et al* identified the PxxP motif as the core sequence recognized by SH3 domains [23]. Since then, structural analysis by crystallography, NMR and biochemical methods were utilized to shed light on the interactions between SH3 domains and their targets, establishing that the majority of Src SH3-binding proteins contains at least one conserved PxxP core motif [14, 21, 23-28], and that the binding surface on SH3 domains is located in the aromatic “hot spot”.

PxxP core binding sequences selected by SH3 domains adopt a pseudo-symmetrical and left-handed polyproline type II (PPII) helix conformation, which contains a two-fold rotation axis and binds to SH3 domain at one of the two opposite orientations. The PPII helix has three residues per turn, which forms a triangle and allows the base of this triangle sitting on the surface of SH3 domains. So the two essential proline residues contact directly the SH3 domain surface, whereas the other proline residues function as molecular scaffold to stabilize the PPII helix. Since polyproline sequences are widely distributed in distinct proteomes from prokaryotes to eukaryotes (approximately 25% human proteins harbor Proline-rich regions), it is mysterious that hundreds of SH3 domains select their physiological partners and involve in specific cellular processes without any cross-interaction. Therefore, many techniques, including chemical synthesized peptides, high resolution structure determination and phage display libraries screening, have been widely used in *vitro* in order to understand how specificity is achieved in the protein-protein interaction mediated by SH3 domains and their PxxP-containing ligands. The interactions mediated by PxxP core motifs with SH3 domains have dissociation constants in the 1-200 uM range [25, 26, 29] (Table 1), comprising hydrophobic interactions and the electrostatic interaction with basic residues flanking the core binding sequence and the acidic residues in SH3 domain [24-27, 30-32]. The PxxP

containing peptides are classified into two groups according to their orientations on the SH3 binding surface: class I and II [27, 31, 32]. The canonical PxxP binding sequence of class I is (R/K)xXPxXP and of class II is XPxXPx(R/K), where the capital 'X' stands for a non-glycine and hydrophobic residue, while the small 'x' denotes any naturally occurring amino acid [27, 30, 33] (Figure 6). Basic residues R/K preceding or following the PxxP core motif determinates the orientation of the peptide. Despite difference in binding orientation, all the peptides containing XPxXP core motifs adopt the same binding modes at SH3 domains. Crystal structures of SH3 domains in the complex with peptides reveal that the surface of SH3 domain is a relatively flat, conserved and hydrophobic ligand-binding surface, which consists of three shallow pockets defined by some conserved aromatic residues (Figure 7). One of the pockets determines the specificity with conserved Asp residues (D100 in Fyn) in SH3 domains, which is distal to the specific selectivity for SH3 binding partners by association with the terminal basic residue R/K flanking the PxxP core motif. The two other pockets consist of four (Y93, W119, P134, Y137) and two (Y91, Y137) conserved aromatic residues respectively and defined as XP binding pockets. The XP dipeptides occupy the two ligand-binding pockets through extensive hydrophobic interactions [16, 30]. The proline residue is the only naturally occurring N-substituted amino acid, which is very specific to the two hydrophobic pockets in the "hot spot" of SH3 domain. Mutagenesis studies demonstrate that the substitution of a Pro by an Ala completely abolished binding. In contrast, replacement of the same Pro residue by a sarcosine (N-substituted glycine) did not significantly alter the affinity of the peptide for the SH3 domain [17, 34]. Therefore, the three key amino acids of two Pro and one basic residue in the core binding motif are the major contributors to the binding affinity. However, considering naturally prevalent SH3 domains and PxxP-containing ligands, there should be extra recognition sites outside the aromatic conserved "hot spot" in a SH3 domain to specifically select their corresponding biological partners in various cellular processes and prevent cross-reactions. Several groups have used phage display or other combinatorial approaches to identify optimal ligands for particular SH3 domains. They found ligand

residues outside the core binding motif can also contribute to the specificity by interacting with surfaces on a SH3 domain outside the PPII binding patch [26, 27, 35, 36]. For example, the appearance of the serine residue, in the downstream of the core binding sequence (PxxPxRxx**S)** in the adaptor protein Nck-SH3 binding peptide, increases the binding affinity to Nck-SH3 domain. However, phosphoserine, acidic and proline residues negatively regulate binding [16, 37, 38]. Moreover, some residues in RT loop and nSrc loop have been proved to be vital for the selectivity and specificity of SH3 domain binding ligands [22, 39].

Table 1. Binding of proline-rich motifs to SH3 domains in 1-200uM range [40].

SH3 domain	Peptides	K_D (uM)
Abl	RAPTMPPPLPP (3BP-1)	34
Abl	PPAYPPPPVP (2BP-2)	5
Fyn	PVRPQVPLRPPMT (Nef)	202
Fyn	PPRPTPVAPGSSKT (p85)	50
Fyn	HSIAGPPVPPR (Sos1-4)	20
Fyn	RAPTMPPPPLPP (3BP-1)	34
Fyn	PPAYPPPPVP (2BP-2)	34
Fyn	PPRPLPVAPGSSKT (p85)	16
Grb2	VPPPVPPRRR (Sos)	5
Grb2	GTDEVPVPPPVPPRRRPESA (hSos)	21
Hck	PVRPQVPLRPPMT (Nef)	91
p85	RKLPPRPSK (libraries)	9
Src	RALPPLPRY (libraries)	8
Src	HSIAGPPVPPR (Sos1-4)	26

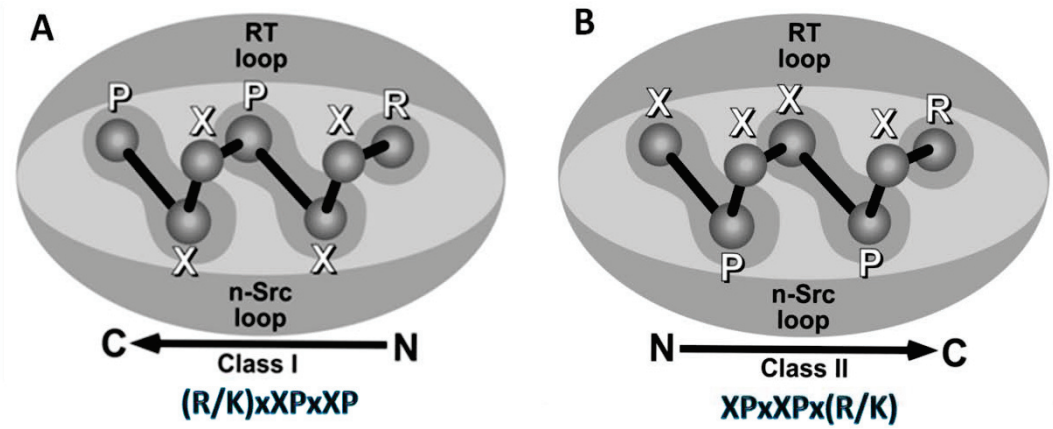


Figure 6.Diagram presented as association of class I (A) and II(B) ligands with the SH3 domain [16].

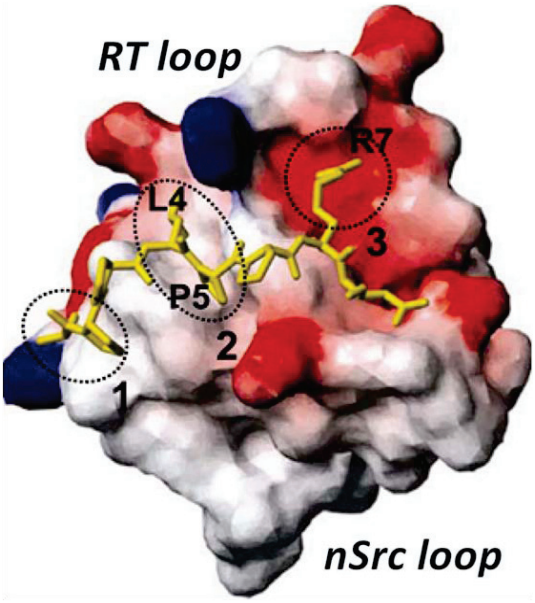


Figure 7.The structure of the c-Src SH3 domain in complex with a PxxP-containing peptide. Site 1 and 2 are the XP pockets and site 3 is the specificity pocked occupied by the basic amino acid Arginine. Red is positive charged; blue is negative charged; light grey is hydrophobic area [17, 27].

3.2.2 Non-canonical SH3-binding peptide motifs

As discussed above, in general SH3 domains associate with PxxP-containing sequences, and these ligands bind to the conserved ligand binding surface on SH3 domains, called “hot spot” comprising three aromatic pockets. Moreover, very subtle changes in the sequence of SH3 domains and the binding peptides determine relative specificity of peptide binding, besides the expanding binding sequence outside the PxxP core binding motif. It is reported that binding affinity between SH3 domains and PxxP-containing peptides is week or modest, spanning 1-200 uM, which is not enough to prevent cross-talking between SH3-containing proteins and PxxP-containing interacting partners in complicatedly cellular processes. Recently, it was found that there are some other atypical binding sequences involving in the recognition of SH3 domains [17, 33, 40] (Table 2). These non-canonical motifs

therefore increase the ligand span of the SH3 domain family.

The atypical SH3 binding motifs are classified into two groups according to whether or not their motifs contain conserved Pro residues. For instance, the proline-arginine motif of PxxxPR was shown to bind to the SH3 domains of the endocytic adaptor protein CIN85/CMS, which has been identified that CIN85 associated with containing the motif effectors to control intracellular trafficking of epidermal growth factor receptors (EGFR) [41]. Very recently, SH3 domain of tyrosine kinase substrate Eps8 was shown to bind to a PxxDY motif rather than to the normal PxxP motif. Another atypical motif is RxxPxxxP that is found in the cytoplasmic regions of the calcium activated potassium (BK) channels and mediates binding to the cortactin SH3 domain (reviewed in [16, 17, 33, 42]).

Besides the above atypical motifs containing the conserved Pro residues which were shown to be essential for binding, there are some other SH3-ligands lacking proline. The peptide region with an RKxxYxxY motif present on the adaptor protein SKAP55 (Src kinase associated protein of 55 kDa) was identified as the SH3 binding motif of Fyn and Fyn-binding protein, and it is demonstrated that up to now it is the only SH3 domain capable of binding to class I peptides that can also bind to the RKxxYxxY motif. Here, the Tyr in RKxxY motif plays a unique feature of the interaction. It has been shown that phosphorylation of the Tyr residue abolishes the interaction between RKxxY containing motif and its partner. The PxxDY motif also shares the same regulation mechanism. Another case is the interaction between yeast protein Pex13p-SH3 domain and its binding ligands of Pex5p and Pex14p. Pex14p contains a canonical PxxP motif and interacts with SH3 domain in a conventional manner, but Pex5p is devoid of a proline and consists of a α-helix formed by the WxxFxxLE motif. It was found that both peptides can bind simultaneously to Pex13p-SH3 domain to form a ternary complex and do not compete for the conventional PxxP binding groove (reviewed in [16, 17, 33, 42]). Of all the atypical SH3-binding motifs, the RxxK motif is found in a large number of proteins such as a family of proteins that are involved in cytokine-mediated signaling and receptor-mediated endocytosis and exocytosis, and it is perhaps appropriate to be

designated as the class III consensus recognized by SH3 domains. The affinity of some RxxK containing ligands is significantly greater than the normal affinity for a canonical SH3-binding ligand (reviewed [17, 33]). For example, the Gads SH3 domain associates with an RxxK-containing SLP-76 (SH2 domain-containing leukocyte protein of 76 kDa, T cell adaptor protein) 11-mer peptide (APSID**RSTK**PA) at unusually high affinity (Kd = 0.24 uM), and 14-mer peptide (APSID**RSTK**PPLDR) at a Kd of 8 nM. This is also another example demonstrating that a sequence extending outside canonical binding motifs can increase the binding affinity between SH3 domain and their binding partners [17, 33, 43, 44]. Crystal structure and NMR solution structure both elucidate the RSTK fragment adopts a right-handed 3_{10} helix. Solution structure of Gads-SH3 domain in complex with 11-mer peptide of (A***P***S***I***D***RSTK***PA) shows that the hydrophobic residues of Pro and Ile in the peptide occupy the two ligand-binding pockets denoted as "XP" binding pockets, and RSTK associates with the specificity pocket found in the conserved "hot spot" of SH3 domains (Figure 8). By comparison with "XP" binding pockets in conventional SH3-ligand interaction, the second-one groove in Gads SH3 domain is significantly different, which is deeper and narrower, and occupied by an aliphatic Ile residue instead of the canonical XP dipeptide unit in ligands. This ligand-binding groove is enclosed on one side by the side chain of Glu275 due to a conformational change in the RT loop. In addition, the side-chains of Arg and Lys in RSTK motif engage a pair of Glu residue in the RT loop of the SH3 domain [45].

Table2. Varieties of SH3 binding motifs [17]

Domain	Binding motifs	Affinity(Kd,uM)	Ligand structrue
SH3	(R/K)xXPxXP (class I)	1-200	PPII
	XPxXPx(R/K) (class II)	1-200	PPII
	RxxK (class III)	0.1-30	3_{10} helix
	RKxxYxxY	20-60	NA
	PxxDY	NA	NA
	(R/K/G)XXPPGX(R/K)	10-200	PPII
	R/K-rich	10-100	N/A
	WxxFxxLE	NA	α-helix

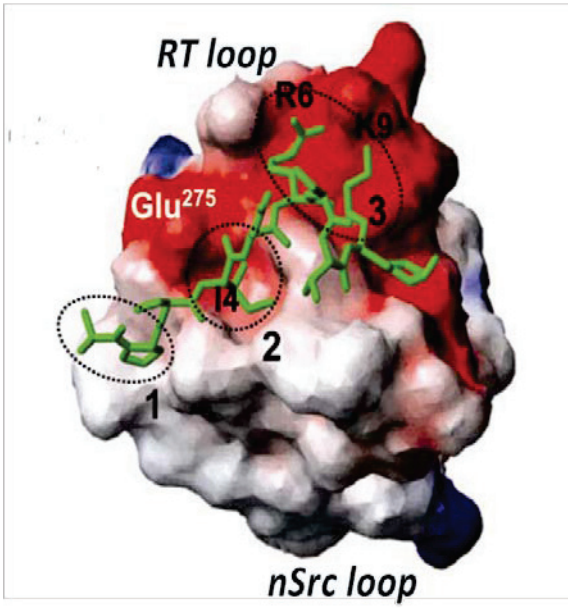


Figure 8.The structure of Gads SH3 domain in complex with the 11-mer peptide of APSID**RSTK**PA. Site 1 and 2 are the hydrophobic grooves and 3 is the specificity one. Red represents as positive charged region and blue is negative charged regions. Light gray is hydrpobic [17].

3.3 SH3 domains recognizing its binding proteins *via* tertiary contacts

SH3 domains can also associate with proteins *via* tertiary contacts that involved defined motifs as discussed above. For example, binding of the Src family kinases SH3 domains to a PPII helix formed by a short PxxP-containing fragment in N-terminal region of the HIV-1 Nef protein makes tertiary contacts at another

non-conventional binding site to increase their binding affinity. This will be discussed in detail below.

In addition to canonical or non-canonical core motif-mediated recognition, SH3 domains can also associate with another protein *via* tertiary contacts that does not involve a specific sequence motif. For instance, in the ubiquitination pathway, the yeast protein Sla1 SH3 domain interacts with ubiquitin through hydrophobic, tertiary contacts between their interfaces instead of the specific core binding motifs (Kd ~40uM) (Figure 9 A) [33, 46]. The critical amino acid of Phe residue on the SH3-binding surface (Phe409 of Sla1) appears to play an important role in mediating ubiquitin-binding since substitution to a Tyr residue abrogated the two proteins interaction. Structural analysis shows the hydrophobic binding site on Sla1-SH3 surface, consisting of Phe and other hydrophobic amino acids, overlaps with its predicted binding site for a PxxP-containing ligand (Figure 9 B). It is likely that ubiquitin and PxxP-containing ligands compete for interaction with a SH3 domain. Another interesting example is the regulation of guanine nucleotide-exchange factor Vav by Grb2 through heterodimerization of its C-terminal SH3 domain with the N-terminal SH3 domain of Vav. The interaction is mediated by a complementary interface between the two SH3 domains rather than a continuous sequence motif, so it can be only rationalized in terms of tertiary contacts [47].

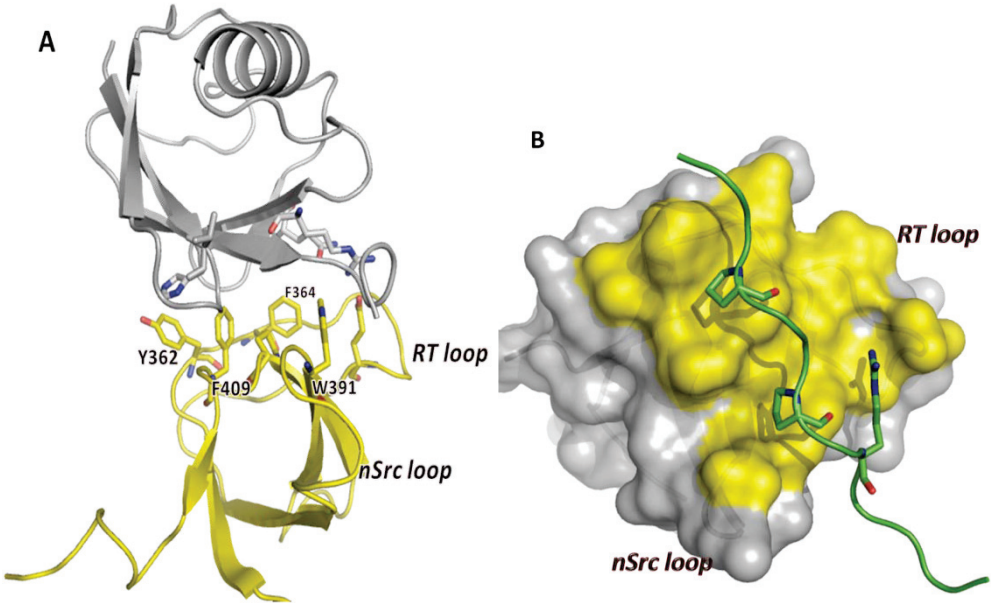


Figure 9. (A) Sla 1-SH3 domain (yellow) bound to ubiquitin (grey) (PDB entry:2jt4). Sla 1 is the yeast ortholog for CIN85. The canonical binding surface of the SH3 domain is used for binding ubiquitin. The hydrophobic amino acids involved in interface interaction are labeled. (B) The ubiquitin domain-binding surface (identified in yellow) mapped onto the p40 phox SH3 domain (PDB entry: 1W70); The peptide is shown in green.

3.4 Tertiary contacts between Src family kinases (SFKs) and HIV-1 Nef protein to increase selectivity and specificity

Typically, the interaction mediated by SH3 domains to their different ligands is weak and differ in strength only by about two orders of magnitude in Kd values. This type of weak interaction may be advantageous for the rapid formation and dissociations of signaling complexes in celluar processes. In addition, weak interactions mediated by inter or intra-molecular domains have been shown to play very important roles in regulating certain processes, and good targets for drug design such as Src family non-receptor tyrosine kinases (SFKs) (we will talk them in detail later). Despite the advantage of weak interactions in certain cellular context, the weak and non-specific interaction may be potentially inclined to cross-react within or between similar proteins in the same cell lines, and mimicked by virus proteins to compete with the normal cellular proteins and disrupt or disorder the normal cellular processes, for example the HIV-1 Nef interacting with SH3 domains of SFKs. Owing

to the weak or moderate affinity between proline-rich motifs and SH3 domains, it is reasonable to assume that greater affinity and greater specificity exists in certain cellular context to prevent or solve the high promiscuity displayed especially by canonical binding motifs. Here, I will discuss the SH3-regulating intra- or inter-molecular interactions in SFKs, and try to elucidate how HIV-1 Nef exploits specificity and selectivity to SFKs SH3 domain *via* tertiary association modes.

3.4.1 A story about tyrosine kinases

SH3 domains of SFKs belong to the subunit of tyrosine kinase. Tyrosine kinases are particularly important because they are involved in the treatment of cancer and signal transduction pathway. The first oncogene of v-Src was discovered as the SFKs activity in the late 1970s [12]. Although an enormous amount of progress has been made, we still do not exactly know how SFKs function to allow a cell to become a tumor cell and how they regulate the signal transduction pathway. There are two classes of tyrosine-specific protein kinases: receptor tyrosine kinases (RTKs) and non-receptor tyrosine kinases (nRTKs). nRTKs can be activated by association with RTKs or other substrates.

RTKs

RTKs are specific trans-membrane glycoproteins and exhibit considerable diversity in ligand binding. RTKs include the insulin receptor and the receptors for many growth factors (e.g. epidermal growth factor (EGF), platelet-derived growth factor (PDGF), fibroblast growth factor (FGF) and vascular endothelial growth factor (VEGF). They play vital roles in many important cellular processes, such as cell proliferation, differentiation, migration and metabolism [48, 49]. RTKs exhibit modular structure and widespread expression, and are in link with regulatory processes. In spite of the various binding ligands, RTKs subfamily shares conserved secondary structure. RTKs have an extracellular N-terminal region, a single hydrophobic trans-membrane spanning domain and an intracellular C-terminal region (Figure 10). The extracellular region can bind to ligands including EGF, PDGF or

insulin *et al* [50], and the intracellular region comprises catalytic activity responsible for the kinase activity of these receptors and sites for PPIs (protein-protein interactions), which catalyses receptor autophosphorylation and tyrosine phosphorylation of RTK substrates or adaptor proteins such as ALIX (we will discuss later in ALIX chapter).

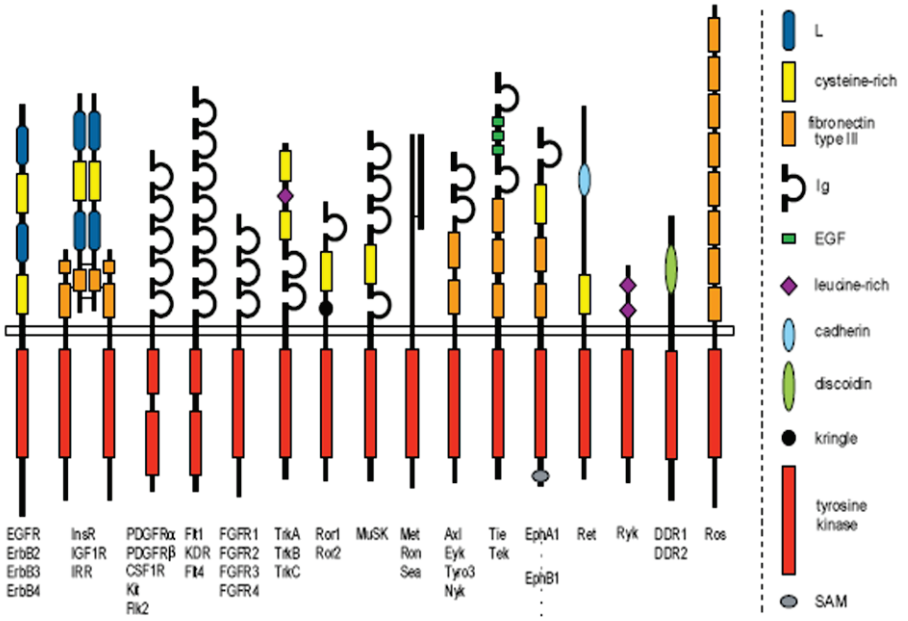


Figure 10.Architecture and domain organization for a variety of RTKs. The extracellular portions of the receptors are on top and the cytoplasmic portions are on bottom. The legend for the domain types is on the right side [50].

RTKs in cellular signaling

In general, activation of RTKs requires two processes: stimulation and transduction. RTKs are activated by the binding of their cognate ligands, expose the intracellular docking sites for downstream signaling proteins and transmit the extracellular signal to the cytoplasm, which is controlled through autophosphorylation of the tyrosine residues on the receptors themselves and the downstream signaling proteins. For example, when the EGF binds to the extracellular domains of EGF receptor, the cytoplasmic kinase domain is rapidly activated and recruits substrates around to the receptor. As a consequence, the activated receptor becomes autophosphorylated on multiple specific intracellular tyrosine residues, which

provides the binding sites for Src homology 2 (SH2) domains or phosphotyrosine binding (PTB) domains containing proteins such as Src and phospholipase C. These proteins send signal into the cell through different signal transduction pathways, such as Ras pathway or phospholipase C pathway. Besides these signal proteins, other adaptor proteins can also interact with the activated RTKs to mediate the downstream signal transduction, for example ALIX. Many groups have reported that PDGF or other RTKs can increase the activated SFKs *via* phosphorylated tyrosine in the kinase domain interacting with SH2 domain of nRTKs to release the repressed form [51].

RTKs crystal structure

RTKs are composed of two parts: extracellular domains and cytoplasmic domains. Crystal structures of extracellular domains of RTKs provide a structural basis for understanding ligand-receptor specificity. Analysis of ligand-receptor complex crystal structures shows that only a subset of domains in the extracellular portion of a RTK are involved in ligand binding [52-54]. For example, the extracellular domain of VEGF-2 receptors consists of seven immunoglobulin homology domains, and only the domains D2 and D3 of VEGFR-2 are bound to its specific binding ligand, VEGF-C. For the cytoplamic domains, crystallographic studies elucidate that the overall structure of the tyrosine kinase domain is similar to that of serine/threonine kinases: an N-terminal lobe comprising a five-stranded β sheet followed by a helix, and a C-terminal α helical lobe (Figure 11). There is a nucleotide-binding loop in N-terminal lobe between β1 and β2. The tyrosine-containing substrates bind to the activating-loop (A-loop) located in the C-terminal lobe. In the insulin receptor, there are three tyrosine autophosphorylation sites in the A-loop. One of these tyrosine, Tyr1162 (ordered in insulin receptor tyrosine kinase) is bound in the active site. When it is bound to the conserved Asp1132 by hydrogen-bond, the RTK adopts in an inactive form [49, 50, 55, 56]. Moreover, biochemical data and mutagenesis experiments are both consistent with the crystal structure showing that Tyr1162 plays a role for auto-inhibition in the tyrosine kinase domain.

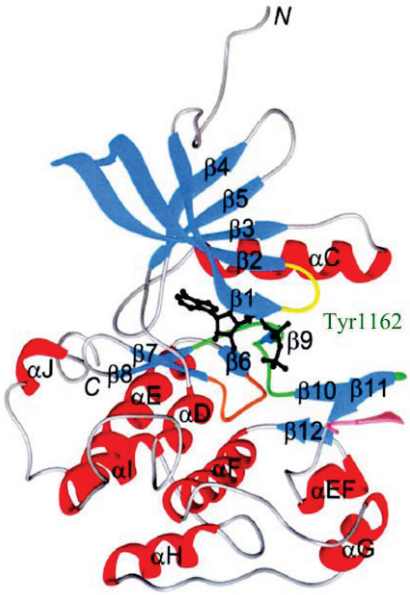


Figure 11 Ribbon diagram of the tyrosine kinase domain of the insulin receptor. The helices are shown in red, the β – strands in blue, the nucleotide-binding loop in yellow, the catalytic loop in orange, the A-loop in green, the ATP analog (AMP-PNP) in black and the peptide substrate in pink. The autophosphorylating Tyr1162 is marked in green. The termini are denoted by N and C [49].

nRTKs

Non-receptor RTKs (nRTKs) are cytoplasmic protein tyrosine kinases localized to a variety of intracellular compartments including the cytosol, the inner leavelet of the plasma membrane and the endoplasmic reticulum. nRTKs are integral components of the signaling cascades triggered by RTKs and by other cell surface receptors such as G-protein-coupled receptors and receptors of the immune system. There are at least eight different sub-classes of nRTKs, including Src-family kinases, Janus kinase (Jaks), Csk family kinase, Btk family kinases and Abl, among others. The largest subfamily of nRTKs is the Src family kinases (SFKs). They participate in varieties of cellular signaling processes by associating with and phosphorylating a number of different substrates. Multiple substrates have been described for Src kinase *in vivo*, including the PDGF and EGF receptors, the nRTK focal adhesion kinase (Fak), p130Cas which is an adapter protein involved in integrin- and growth factor-mediated signaling, and cortactin that is an actin-binding protein important for the proper

formation of cell matrix contact sites [50]. nRTKs do not contain extracellular and trans-membrane spanning regions. Some nRTKs harbor one or more amino-terminal acylation sites, such as needed for myristoylation or palmitoylation, (required for anchoring to the cell membrane). In addition to a tyrosine kinase domain, nRTKs possess domains that mediate protein-protein, protein-lipid, and protein-DNA interaction. Of the various protein-ligand binding domains, SH2 and SH3 domains are the most common ones in nRTKs [50]. SH2 and SH3 domains are both noncatalytic conserved domains and interact with different partners. SH2 is a compact domain of about 100 residues that binds to phosphotyrosine-containing substrates to regulate the functional activity of the substrates and localize the substrates within the cell. During the search for the first binding partner of tyrosine kinase Abl, it has been demonstrated that SH3 domains bind to a proline-rich motif [23, 28]. The functions of SH3 domains are not very clear yet, although more and more binding protein partners are found. In cellular processes, SH3 domains do not act independently. They have been reported to collaborate with intra-SH2 domain and help the kinase in repressed form. They also help recruiting substrates in cooperation with other domains in kinase proteins, such as the PDGF and EGF receptors *et al*. Whereas SFKs contain both SH2 and SH3 domains, some nRTKs lack SH2 and SH3 domains but possess other specific domains mediating protein-protein interaction. For instance, Jak subfamilies contain specific domains targeting the cytoplasmic portions of cytokine receptors. Fak possesses two domains regulating protein-protein interaction, which are the integrin-binding domain and the focal adhesion-targeting domain. Abl contains a nuclear localization signal, recruiting it to the nucleus and cytoplasm. Abl possesses not only SH2 and SH3 domains, but also an F actin-binding domain and a DNA-binding domain. (Figure 12)

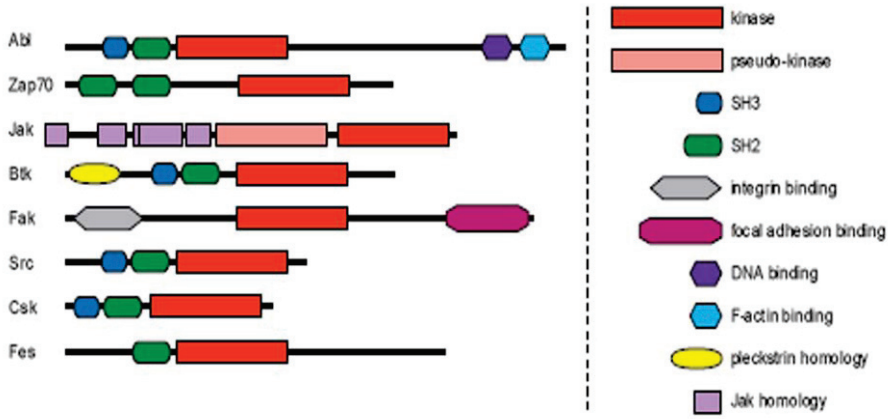


Figure 12.Domains organization for the major subfamilies of nRTKs. The amino terminus is on the left and the carboxy terminus is on the right [50].

Types, distribution and architecture of SFKs

SFKs are proto-oncogenes and participate in varieties of signaling processes, including mitogenesis, T- and B-cell activation, and cytoskeleton restructuring. There are 11 member proteins indentified as SFKs. Among them, Src, found in humans, chickens and other animals, is the cellular homologue of the viral v-Src. Src is expressed ubiquitously, but in some tissues such as brain, osteoclasts and platelets, it expresses to very high levels compared to most other cells. In contrast to the ubiquitous presence of Src, Fyn and Yes, the other members (Srm,Brk, Frk, Fgr, Lck, Lyn, Hck and Blk) show restricted tissue distribution. For instance, Srm is found in keratinocytes. Frk exists mainly in bladder, brain, breast, colon and lymphoid cells. Brk occurs chiefly in colon, prostate and small intestine; however it was initially isolated from a breast cell line. Lyn is expressed in T-cell. Lck is in T lymphocytes. Blk is expressed in B cells. The others express primarily in hemopoietic cells. SFKs share a high degree of structural similarity with common domain architecture and regulatory mechanisms. From the N- to C-terminus, SFKs contain a 14-carbon myristoyl group (SH4) that is required for membrane attachment; a unique domain which may impart distinct localization properties to the individual family members; an SH3 domain, an SH2 domain, an SH2-kinase linker, a protein-kinase domain and a C-terminal regulatory tail (Figure 13). Except for the N-terminal unique region, the sequences of the SFKs are highly conserved. It is well-accepted to use residue

numbering corresponding to chicken c-Src for SFKs.

The first discovered protein tyrosine kinase was the v-Src oncoprotein (a viral protein), which is encoded by the Rous sarcoma virus [57]. Although, v-Src possesses functional SH2 and SH3 domains, it lacks the autoinhibitory mechanism that controls the activity of the normal cellular form of the protein (c-Src), because the C-terminal tail of c-Src is replaced in v-Src by an unrelated sequence that lacks the inhibitory phosphorylation Tyr527 (according to c-Src) (Figure 14) [13, 51, 58, 59]. As a consequence, v-Src is constitutively active and is a potent transforming protein, leading to uncontrolled growth of infected cells. Conversely, the homologues of v-Src have a negative-regulatory domain in their C-terminus, providing these kinases with a regulatory motif. Mutations and disruption of the intra-interaction mediated by SH2 and SH3 domains, both of which disrupt the autoinhibitory mechanism, can invoke the Src kinases activity.

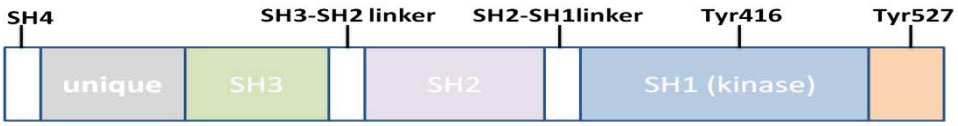


Figure 13 Organization of Src. The chicken numbering system is displayed. The tyrosine is numbered as c-Src.

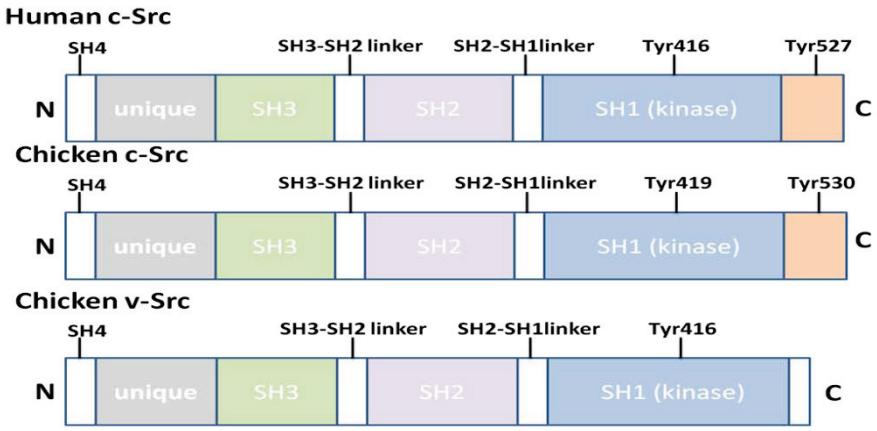


Figure 14.The architectures of Src homologues [59]. c-Src is denoted as the normal celluar homologue. Compare the molecular structures of human c-Src, chicken c-Src and chicken v-Src. All three proteins contain four SH domains and a unique amino-terminal domain between SH4 and SH3 which is of unknown function. The SH1 domain contains the kinase domain and a conserved tyrosin residue involved in autophosphorylation (Tyr419 in human c-Src; Tyr416 in chicken Src proteins). Chicken v-Src lacks the carboxy-terminal negative-regulatory domain and is constitutively active.

Regulation of SFKs

In normal cells, the level of SFKs can be regulated by ubiquitin-proteasome pathway. For example, the E3 ubiquitin ligase Cbl which can mediate Src polyubiquitylation and subsequent proteasome-mediated degradation.

It has been reported that the protein tyrosine kinases Csk and its homologue CHK phosphorylate the tail tyrosine and represses the catalytic activity of all members of SFKs [12, 60, 61]. Csk gene deletion experiments in mice show an embryonic lethal phenotype and constitutive activation of SFKs, confirming the importance of their regulation *via* this kinase. SFKs possess two important regulatory tyrosine phosphorylation sites: Tyr527 in the carboxy-terminal tail and Tyr416 in the catalytic domain (according the chicken Src numbering). When Tyr527 is phosphorylated by Csk, the SFKs adopt a closed inactive conformation *via* association with the SH2 domain (Figure 15). In addition to the SH2 domain, the SH3 domain is implicated as negative regulation, which stabilizes the intra-molecular interaction by

association with the linker between SH2 and kinase domain. Tyr527 is a very important negative regulatory site in SFKs. Under basal conditions *in vivo*, 90%-95% of SFKs are phosphorylated at this conserved residue. In contrast to the negative regulatory Tyr527 residues, a second regulatory phosphorylation site Tyr416 (in chicken c-Src) in the kinase domain plays a positive role. Its phosphorylation promotes kinase activity. Tyr416 localizes within the activation segment of the catalytic domain (Figure 14 and 15). Some groups reported this autophosphorylation site of Tyr416 is less of a positive control for kinase activity, because it is generally thought to be phosphorylated only in the absence of phosphorylation of Tyr527. Moreover, it can be compatible with Tyr527 in the inactive repressed conformation.

Besides the phosphorylation of Tyr 416, there are other ways to active SFKs: dephosphorylation of the tail by protein tyrosine phosphatases (PTPases); competition for the SH2 domain by a higher-affinity phosphotyrosine-containing ligand; and competition for the SH3 domain by a higher-affinity binding ligand. Once the SH2 or SH3 domain binds to their competitive binding partners or the phosphorylated tyrosine in tail is dephosphorylated, the autoinhibitory interaction is disrupted, leading to activation (Figure 15). Previous studies show that the deletion or mutation in either the SH3 or SH2 domain causes an increase in the catalytic activity of SFKs *in vitro* and *in vivo* and activates its oncogenic potency, suggesting that the SH3 binding interaction collaborates with the SH2 domain interaction to negatively regulate SFKs catalytic activity. Mutations of the SH3 domain also prevent phosphorylation of SFKs substrates and inhibit their partners binding to the oncogenic form of SFKs [12]. The active SFKs adopt an “open” conformation, and the released SH2 and SH3 domains are also exposed to the ligand-binding surface to interact with downstream or substrate signal proteins. For example, SH2 domains can associate with the activated PDGF receptor. The association of PDGFR with SFKs is to transmit the necessary mitogenic signal [51] (Figure 17). Since the first oncogene v-Src was discoveried, more and more studies are targeted on SFKs and a wide range of their substrates have been identified, including growth factor receptors, cytokine receptor and focal adhesion factors. These substrates interact with SFKs either directly or indirectly, and

regulate their activity [59]. Other proteins, for instance CRK-associated substrate (CAS) and focal-adhesion kinase (FAK) also interact directly with SFKs by association with their SH2 and SH3 domains to activate the kinases. In addition, they are also important for integrin signaling. SFKs are also activated by antigen coupled receptors and receptors that are coupled to G proteins, and related to the cytoskeletal rearrangements. The association of SFKs with the plasma membrane is considered essential for cellular transformation. In these processes, active SFKs highly transform cells. Moreover, these substrates involve in various vital cellular processes, including adhesion, invasion and motility. These functions might contribute to the tumor progression and metastasis. So SFKs have been linked with the development and progression of cancer, and targeted for cancer therapy. Moreover, the week interaction between subdomains of SFKs and these substrates have a tendency to be mimicked by virus and disrupted.

In summary, identification of SFKs binding proteins will facilitate understanding the mechanisms regulating and regulated by SFKs.

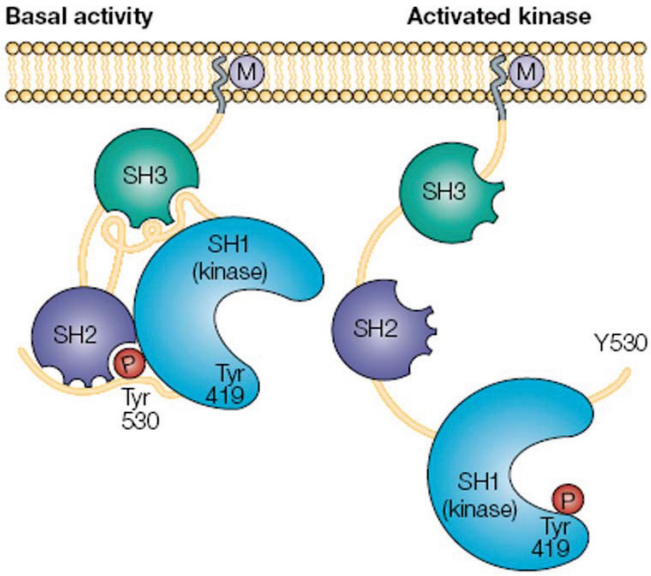


Figure 15. Schematic presentation of human c-Src kinase from repressed form to active from. Inactivation of human c-Src occurs when its C-terminal Tyr530 is phosphrylated and binds back to the SH2 domain, which provokes the interaction between the SH3 domain and the kinase domain. The intramolecular interactions result in a closed repressed kinase. Conversely, c-Src activation occurs with removal of inhibitory intramolecular interactions and the C-terminal phosphotyrosine. Tyr419 involve in activating kinase [59].

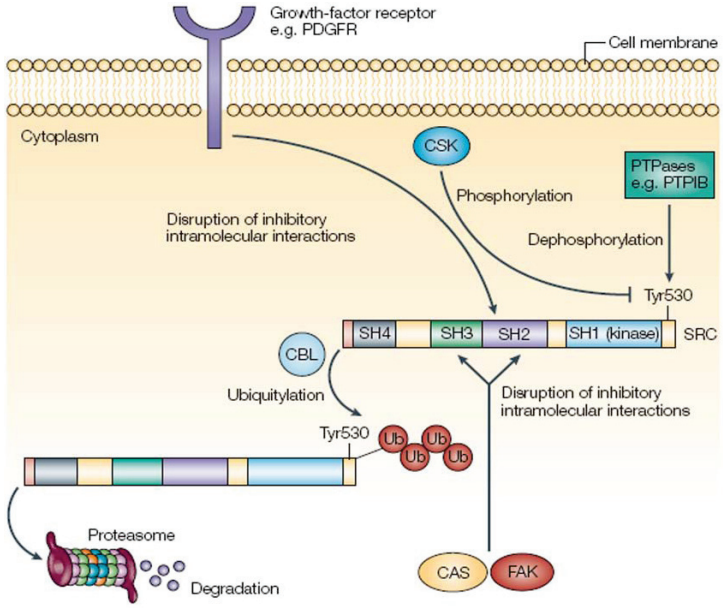


Figure 16.Regulation of SFKs in cellular processes [59].

3.4.2 Structural information about tertiary contacts mediated by SFKs SH2 or SH3 domain

Structure of Csk

The architecture of Csk is similar to Src-family kinases, consisting of a SH3, a SH2 and a kinase domain. Csk is regulated through its SH2 domain. The crystal structure of full-length Csk elucidates the molecular basis of the Csk-mediated regulation of SFKs through the tertiary contacts between intra-molecular subdomains [62].

The crystal structure[62] contains six molecules per asymmetric unit, which can be divided into three pairs of putative dimers. Two pairs adopt an active conformation, and one pair has an inactive conformation. In each molecule, the three domains, SH2, SH3 and kinase, adopt similar constellation, independent of the activation state. The SH2 and SH3 domains are diametrically opposite on the top of the small lobe of the kinase domain. There is no direct contact between the SH2 and SH3 domains (Figure 17, 18). Both peptide-binding pockets of Csk SH2 and SH3 domains are oriented outward permitting inter-molecular interactions. The most remarkable difference between Csk and other tyrosine kinases is that there is no phosphotyrosine in the activation loop of Csk (corresponding to Tyr416 of c-Src) and that Csk lacks the C-terminal tail inhibitory phosphotyrosine (corresponding to Tyr527 of c-Src). Based on these, it is obvious that Csk functions in a different mechanism to SFKs. In active molecular forms, the interactions between the kinase domain, the SH3-SH2 linker and the SH2-kinase linker stabilize its active conformation. Superimposing the active and inactive molecules of the Csk asymmetric unit, shows that the SH2 domain and kinase domain in both inactive molecules have shifted and that there is no interaction between SH2-kinase linker and kinase domain. In contrast, the SH2-kinase linker is involved in the interaction network of the active kinase. That SH3-SH2 linker and the SH2-kinase linker are tightly associated with the N-lobe of the kinase domain, which may predispose the SH2 domain to associate with its binding ligands. At the same time, the interactions between SH2 and kinases or other binding partners may help Csk activation. In other words, it is possible that ligands for the Csk SH2 and SH3 domains can stabilize the active conformation. This can be proved by the interaction

between the SH2 domain of Csk and its strong ligand of the SH2 domain, the phosphorylated Cbp/PAG. In addition to the absence of activating tyrosine in kinase domain, there is no intra-interaction between SH3 domains and the SH2-kinase linker in Csk like usual SFKs to perform a restrain state. Therefore, it is possible that Csk activates SFKs by first association with SFKs Tyr416 *via* its SH2 domain, and then phosphorylating Tyr527 in SFKs C-termini to inactivate SFKs [62, 63].

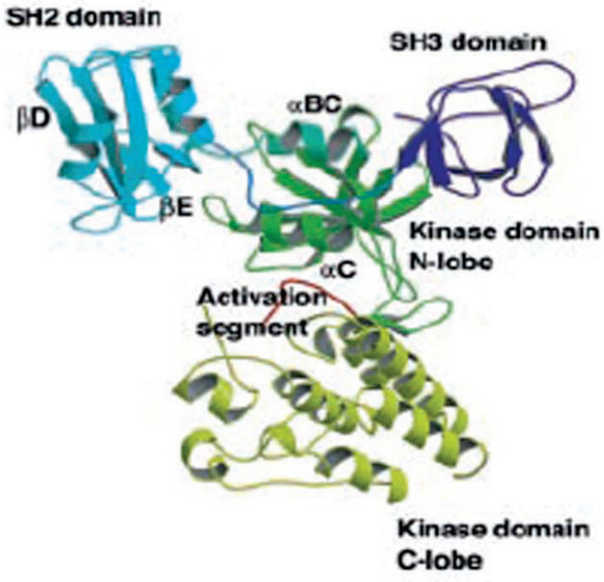


Figure 17. Ribbon diagram of Csk structure in its active conformation. Domain structures are represented by different colors: SH2 in blue, SH3 in purple, N-lobe in dark green, C-lobe in light green, and the activation segment in red [62].

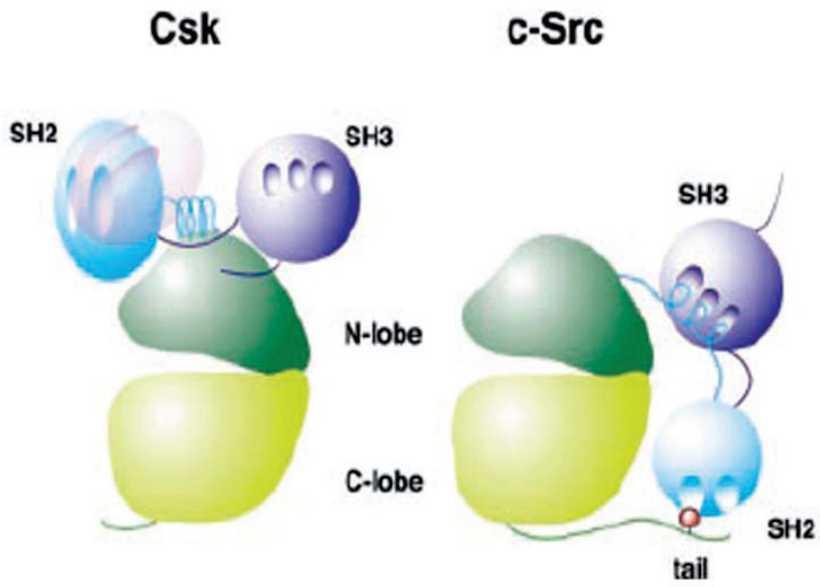


Figure 18 Comparison of the domain structures of Csk (left) and Src (right). Peptide binding pockets of the SH2 and SH3 domains are shown by hollows. Phosphatyrosine 527 on the tail is colored red. The SH2-kinase linkers are light blue, and the SH3-SH2 connectors are gray. The rotated SH2 domain of inactive molecular is colored pick [62].

Structure of SFKs

The crystal structures of SFKs provide a molecular framework for understanding the regulation of this class of nRTKs by intra-molecular association, which is consistent with the results obtained from mutation and enzymology studies, and provide clear image of the specific interaction between various SFKs and ligands, such as the HIV-1 Nef protein. As in the Csk crystal structure, there are two lobes in the kinase domain of SFKs: the smaller amino-terminal kinase lobe and the large carboxy-terminal lobe. The smaller lobe has a predominantly antiparallel β-sheet structure. The large lobe is predominantly α-helical and responsible for binding the protein substrates. The catalytic sites of SFKs lie in a hydrophobic cleft between the two lobes, which can be protruded or embedded depending on the orientation of the two lobes. When the kinase is inactive, Tyr416 situated in the activation loop is buried into the hydrophobic pocket to occlude the substrate-binding site and protect Tyr-416 from autophosphorylation. The crystal structures of restrained Src and Hck have been determined. In both cases, the proteins were phosphorylated on Tyr527 in the

C-terminal tail, but not on Tyr416 in the activation segment (Figure 19) [64, 65]. Tyr527 is associated with the SH2 domain by hydrogen bonds to help the kinase domain adopting a compact conformation in the restrained state. Simultaneously, the SH3 domain is docked onto an internal polyproline type II helix that is formed by the SH2-kinase linker. The polyproline helix is at the back surface of the the small lobe of the kinase domain. It is reported that Trp260 situated in the SH2-kinase linker is fixed by interaction with residues at the SH3-binding site in inactive kinases. Mutation of Trp-260 to alanine in Src/Hck results in constitutive activation [50].

SH2-kinase linkers are different in various SFKs, such as Hck and Src. The SH2-kinase linker in Hck contains a PxxP motif, the core binding sequence for SH3 binding, but the linker in Src contains only one proline. There is no basic amino acid flanking the proline-rich sequence, which is responsible to high-affinity binding of SH3 domain to its ligand. Because of the sub-optimal binding properties of SH2-kinase linkers, SFKs tend to associate with higher affinity SH3-binding ligands, which may lead to transformation of cells, as illustrated by the ligand HIV-1 Nef.

In summary, the structures of Src and Hck clearly show that it is the SH3 domain that plays more direct role in the control of catalytic activity. SH3-binding protein, such as HIV-1 Nef, can activate Hck, which is shown that SH3 domain displacement provides maximal inactivation of the kinase. HIV-1 Nef can transform cells by activating Hck *in vivo*. The SH2-kinase linker is poorly conserved between Hck, Src and other SFKs, which reflects by the differences in the target specificities of the SH3 domains. For example, HIV-1 Nef can bind strongly to the Hck SH3 domain, but only weakly or not at all to the other Src family kinases (we will discuss this later).

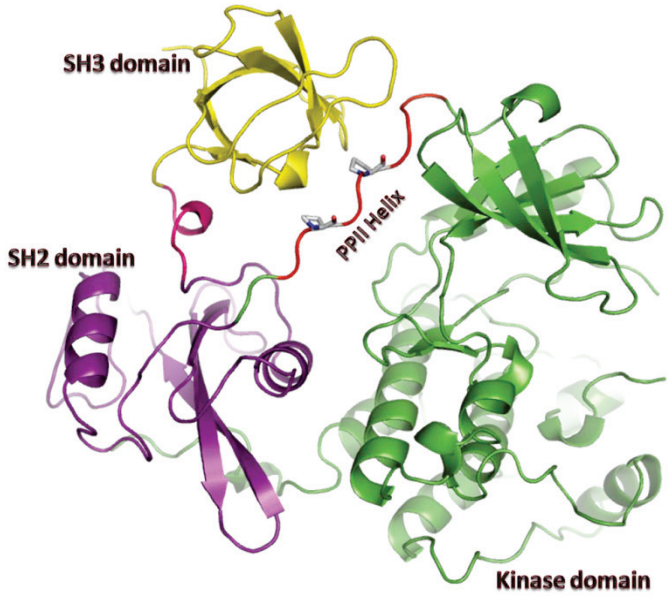


Figure 19. Ribbon diagram illustrating the structure of Hck (PDB code 1AD5) [50].

3.4.3 Analysis of tertiary contacts between SH3 domains and HIV-1 Nef

As we have discussed above, despite the lack of distinguishing features, the ligand binding surface on SH3 domain can accommodate various core binding motifs. However, little or low specificity is observed by the linear interactions induced by normally canonical and/or non-canonical binding core motifs *in vitro*. Besides, in most cases, these differences in binding affinity are not great enough to explain the selectivity and specificity interactions between SH3-containing proteins and their binding partners *in vivo*. Starting from about 13 years ago, SH3 binding partners that engage so-called tertiary contacts are found to expand the pool of SH3-recognition ligands. One of them is HIV-1 Nef, which bind to SH3 domains using also another surface in addition to the canonical PxxP core motif. This tertiary interaction allows Nef to definitely discriminate proteins that share the conserved hot-spot on the SH3 surface, and to significantly improve the specificity and selectivity of its interactions with different SH3 domains.

The *nef* gene of human (HIV-1 and HIV-2) and simian immunodeficiency viruses (SIV) encodes the Nef protein, which is needed for respond for optimal viral replication and onset of AIDS pathogenesis. Nef can interact with cellular protein kinases and transform cells to increase viral infectivity and replication. For example

Nef binds the SFKs Hck, Lyn, Fyn and Lck. The Nef protein contains a PxxPxR core motif which is well-known to interact with SH3 domains of SFKs, but presents different binding-affinity to various SH3 domains of SFKs. The discrimination for binding affinity was explained by binding studies between SH3 domains and proteins or peptides, which suggest that some extra specific binding sequences outside the core binding motifs interact with SH3 domains *via* tertiary contacts to distinctly strengthen specificity and selectivity. For example, binding studies show that a 12-residue peptide encompassing the Nef core binding motif of PxxPxR bound with a 300-fold reduced affinity to the SH3 domain from Hck than did the intact Nef protein. Meanwhile, HIV-1 Nef binds strongly to the Hck SH3 domain, but only weakly or not at all to SH3 domains of certain other SFKs such as the Fyn SH3 domain. Based on the above results, the HIV-1 Nef protein was shown to target various SH3 domains of SFKs by using not only a PxxPxR motif for the interaction, but also a supplementary "tertiary' surface region [39, 66]. This tertiary recognition, which is provided by the folded core domain of HIV-1 Nef, allows this protein to greatly enhance its specificity and selectivity towards SH3 domains. That is to say, HIV-1 has potentially evolved a novel mechanism to exploit additional interactions to allow it to better compete with pre-existing cellular proteins.

X-ray crystallography and mutagenesis experiments have revealed that the different specificity and selectivity lies in a variable loop, the "RT loop" in SH3 domains, positioned close to the conserved SH3 residues implicated in the binding of proline-rich (PxxPxR) motifs. Binding experiments demonstrate that despite the weak affinity of $Fyn_{(WT)}$SH3 for HIV-1 Nef (Kd>20uM), a mutant $Fyn_{(WT)}$SH3 with a single amino acid substitution (R96I) in its RT loop greatly increases its affinity (Kd 380nM) comparable with that of Hck-SH3 (Kd 250nM) [39, 67]. The selectivity and specificity between various SH3 domains come from the difference between Ile96 and Arg96 residues, which both are located in the RT loop of SH3 domains and interact with HIV-1 Nef by inserting and interacting with a hydrophobic pocket on the surface of HIV-1 Nef.

In 1996, Lee, C. H. and his colleagues determined the crystal structure of HIV-1

Nef in complex with $Fyn_{(R96I)}$SH3, which firstly shed light on the interaction between HIV-1 Nef and SH3 domain of SFKs through structural information. In the complex structure, the RT loop-interacting hydrophobic crevice on the HIV-1 Nef surface is formed by two helices, and is bordered by eight conserved residues (Trp113, Leu110, Arg106, Leu100, Leu97, Leu91, Phe90; ordered from the left to right side). Moreover, the hydrophobic pocket is separated from the SH3 interaction surface by the side chains of two aromatic residues (Trp -113 and Phe – 90) (Figure 20 and 21). Both the N-terminal PxxP region and the two helical layer of HIV-1 Nef form the SH3 interaction surface (Figure 21). Residues 71-77 of HIV-1 Nef span the conserved PxxP core motif and adopt a left-handed PPII helix. As we mentioned above, the PxxP motif occupies the conserved hydrophobic surface between RT loop and nSrc loop. Of them, the two proline residues that define the PxxP motif of Pro72 and Pro75, pack against the conserved hydrophobic residues (Tyr91, Trp119, Pro134 and Try127) on the SH3 aromatic pocket. The specific Arg77 side chain of the HIV-1 Nef PxxPxR motif is stacked closely against the face of Trp119 side chain on $Fyn_{(R96I)}$-SH3, and interacts with Asp100. The interaction between Arg77 (HIV-1 Nef) and Asp100 ($Fyn_{(R96I)}$SH3) disrupts the extensive hydrogen-bond network mediated by Asp100 and its surrounding amino acid in $Fyn_{(R96I)}$SH3. Then, the RT-loop flexibility is increased, which facilitates residue Ile96 to insert to the hydrophobic pocket on HIV-1 Nef surface and to be packed against the conserved side chains of Leu87, Phe90, Trp113 and Ile114 of HIV-1 Nef [66, 67].

Shortly after, Stefan Arold solved the crystal structure of HIV-1 Nef in complex with $Fyn_{(WT)}$SH3, which provided further information about the specificity and selectivity between HIV-1 Nef and SH3 domains. By comparison of the structures of $Fyn_{(R96I)}$SH3 and $Fyn_{(WT)}$SH3 in complex with HIV-1 Nef, it showed that residue Arg96 in the former complex mimics Ile96 in later one to insert into the hydrophobic pocket on the Nef surface. And the major difference between the two complexes comes from the hydrogen-bond network caused by the interaction of Arg77 (Nef) and Arg96/Ile96 (SH3). The side chain of Arg96 in $Fyn_{(WT)}$SH3 forms a hydrogen bond with its neighboring Thr94 ($Fyn_{(WT)}$SH3), which maybe decrease the flexibility of

Arg96-containing RT loop. In contrast, Ile96 in $Fyn_{(R96I)}$SH3 which is shorter and no charged, can insert deeply into the hydrophobic pocket. In addition, Ile96 in $Fyn_{(R96I)}$SH3 has any hydrogen-bond formation with neighborhood, which increases the flexibility of I96-containing RT loop. Comparison of the crystal structures of HIV-1 Nef in complex with $Fyn_{(R96I)}$SH3 and $Fyn_{(WT)}$SH3 shows that it is a water, which displaces the Arg96 side chain position on $Fyn_{(R96I)}$SH3-Nef complex to interact with neighboring residue Thr94 (Figure 22) [66]. So, the increase in affinity between $Fyn_{(R96I)}$SH3 and HIV-1 Nef is attributed to not only the inherent flexibility in the RT loop, and but also the match or mismatch of the hydrophobic pocket on HIV-1 Nef surface with the amino acid 96 in RT loops of SH3 domains.

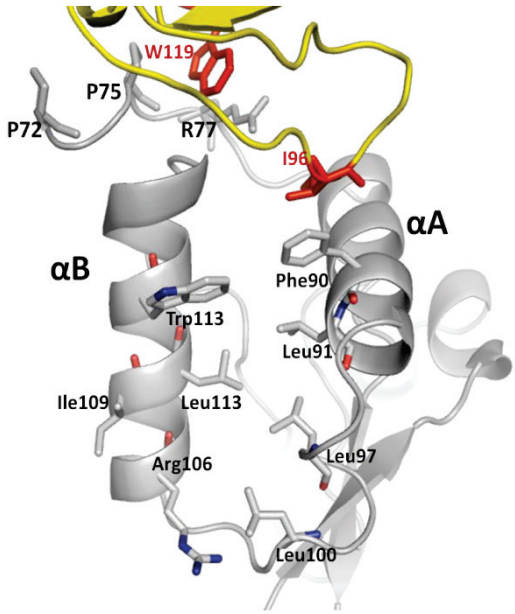


Figure 20.The hydrophobic crevice is formed between two helices of αA and αB in Nef (PDB entry IEFN). $Fyn_{(R96I)}$SH3 is in yellow; HIV-1 Nef is in grey. All the important residues are labeled. The bordered eight conserved residues on αA and αB are labeled in black [67];

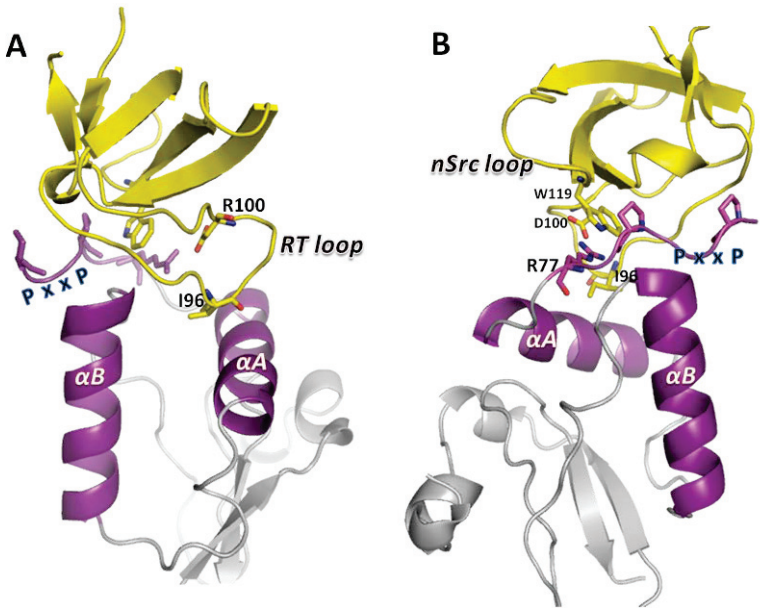


Figure 21. Overall structure of HIV-1 Nef in complex with $Fyn_{(R96I)}$SH3 (PDB entry IEFN). (A) and (B) is approximately orthogonal. $Fyn_{(R96I)}$SH3 is presented in yellow; The αA and αB in HIV-1 Nef is marked as purple and others are in grey.

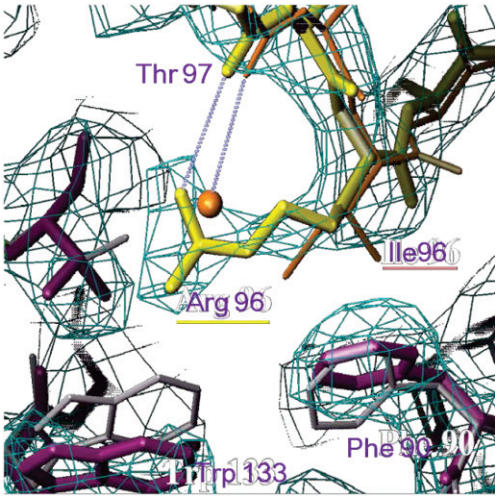


Figure 22.Superimposition of structures of $Fyn_{(R96I)}$SH3 (PDB entry IEFN) and Fyn SH3 (PDB entry 1AVZ). Hydrogen bonds are drawn as dashed lines. A water is presented as orange sphere [66].

In addition, other studies also demonstrated that HIV-1 Nef protein exploits a novel mechanism to increase its specificity and selectivity to various SH3 domains through its N-terminal PxxPxR motif and its hydrophobic pocket on the surface. In 2000, Yves Collette *et al* found that Nef proteins derived from HIV-1, HIV-2 or SIV

exploit different mechanism to interact with SFKs. Sequence alignment shows all of the three Nef proteins contain the N-terminal conserved PxxPxR core motif, but the amino acids in the RT loop-recognition pockets ("specificity pocket") are varied. HIV-1 Nef has the three residues of W113, T117 and Q118 in this pocket, which are replaced by Y, E and E in HIV-2/SIV Nef, respectively. The former "specificity pocket" is more hydrophobic than the latter two, which contributes mainly to the different mechanisms explored by HIV-1, HIV-2 or SIV. As a result, and in contrast to HIV-1 Nef, the Nef protein encoded by the SIV and HIV-2 poorly interact with Hck-SH3, but preferred for Fyn- or Src-SH3 that contains Arg96 in their RT-loop. Combination of mutagenesis, yeast two-hybrid system and immunoblotting were used to explore the different binding affinity between HIV-1, HIV-2 or SIV and SH3 domain. First, the conserved PxxPxR core motif was demonstrated again to be integral for the interaction between HIV-1 Nef and Hck SH3. Mutating the two core proline residues abolished the HIV-1 Nef- Hck SH3 association. Conversely, the interaction of the SIV Nef-Hck SH3 is PxxP-independent. Second, a three amino acids substitution in the specific hydrophobic pocket of SIV Nef to HIV-1 Nef (Y113W, E117T, E118Q) explains the change in selectivity (Figure 23) [68]. Since the "specificity pocket" in HIV-1 Nef is dominantly hydrophobic and Hck-SH3 have a hydrophobic amino acid Ile96, the interaction between Ile96 of Hck and the hydrophobic crevice of HIV-1 Nef displays higher affinity than the hydrophilic-residue containing SH3 domains, such as Fyn and Src SH3. The one of SIV/HIV-2 Nef is less hydrophobic and more negatively charged. So the substitutions of the three amino acids in SIV/HIV-2 Nef to HIV-1 Nef corresponding residues are consistent with the above property, which means the mutation increases the hydrophobic character and diminishes negative charges. So the fundamental mechanisms targeting SFKs by HIV-1 Nef, HIV-2 Nef and SIV are different. Nef interacts with a SH3 domain *via* a "pocket" proximity to the canonical proline-rich motif to increase its selectivity. Besides, HIV-1 Nef explores a unique mechanism to improve the virus infectivity and prevalence by combination of its N-terminal PxxPxR motif and hydrophobic pocket on the surface. Furthermore, in 2004, Choi, H.J. and his colleagues elucidated that the triple amino acids of Ala83

His116 and Tyr120 on HIV-1 Nef conserved hydrophobic surface mainly contribute to the high affinity binding by comparison of the allelic variants of HIV-1 Nef in terms of Hck binding and activation *in vivo* [69]. Among them, Ala83 and Tyr120 form part of hydrophobic pocket that contacts Ile96 in the RT loop of the Hck SH3 domain in the Nef-SH3 crystal structure. Then hydrophobic pocket is critical for HIV-1 Nef interaction with cellular partners and good target for disrupting Nef signaling through Hck in HIV-infected macrophages, potentially slowing disease progression.

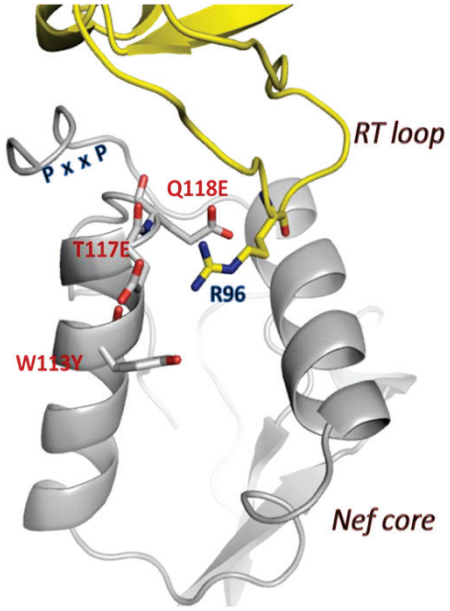


Figure 23.Cartoon representation modeling the interaction between Fyn SH3 domain and HIV mutant. The Fyn SH3 domain is on the top, which stretchs its R96 located in RT loop into the hydrophobic crevice of SIV on the bottom. Three residues if Tyr, Glu and Glu in SIV, were changed as the corresponding residues of Trp113, Thr117 and Gln118 in HIV-1 Nef。

As mentioned above, a SH3 domain of SFKs constitutes a conserved and critical component of HIV-1 Nef's action [69-71], and is required for development and progression of AIDS *in vivo* [72-74]. Therefore, the Nef:SH3 interface appears a promising target for anti-HIV drugs. This prompted targeting the specific hydrophobic pocket of Nef with small engineered RT-loop fragments selected by phage-display, and with chemical mimics obtained by combined high-throughput docking and experimental screening of chemical libraries [70]. Interestingly, these compounds act

as potent inhibitors of HIV-1 Nef function when introduced in cultured cells, but their potential side effects are not known. To avoid high risks of side effects such drugs require the SH3 recognition mode of HIV-1 Nef to be fairly unique among cellular SH3-ligand interactions. To date, no 'Nef-like' SH3 recognition has been documented, even among the several atypical SH3-binding motifs described. However finding the answer to this question is challenging because detection of such tertiary contacts is not possible by sequence analysis alone, and requires the difficult structure determination of protein-protein rather than protein-peptide complexes.

. Our collaborators took advantage of the well-defined $Fyn_{(WT)}$- and Hck-SH3 domains to design a large-scale screen for Nef-like RT-loop dependent SH3-ligand interactions in a human fetal expression library. Yeast two-hybrid screening using SH3 domains as baits revealed that only 2-3% of clones from a whole human fetal library have possible SH3 selectivity similar to HIV-1 Nef. Here, we will present one of the identified proteins: ALIX (**AL**G-2-**i**nteracting protein **X**) as a novel Hck SH3 binding protein.

4. Structural and Biochemical studies of ALIX

4.1 History background

ALIX is a multifunctional adaptor protein that has been identified as a binding partner of various proteins and accordingly implicated in diverse cellular processes. Initially, it was isolated as a binding partner of ALG-2(apoptosis-linked gene-2) by two groups in 1999. ALG-2 is a small calcium-binding protein whose expression is required for apoptosis in lymphocytes in response to variety of stimuli. In their paper, the authors used full-lengh ALG-2 gene as bait in a yeast two hybrid screen of a mouse brain cDNA library to obtain ALG-2-interacting protein X, named ALIX [75, 76]. So ALIX was initially implicated in the regulation of apoptosis by interaction with ALG-2 which links ALIX to apoptotic processes. Since it has been demonstrated the association is mediated by a SH3/proline-rich recognition mode (Suzuki et al. 2008). Shortly after, ALIX is extensively studied by many groups. The C-terminal half (ALIX-CT, denoted as PRR [PR region] now) is related to apoptosis process in

another way, interacting with endophilin through a SH3/proline-rich domain interaction. Overexpression of ALIX-CT induces cytoplasmic vacuolization, which protects cells from death [76, 77]. In 2000, Chen *et al* [78] isolated ALIX as a binding partner of SETA (the src homology 3 (SH3) domain-containing expressed in tumorigenic astrocytes) by yeast two-hybrid screen. SETA proteins locate to the actin cytoskeleton and sensitize astrocytes to apoptosis induced by UV irradiation. Interference in the interaction between ALIX and SETA will also lead to apoptosis process. Shortly after the discovery in 2003, ALIX are found not only involved in the endosomal sorting processes by its binding partners of CHMP4 (charomatin-modifying protein; charged multivesicular body protein) [79], but also involved in the human immunodeficiency virus budding by serving as a link to L domain of HIV-1 p6, EIAV p9 or Tsg101(tumor susceptibility gene101), a component of the class E vacuolar protein sorting (Vps) machinery [80, 81]. Later on, more and more studies elucidate other cellular functions of ALIX by cell biology, biochemistry and crystallography methods. ALIX C-terminal containing-polyproline region is associated with various SH3 domain, which plays a key role in determining of the functional association of ALIX with its partner proteins, such as the receptor and cytoskeleton-associated tyrosine kinases.

4.2 Architecture

ALIX is a cytosolic protein in mammalian cell and travels between cytoplasm and cytoskeleton by interaction with different partners to participate in vital cellular processes including endocytosis, cell adhesion, cell death and other cellular processes [77-79]. Recent crystal structures revealed that human ALIX is comprised of three domains. From its N to C terminus, it is the N-terminal Bro1 domain (residues 1-358), a central V domain (residues 362-702) and a C-terminal proline-rich region (PRR; PR domain; residues 703-868) (Figure 24). The ability of ALIX to bind various partners stems from its domain architecture (Table 3). Bro 1 domain and V domain have been described to interact with cellular proteins and virus composition respectively that facilitates virus budding, but C-terminal PRR is critical for interaction with the

majority of proteins that connect ALIX to cellular processes [82, 83], one of which is its association with SH3 domains [84].

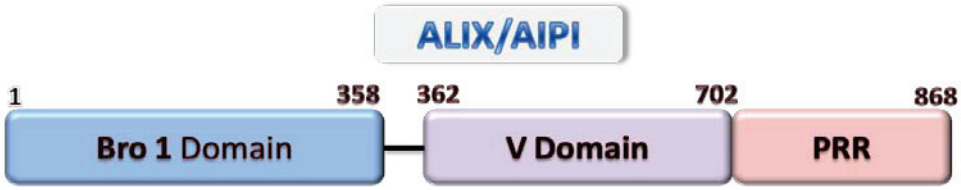


Figure 24.Schematic diagram depicting the sub-domains of ALIX.

Table 3 Summary of ALIX protein-protein interaction

ALIX-binding protein	ALIX-binding motif	Binding site in ALIX	Cellular activity
CHMP4 (ESCRT-III)	M/L/IxxLxxW	Bro 1 domain, Patch 1	MVB sorting and viral budding
tyrosine kinase Src	SH2 domain	Bro 1 domain, Patch 2	Growth factor receptor endocytosis; focal adhesion remodeling
Gag	YPXnL	V domain, Phe676 pocket	Viral budding
SETA	SH3 domain	P_{740}TPAPR$_{745}$	Growth factor receptor endocytosis; focal adhesion remodeling
Endophilin	SH3 domain	P_{755}ARPPPP$_{761}$	Growth factor receptor endocytosis
Src	SH3 domain	P_{752}QPPAR$_{757}$	Growth factor receptor endocytosis; focal adhesion remodeling
ALG-2	Unkown	PGY repeats (aa 802-813)	Apoptosis
TGS101 (ESCRT-I)	UEV domain	P_{716}SAPSIP$_{722}$	MVB sorting and viral budding
ubiquitin E3-ligase POSH	SH3 domain	PRR	MVB sorting and viral budding
ubiquitin E3-ligase POSH	Unkown	Bro 1 domain	MVB sorting and viral budding

4.3 Multiple personalities of ALIX in cell

ALIX is expressed widely in tissues. It can act as an adaptor protein that links different cellular proteins into networks or directly contributes a specific function to distinct molecular machineries (Figure 25 and Table 3). Much of the interest in ALIX has focused on its role in a functional connection to membrane trafficking regulated

by ESCRT machinery, and its PR domain mediated SH3/Proline-rich mediating mechanisms such as ALG-2 related apoptosis and Kinase-related regulation *etc*.

ESCRT-mediated protein sorting and virus budding

ESCRTs (endosomal sorting complexes required for transport), assembled by class E Vps (vacuolar protein sorting) proteins, are involved in membrane trafficking. ALIX is involved in the ESCRTs-related membrane traffic for degrading the cell-surface receptors and retrovirus budding from the plasma membrane. Both of them are *via* the interactions with ESCRT components and ALIX sub-domains. There are two ESCRT components involved in interacting with ALIX. ESCRT III binds to N-terminal Bro 1 domain *via* its CHMP4 subunit, and ESCRT I is associated with the C-terminal PR domain *via* its TSG101 subunit. The P[T/S]AP motif in the C-terminal region of ALIX is the conserved binding motif for TGS101. TGS101 has a ubiquitin-binding module that can be ubiquitylated and then recognized by MVB (multivesicular body) cargoes. Rather than associating with MVB cargoes, ESCRT III recruits ALIX directly to endosomes by interaction of its CHMP4 subunit with the Bro 1 domains of ALIX. Retrovirus can hijack the endosome sorting machinery for their virus budding purpose by interaction with ALIX. The interaction between ALIX and virus is mediated by late domains of viral Gag proteins, which promote the release of virion from the infected cell. To date, there are three different late domains characterized, containing P[T/S]AP, YPXnL and PPPY motifs respectively. HIV-1 $p6^{Gag}$ comprises of P[T/S]AP and YPXnL motifs simultaneously, which are demonstrated to interact with Tsg101 and ALIX respectively. Unlike to HIV-1 $p6^{Gag}$, EIAV P^{9Gag} contains only YPXnL domain and is strictly depend on ALIX for its budding. Both of the two virus YPXnL domains recognize the hydrophobic Phe676 pocket in V domain of ALIX (below, the structure information will explain it in details). The PPPY motif is found to associate with Nedd4-like ubiquitin ligases.

In addition, ALIX also binds to lysobisphophatidic acid (LBPA), which may contribute to its targeting to inward-budding vesicles in the late endosome [85].

The relation between ALIX and tyrosine kinases

RTK signaling plays a central role in cellular growth control and is often deregulated in cancer. Escape from endocytosis is one common characteristic of RTKs that have undergone oncogenic deregulation. Endocytosis of RTKs is triggered by Cb1-mediated ubiquitination. Cb1 binds to phosphotyrosine residues on active EGFR, which activate Cb1 by phosphorylation. Activated Cb1 associates with SETA-endophilins complex, and then ubiquitates itself and SETA, which serves to promote rapid internalization of activated EGFR. ALIX negatively regulates the EGFR endocytosis by interfering with the Cb1-SETA interaction. ALIX comprises proline-rich motifs in its C-terminal which can associate with SH3 domains of SETA and endophilins. Therefore, over-expression of ALIX allows it to compete with Cb1 for SETA-endophilin complex binding and inhibits the concomitant activity of Cb1, consequently reducing the EGFR internalization.

Schmidt *et al* found that the Src kinase can interact with ALIX and then regulates the process of ALIX-mediated regulation of EGFR internalization [84, 86]. Phosphorylation of ALIX by Src prevents ALIX from binding to SETA, thereby enabling the association of Cb1 with the SETA-endophilin complex. ALIX is phosphorylated at Tyr319 in the presence of normal levels of Src. The interaction between ALIX and Src is a two-step process: first, the SH2 domain of Src binds to ALIX phosphorylated Tyr319 locating in Bro1 domain, thus activates Src and phosphorylates ALIX in turn; second, a proline-rich motif ($_{752}$PQPPAR$_{757}$) in the ALIX C-terminus binds to SH3 domain of Src. As a consequence of elevated Src levels, Src continues phosphorylating ALIX at its PR domain rich in tyrosine. The hyperphosphorylation causes ALIX to translocate from the membrane and cytoskeleton to the cytoplasm, which reduce the interaction between ALIX and SETA-endophilin complex [84, 86]. The interaction between ALIX and SETA is also found to be negative regulation of FAK and Pyk2 (phosphotyrosine kinase-2) that are the kinases regulating cell adhesion [84, 87].

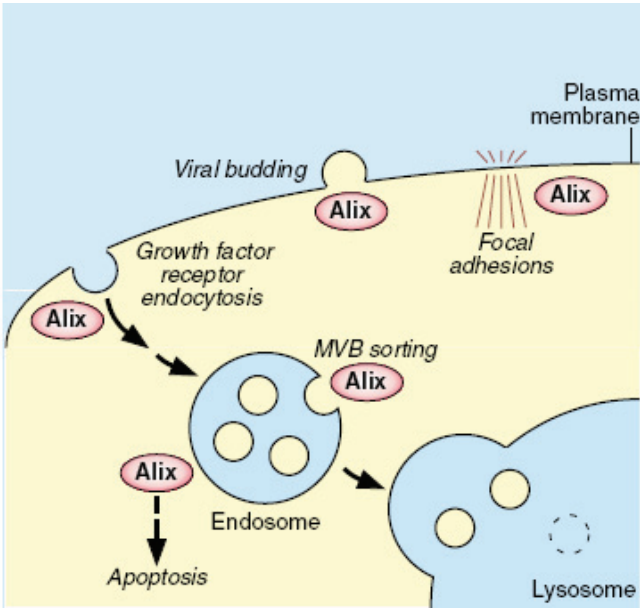


Figure 25. Overview of the cellular mechanisms involving ALIX [87]

4.4 X-ray structural analysis of ALIX

During the past decade, interest in the multiple roles of ALIX has risen steadily. Combination of mutagenesis and cellular biology methods has identified some of the cellular binding partners and the cellular pathways for ALIX. However, the details of these mechanism regulated by ALIX were unknown until crystal structures of various constructs of ALIX were published in 2007 in the journal Cell [88]. Robert D.F. and his colleagues elucidated the atomic structure or the Bro1 domain, V domain and Bro1-V domain. They provided a platform for structural analysis of ALIX-partners binding module, even though there is no ligand bound. Later on, some groups released the crystal structures of ALIX in complex with its known binding partners CHMP4 and the viral YPXnL late domain, respectively [82, 89, 90]. All the structures (Table 4) shed light on the ALIX-mediated cellular processes, and facilitate a mechanistic understanding of ALIX. However, there is still no structural analysis available of the PRR region, especially, its interaction with its binding partners.

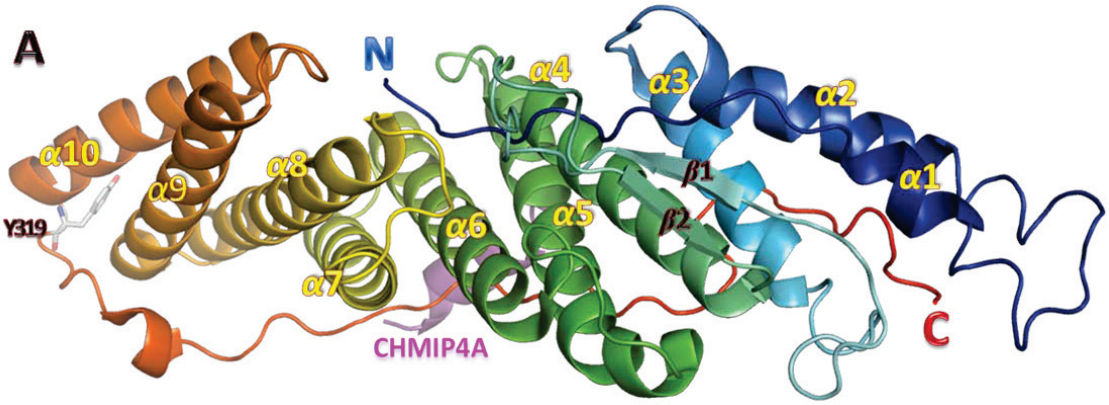
Table 4 Release ALIX-related crystal structure in PDB

Release Date	PDB entry	Compounds	Publish jounral
2007-02-20	2OJQ	V domain	Nat. Struct. Mol. Biol
2007-03-27	2OEX	V domain	Cell
	2OEW	Bro 1 domain	
	2OEV	ALIX (Bro 1 and V domain)	
2007-12-18	2R05	ALIX in complex with the HIV-1 YPLASL Late domain	Nat. Struct. Mol. Biol
	2RO3	ALIX in complex with the YPDL Late domain	
	2R02	ALIX in complex with the HIV-1 YPLTSL Late domain	
2008-06-10	3C3R	Bro 1-CHMP4C complex	Proc.Natl.Acad.Sci.
	3C3Q	Bro 1-CHMIP4B complex	
	3C3O	Bro 1-CHMIP4A complex	

Bro 1 domain

The overall crystal structures of unligand ALIX and ALIX in complex with different binding partners are very similar. The N-terminal Brol 1 domain is shaped like a banana and organized about a core tetratricopeptide helical hairpin (Figure 26A). There are three helix hairpins involved in the core tetratricopeptide repeat (α4/α5, α6/α7 and α8/α9), each of which comprises about 50 residues. The core is flanked by α10 and a small β sheet (β1, β2) at its N-terminal site and C-terminal site, respectively. The first three helices are separated into a three-helix bundle. N- and C-terminus are both adopted extended conformation, traversing the opposite sides of the domain. The overall structure of Human Bro1 domain resembles its yeast counterpart, albeit Human Bro1 C-terminus is terminated earlier. In contrast, yeast Bro 1 counterpart forms consecutively three helical segments and turns up into the domain [91].

It is reported that there are two exposed hydrophobic patches (Patch 1: convex surface and Patch 2: concave surface) on the surface responsible for association with ALIX ligands [83, 91]. Patch 1 is located in the middle of Bro 1 domain, which contains three hydrophobic pockets between α5，α6 and α7 (here denoted P1, P2 and P3 respectively) (Figure 26B, C). Patch 1 consists of non-linear residues topologically centered on F199 (sited in α6). This three-dimensional structure motif is the docking site that interacts with the ESCRT-III component CHMP4. Analysis of the complex structures of Brol-CHMP4 peptides reveals important interactions in all complexes are made by the three hydrophobic residues located in three successive turns of CHMP4 peptides (W220, L217 and L214). W220 is inserted into P1 and associates with D143 and K147 (both sited in α5) in P1 by hydrogen bonds. L217 from CHMP4 peptides occupies P2 and L214 is related to P3. Mutagenesis experiments corroborate the structural results. Besides, D143 and K202 residues (sited in α6) in Bro 1 domain forms a salt bridge that is absolutely conserved as an acidic/basic residue pair in different Bro 1 domains. Patch 2 is located on the top of Bro 1 domain and contains a conserved Tyr319 residue (sited in the end of α10) which is identified to be phosphorylated and associate with Src SH2 domain.



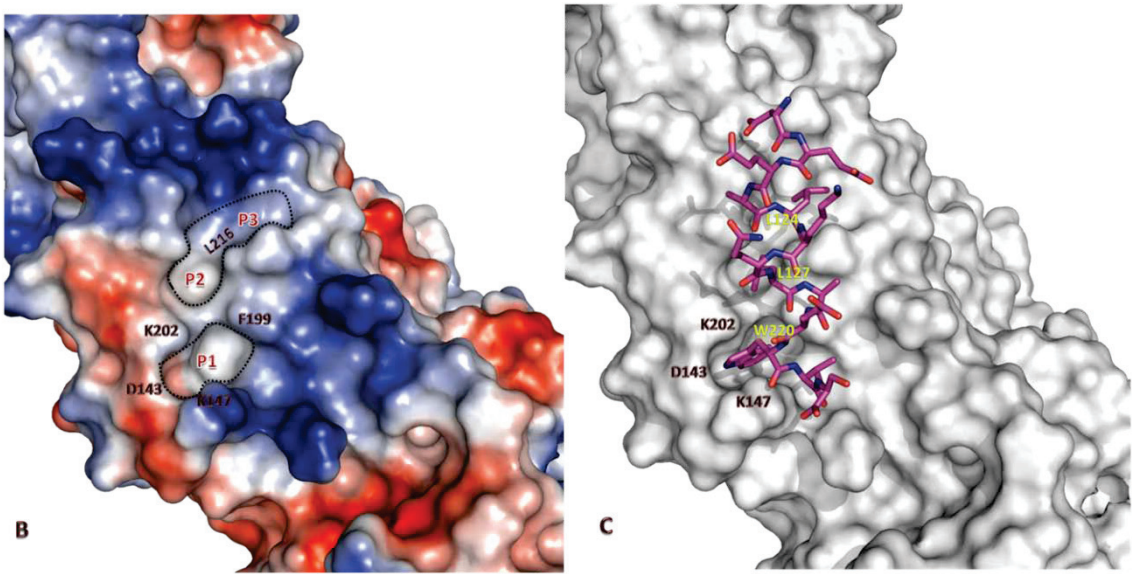


Figure 26.(A)Ribbon representation of $ALIX_{Bro\ 1}$ in complex with C-terminal helix from CHMP4A (purple) (PDB entry 3C3O). Secondary structures are shown here [90]; (B) The surface rendering shows the binding sites of CHMP4A, which are denoted as P1, P2 and P3 here. The key residues involved in protein-ligand interaction are hightlighted; (C) CHMP4A (purple) interacting with Bro 1 domain. The residues in CHMP4A of W220, L127 and L124 are highlighted in yellow.

ALIX V domain

Previous studies have demonstrated that over-expressed ALIX-V domain was inhibited virus budding, and blocking the binding of ALIX-V domain to virus could reverse the inhibition [80, 92]. Combination of the crystal structures of unbound or bound V domain and mutagenesis elucidates the virus budding mechanism regulated by ALIX [82, 88-91]. For instance, Lee *et al* showed mutations of Phe676 and V509 in ALIX-V domain completely abolishe or greatly decrease the function of V domain and promote virus budding. In contrast, mutant of C512A has exhibited a slight increase in the affinity between viral late-domain protein and ALIX-V domain, and the capability to inhibit virus budding [89]. The crystal structure analysis are consistent with the mutagenesis studies. The central V-shaped domain contains eleven helices with connecting loops, forming into two extended three-helix bundles denoted as arm1 and arm2 respectively. The connecting loops cross three times between the two arms of the ‘V’, and stabilize the two bundles by hydrophilic interaction between

helical component residues (R649) and flexible loop residues (D407, T412, P535) (Figure 27) [91]. There is a deep and extensive hydrophobic crevice on the arm2 surface facing to arm1 which stabilizes the V domain packing in the crystal lattice space and also acts as the ligand binding site of viral late-domains. In the crystal packing arrangement, the loop between α14 and α15 is reported to protrude into the hydrophobic pocket [89]. The hydrophobic crevice is formed by the conserved residues (Figure 28). Of them, the conserved residue Phe676 locating in the hydrophobic pocket is critical for binding to viral late domains (YPXnL containing ligands) [82, 89, 91]. The structures of V domain in complex with YPXnL containing peptides from HIV-1 and EIAV show both of the peptides bind to the same ALIX V domain surface and form similar interactions (Figure 29A and B). Mutation and crystallography analysis both demonstrate the residues of V498, V509 and Phe676 on the arm2 conserved hydrophobic pocket are vital for the viral late domain binding. The late domain Tyr in **Y**PXnL motif is essential for the virus budding function. Its side chain projects into a hydrophobic pocket consisting of Ala502, Ala505, Asp506, Val509, Glu672, Gly673, Phe676 and Tyr677 on the arm2 hydrophobic crevice, and forms a hydrogen bond with Asp506 directly (Figure 29C). And the hydrogen bond formed between the side chain of tyrosine in **Y**PXnL and Asp506 in ALIX enhances specificity and affinity between viral later domain and ALIX. Besides, the Asp506 also associates with Tyr677 whether ligand bound or unbound (Figure 29C). In addition, the NH group of tyrosine in **Y**PXnL also forms a hydrogen bond with ALIX at the conserved side chain of Glu672. Besides tyrosine residue in **Y**PXnL, another conserved residue P of the late-domain proline (Y**P**XnL) packs against Phe676, Ala502 and Val498. The different late domain peptides decided by n perform quite different conformation at the same binding site between the conserved residue Pro and Leu (Y**P**Xn**L**). For example, EIVA late domain peptide is extended, but HIV-1 forms a helical turn, which allows them to make equivalent contacts with ALIX (Figure 29A and B). It is reported that protein context modulates ALIX-late domain interaction. The conserved resides leucine and proline flanking this tyrosine in late-domain sequences (***L***Y**P**XnL) are also make equivalent hydrophobic contacts with ALIX.

Above all, Phe676 and V509 are located in viral late domain binding pocket and play key roles for association with the YPXnL motif, so their mutation affect the ability of V domain to disrupt budding. In addition, Cys512 is located at the end of the hydrophobic pocket closest to the hinge region of the V, and is highly solvent exposed which maybe facilitate Cys512 to post-translational modifications so that it could interfere with a partner binding [89].

However, the previous observation by mutation and crystallography could not explained the $ALIX_{V+PRR}$ and $ALIX_{V}$ induced the dominant negative effects on retroviral budding by the expression of truncated ALIX constructs until Pires *et al* showed that ALIX performed dimerization *in vitro* and *in vivo via* the V domain [93]. In their studies, the SAXS analysis demonstrated that mutation of $ALIX_{V}$ within the regions of 638-**KMK**QSNNE-645 located in the dimerization surface of V domain prevents dimerization and results in an elongated opening form which stabilizes the open monomer conformation. Conversely, the dimeric form of the wild-type recombinants of $ALIX_{Bro1-V}$ displayed an elongated crescent-shaped structure in SAXS modeling. Based on the SAXS analysis, they confirm the V-domain is sufficient for dimerization (Figure 30). Furthermore, their data suggest that ALIX dimerization is the active form that interacts with CHMP4 and functions an adaptor protein in endosomal sorting processes. However, the truncated forms of ALIX that retain the V-dimerization domain are prone to form hetero-dimers with endogenous ALIX. The mixed dimeric conformations of ALIX may block adaptor protein binding, which regulates the ALIX complex in the correct endosomal sorting pathways.

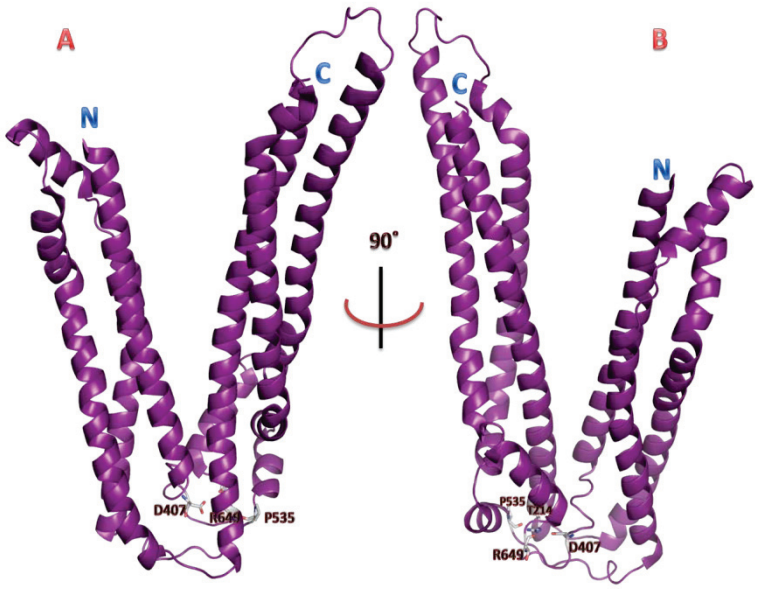


Figure 27. Overall fold of the ALIX V domain. The residues involved in the packing interactions between two arms are shown, which stabilize the conformation of the V domain [91].

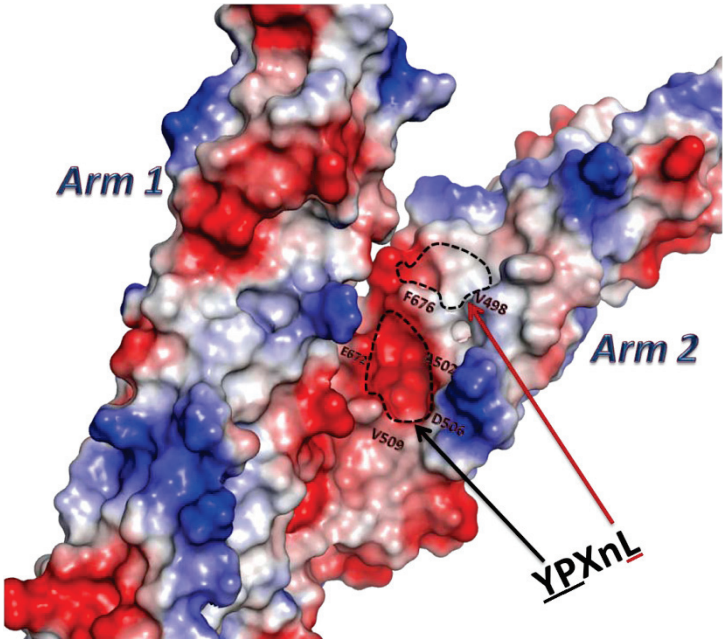


Figure 28 Identification of the conserved hydrophobic pocket on ALIX V domain. The hydrophobic pocket is surrounded by some key residues shown here. The pocket under reside F676 is occupied by residues Y and P of the Gag late domain sequence YPXnL, and the one adjacent to residue F676 is occupied by residue L in YPXnL [89].

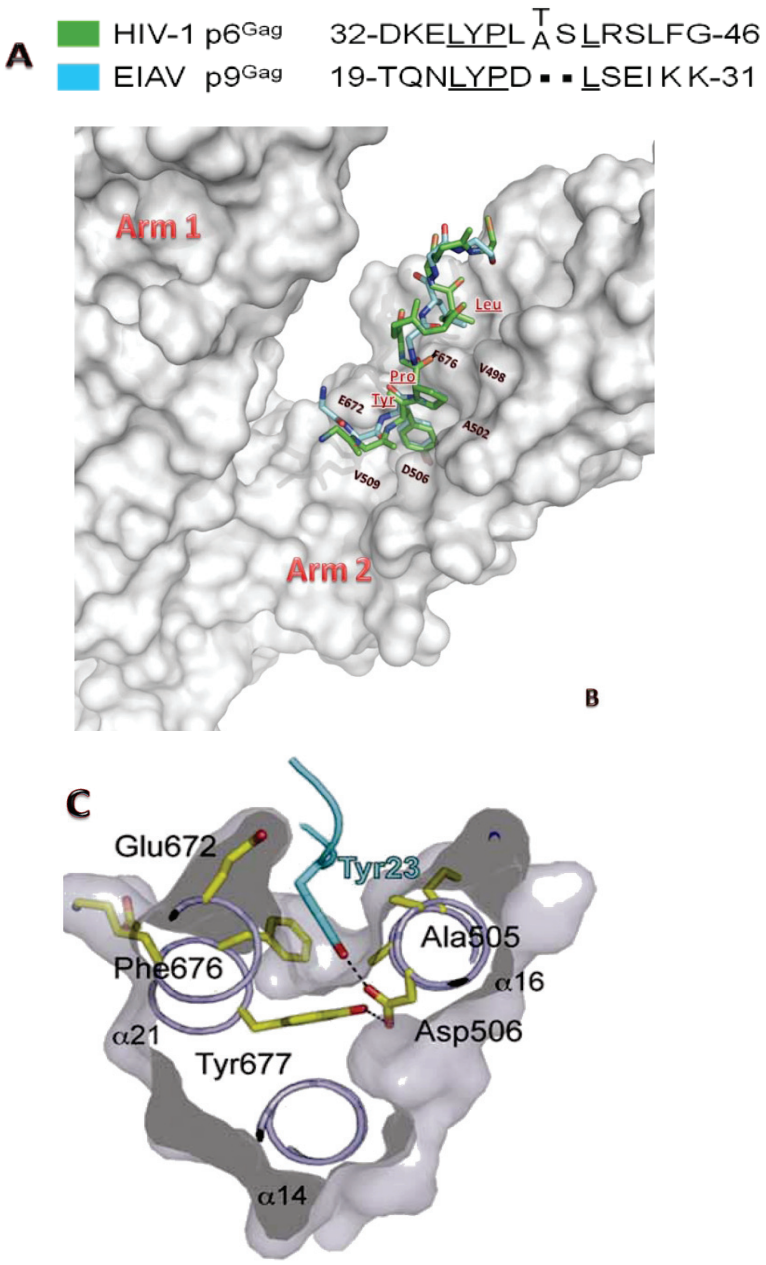


Figure 29 (A) sequences of HIV-1 and EIAV late-domain peptides; (B) Interaction between viral late domain peptides and ALIX V domain.; (C) close-up view of the interaction between the conserved late-domain Tyrosine and ALIX.

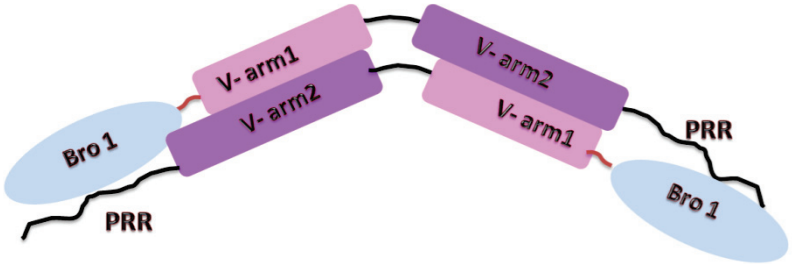


Figure 30. Model of dimeric ALIX. Dimerization requires that $ALIX_V$ opens and allows antiparallel interaction of the ALIX V-domain arms.

ALIX Bro 1-V domain

Recently, Zhou *et al* showed that the native state of ALIX is inactive and has no capacity to perform its cellular functions. They demonstrated that the ligand-binding sites in Bro 1 and V domains were all unavailable in the native state of ALIX. These masked sites could be unmasked by detergents and then exposed the Patch1/2 docking sites and F676 pocket to remove auto-inhibition of ALIX. The auto-inhibition of ALIX was speculated on the basis of intra-molecular interactions regulated by PRR, which masked the partner binding sites in both Bro 1 and V domain (Figure 31). Above all, it predicts that there should to be an activation step to remove the auto-inhibition conformation before ALIX performs its cellular functions.

As illustrated in Figure 32, the $ALIX_{Bro\ 1\text{-}V}$ adopts an extended conformation. There are three linker residues ($_{359}$VPV$_{361}$) between Bro 1 and V domain, which can change their relative orientation upon their partner binding or other environmental variation. Moreover, the conformation of the middle V domain brings the N-terminal Bro 1 domain and the C-terminal PRR into spatial proximity. Therefore, for one hand, Src kinase can simultaneously bind to the Bro 1 domain and PRR by SH2 and SH3 domain respectively and phosphorylate Tyrosine residues in ALIX C-terminus; on the other hand, the flexibility maybe mediates the intra-molecular interaction and then inhibits the ALIX activity. So association with either Bro 1 domain or PRR domain will release inactive conformation of ALIX. Of them, phosphorylation of Tyr319 is reported to be necessary for interaction with intact Src and activates ALIX as cellular regulators. This is another example to prove that protein phosphorylation and partner

protein binding are common mechanisms to relieve the auto-inhibitory and intra-molecular interaction of a protein.

Based on these data, Piere *et al* proposed a dimerization model: first, displacement of the PRR serves as the first signal to release auto-inhibition; second, an opening of the V-domain and subsequent antiparallel dimerizaiton via the V-domain. Finally, the dimeric ALIX stabilizes the active form of ALIX and allows it to interact with CHMP4 subunits and recruit of a variety of effectors via its PRR.

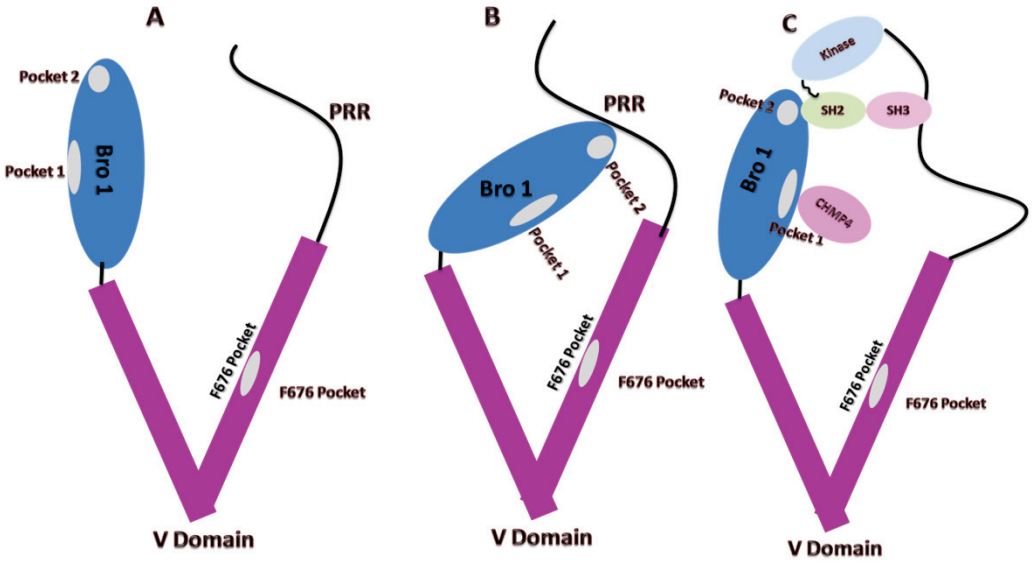


Figure 31.Illustration of the open-conformation models of ALIX domains (A). (B) The inhibition models; (C) association with Src kinase. The patches and pocket on these domains [83]. Patch1 is the docking site for CHMP4 proteins regulating ESCRT pathway. Patch 2 is the binding site for Src kinase SH2 domain by phosphorylation at Tyr319. F676 pocket is related to virus budding pathway. (A) .

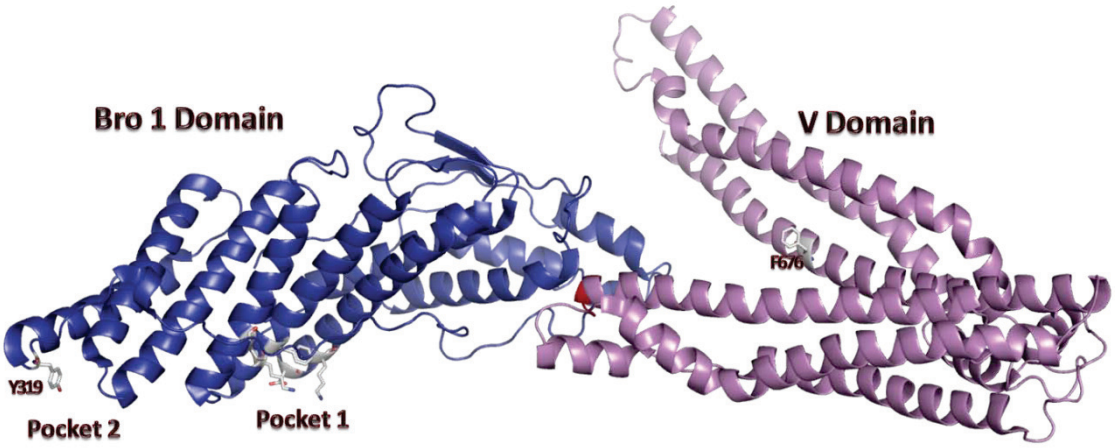


Figure 32. Overall structure of $ALIX_{Bro\ 1\text{-}V}$ (PDB entry 2OEV). Bro 1 and V domain are presented in blue and violet, respectively. Residues implicated in CHMP4 binding (pocket 1, D143/K147/K202/F199/L216), Src SH2 domain binding (pocket 2, Y319) and Gag YPXnL motif binding (F676 pocket) are highlighted.

Despite the great interest and recent advances concerning the biology of ALIX, the structural information of C-terminal PRR is still unclear. It has been well known that the PRR is involved in multiple cellular processes via its polyproline sequences associating the corresponding partners. So it is still mysterious about the mechanisms that how the polyproline regions of ALIX select their binding partners and how they function? Does the ALIX possess of the similar selectivity and specificity as HIV-1 Nef toward SH3 domain? Is the interaction between ALIX and SFK SH3 domains is biologically relevant in *vivo*? What is the difference at the interface between ALIX-SH3 and HIV-1 Nef-SH3? To answer all these questions, we therefore designed a series of biological studies to elucidate the interaction between ALIX and various SFK SH3 domains

Part II Materials and Methods

5. Experimental Materials

5.1 Host Strains and Vectors

All strains and plasmids used for cloning and expression were from Novagen company, such as *E.coli* Top10F, BL21(DE3) and pET 42a. Competent cells were prepared for transformation according to the pET system manual. The modified expression vector of pET11d-Sumostar was preserved in our Chinese lab (FJIRSM). The plasmids of pGEX-Hck, pGEX –$Fyn_{(R96I)}$, pGEX –$Fyn_{(WT)}$ and pGEX –$Fyn_{(R96W)}$ were the kindly gifts from our collaborator Dr. Yves Collette.

5.2 Enzyme, Chemicals and Experimental Kits

All enzymes were purchased from TaKaRa except that *BseRI* was from New England Bialabs and *Pfu* DNA polymerase was from Sangong Shanghai. Other chemicals used were of analytical grade and purchased from Shanghai Reagents Inc, China or Sigma-Aldrich. Ni-NTA resin was from Invitrogen. Superdex75 HR 10/30 and 10/26 size exclusion columns were from Amersham Pharmacia Biotech.

5.3 Growth Media and Stock solutions

LB media was used for plasmid production and protein expression. M9 medium was used for expression of the ^{15}N labeled SH3 proteins for NMR HSQC. Ampicillin at 100 ug/ml and kanamycin at 30 ug/ml were used for colony selection.

⑴ LB (Per liter):

10 g Tryptone;

5 g Yeast extract;

10g NaCl;

Adjust pH to 7.5 with 1N NaOH , Autoclave

⑵ 10×M9 salts (Per liter):

60 g Na_2HPO4

30 g KH_2PO4

5 g NaCl

Autoclave.

⑶ M9 (Per liter):

100 ml 10×M9 salts;

1ml 1M MgSO4;

1ml 0.1M CaCl2;

1ml thiamine HCl;

10 ml 20% Glucose;

1g NH4Cl

900ml ddH_2O

⑷ 1.0 M IPTG (isopropyl-β-D-thiogalactopyranoside): 23.8 g IPTG in 100 ml deionized water. Filter sterilize and store at -20℃.

⑸ Ampicillin (sodium salt): 100mg/ml in deionized water. Filter sterilize and store at -20℃.

⑹ Kanamycin (sulfate): 30 mg/ml in deionized water. Filter sterilize and store at -20℃.

5.4 Oligo primers

All the primers were synthesized by Shenggong shanghai company.

① ALIX constructs primers

362 sen: 5 ‘TCAGTACAGCAGTCTTTGGCTGCCTATAATC ’3

Mid AS: 5‘797 TTAAGAAGTCCCTCCTGTTTCTTTAGAGATTC ’3 766

Mid sen: 5‘762 CCAAGAATCTCTAAAGAAACAGGAGGGACTTC’3 793

1211AS: 5‘1235 GCAGTAGCAGAAGGAGCTCGATTTG ’3 1211

$ALIX_{V+PRR}$ Sense: 5‘AGATTGGTGGC TCAGTACAGCAGTCTTTGGC ’3

$ALIX_{V+PRR}$ Antisense: 5‘GAGGAGAGTTTAGAC TGGTGGAGGCCTGG’3

$ALIX_{V}$ antisense: 5‘GAGGAGAGTTTAGAC TCTTTCTGTCTTCCGTGC ’3

② C-termianl His_6 SH3 constructs primers

Hck-Sense2 : 5‘GGAATTCCATATGATCGTGGTTGCCCTGTATG ’3

Hck-Antisense2: 5‘CCGCTCGAGGTCAACGCGGGCGACATAG ’3

Fyn sense : 5‘GGAATTCCATATGACAGGAGTGACACTCTTTGTGG’3

Fyn antisense : 5‘CCGCTCGAGGATAGAGTCAACTGGAGCCAC ’3

6. Experimental Procedures

6.1 DNA Construction

6.1.1 Sumo fusion ALIX recombinants

ALIX cDNA spanning from the V domain to the Proline Rich Region (PRR) was amplified from human fetal liver library as template by three-round PCR. First, the two overlapping gene fragments, corresponding to residues 362-621 and residues 615-774 of the full-length ALIX protein, were amplified with the primers 362 sen & Mid AS and Mid sen & 1211AS, respectively. Then both PCR products were purified and mixed at 1:1 equivalent as templates to amplify the final ALIX gene consisting of residues 362-774 with the primes 362 sen and 1211AS.

A sticky-end PCR strategy was applied against the ALIX gene of residues 362-774 in order to obtain recombinants gene encoding i) $ALIX_V$ (residues 362-702) and ii) $ALIX_{V+PRR}$ (residues 362-760), respectively. Both of the PCR products were digested with two universal cloning sites (5'*NdeI* and 3'*BamH1* restriction sites) and permitted to be directionally sub-cloned into the modified pET11d-Sumostar expression vector using the two restriction enzymes, generating pET11d-Sumo-ALIX recombinants. Transformation was performed according to standard techniques. Both clones were verified by sequencing.

6.1.2 C-terminal His_6 -SH3 Recombinants

For SPR (surface Plasmon resonance) and GST (glutathione S-transferase) pull down experiments, GST and the various GST-SH3 recombinant proteins were used, including GST-((SH3-Hck, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3)). They were provided by our collaborator Dr. Yves Collette.

To produce recombinant proteins of the different recombinant SH3 domains and simplify the purification procedure for them, cDNAs encoding the various recombinant SH3 proteins (Hck-SH3, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3) fused with His tag were amplified by PCR using the corresponding pGEX-SH3 constructs (pGEX-Hck, pGEX-FynR96I, pGEX-Fyn and pGEX-FynR96W) as templates . Each amplified DNA fragment was sub-cloned into an expression vector,

pET42a using 5'*NdeI* and 3'*XhoI* restriction sites. All constructs were transformed into *E.coli* Top10 F' and verified by DNA sequencing.

6.2 Protein Expression and Purification

6.2.1 Sumo Fusion ALIX recombinants

Human ALIX recombinant constructs ($ALIX_V$ and $ALIX_{V+PRR}$) were both expressed in *E.coli* BL21 (DE3) with a His_6-Smt3 tag (designated as SUMO-tag) at their N-terminal. This tag can be removed by SUMO protease. A single colony was from a fresh LB-agar plate and inoculated into 10 ml LB containing the 100 ug/ml Ampicillin at 37°C overnight. This overnight culture was transfered to 1 liter LB medium and shaked at 37°C until the OD_{600} value reached 0.4-0.6. Cells were induced with 0.2 mM IPTG for 12 hours at 16°C. The cells were harvested by centrifugation at 5000 g for 10 min at 4°C. Harvested cells were resuspended in 20-30 ml lysis buffer (50 mM Tris-HCl pH8.0, 150 mM NaCl, 5 mM imidazole, 2 mM β-mercaptoethanol, 0.1 % *DNAase I*). After the cells were broken-up by French press or sonication, the supernatant was mixed with 5 ml Nickel-NTA resin (Qiagen) at room temperature for 2-3 hours. After extensive wash by lysis buffer, the captured SUMO-ALIX fusion proteins were incubated with 0.5 mg of SUMO-protease in cleavage buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM β-mercaptoethanol) for 1-2 hours at 30°C. Finally, the cleaved ALIX recombinant proteins were released from the Nickel-NTA column. The flow-through was collected, concentrated and applied to a Superdex 75 column (Amersham Biosciences) equilibrated in 50 mM Tris-HCl pH 8.0, 20 mM NaCl, 2 mM β-mercaptoethanol. Purified proteins were assessed by SDS-PAGE followed by Coomassie blue staining. The target proteins were concentrated using a Millipore Ultrafree concentrator with a 3kDa molecular-weight cutoff, and stored in -20°C for all kinds of measurement.

6.2.2 Sumo-protease

The shuttle vector of sumo-protease (pET28d-UPL) was constructed and preserved by our laboraty in FJIRSM. Sumo-protease was expressed with both N- and C-terminal His_6-tag. Cells were induced with 0.5 mM IPTG at 16 °C for overnight.

Harvested cells were resuspended in lysis buffer (50mM Tris-Hcl pH7.4, 300mM NaCl, 20mM imidazole, 2mM β-mercaptoethanol), then French pressed or sonicated. The soluble supernatant was loaded on the Ni-NTA column equilibrated by the lysis buffer. After extensive wash by wash buffer (50mM Tris-Hcl pH7.4, 300mM NaCl, 50mM imidazol, 2mM β-mercaptoethanol), the recombinant protein was eluted by 50mM Tris-Hcl pH7.4, 300mM NaCl, 300mM imidazole, 2mM β-mercaptoethanol. Finally, the captured sumo-protease by Ni-NTA was conserved in 50% glycerol and stored in -20 °C.

6.2.3 C-terminal His_6 -SH3 Recombinants

The various SH3 recombinant proteins (Hck-SH3, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3) were all expressed as C-terminal His_6 fusion proteins in *E.coli* BL21 (DE3). Cells were induced with IPTG (0.5 mM) and grew overnight at 20 ℃. Harvested cells were re-suspended in lysis buffer (50mM Tris-Hcl pH8.0, 150mM NaCl, 5mM imidazole, 2mM β-mercaptoethanol) and then French pressed or sonicated. The soluble fraction was loaded on Nickel-NTA column, and obtained by one-step elution with elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 500 mM imidazole, 2mM β-mercaptoethanol). The protein captured by Nickel-NTA column was concentrated and further purified by gel filtration chromatography on a Superdex 75 column. Purified proteins were assessed by SDS-PAGE and concentrated by Millipore Ultrafree concentrator with 3 kDa molecular-weight cutoff.

6.3 GST pull-down assay

Purified $ALIX_{V+PRR}$ (362-760) was pre-incubated at room temperature for 2 hours in 500 µl of reaction buffer (25 mM Hepes pH 7.8, 150 mM NaCl, 10 mM EDTA, 1 mM EGTA supplemented with 1 % triton X100) with Sepharose-coupled recombinant GST or GST-Hck SH3, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3 (2 µg per reaction). After three washes using reaction buffer, complexes were resolved by 12% SDS-PAGE, colored by commassie blue (Biorad) for 1 hour and destained in 40% methanol/10% acid acetic.

6.4 Isothermal Titration Calorimetry (ITC)

ITC can directly measure heat evolved or absorbed in liquid samples as a result of mixing precise amounts of reactants. Before operation, both samples and reference cells have to be cleaned extensively. Meanwhile, all the samples used in experiments should be dialyzed in the degassed assay buffer. Sample cell has been filled with the corresponding assay buffer. The assay buffer is 10 mM Hepes pH 7.5, 150 mM NaCl, 2 mM EDTA with (for $ALIX_{V+PRR}$) or without (for PxxP petides) 2 mM *β*-mercaptoethanol. After dialysis, the protein concentrations were measured using a Nanodrop 1000 (Thermo Scientific). Data were recorded by the MicroCal VP microcalorimeter. Peptides were dissolved directly in the corresponding assay buffer. The SH3 recombinant protein fused with its C-terminal His_6 (Hck-SH3, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3) at concentration of 600 μM filled the syringe and were injected into the measurement cell containing $ALIX_{V+PRR}$ at 50 μM. Titrations with Hck-SH3 and PxxP peptides were performed with the peptides (at 3.8 mM concentration) being injected into the cell containing Hck-SH3 (at 200 μM). All titrations were carried out at 25 °C, using 10 μl to 15 μl injections. Blank experiments were carried out by titration of the corresponding syringe ligands at the same concentration as the former into the assay buffer, and data were subtracted to assay data obtained for the interaction between SFKs SH3-$ALIX_{V+PRR}$ or Hck-SH3-peptides. All data were analyzed using the Microcal Origin Software.

6.5 Surface Plasmon Resonance (SPR)

SPR was used to measure the binding interaction and to evaluate the kinetic parameters between $ALIX_{V+PRR}$ (362-760) and the different SH3 domains (Hck-SH3, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3). The measurements were processed with a Biacore 2000 system at 25 °C with a flow rate of 20 μl/min. The CM5 chip was activated by a mixture of 1:4 molar ratio of N-Hydroxysuccinimide (NHS) /N-(3-Dimethylaminoproply)-N'-ethylcarbo-diimide (EDC) at a flow rate of 5μl/min for 7 min. After activation, anti-GST antibody was immobilized to the chips by coupling buffer. Then GST or GST-SH3 fusion proteins were flown over the anti-GST surface with running buffer, respectively. The recombinant GST was performed as a

negative control. Before injecting the binding partner of $ALIX_{V+PRR}$, it was important to wait until the response of injecting GST fusion protein reached stable.

The anti-GST antibody was diluted in coupling buffer (10 mM sodium acetate, pH 5.0) to 30 μg/ml and directly immobilized on the activated chip at a flow rate of 5μl/min, for 7 min. This coupling/immobilization was followed by a blocking step using ethanolamine (1 M, pH 8.5) at 5 μl/min, for 7 min. Then, recombinant GST and GST-SH3 fusion proteins were diluted in running buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005 % (v/v) tween) to 20 μg/ml and passed over separate flow cells over the anti-GST surface at 20 μl/min for 6 min in running buffer. Binding of $ALIX_{V+PRR}$ was measured in running buffer. Finally, surfaces were regenerated by injection of 10 mM glycine-HCl pH 2.5 at 20μl/min for 1 min. Experiments were repeated at different concentrations of $ALIX_{V+PRR}$ from 20 uM to 70uM diluted with running buffer.

6.6 Nuclear Magnetic Resonance spectroscopy (NMR)

To further elucidate the interactions between SH3 domains (Hck-SH3, $Fyn_{(WT)}$-SH3, $Fyn_{(R96I)}$-SH3 and $Fyn_{(R96W)}$-SH3) and $ALIX_{V+PRR}$ or PxxP peptides derived from ALIX PR region, ^{1}H-^{15}N HSQC experiments were recorded. Each H-N correlation peak is associated to the NH group of an amino acid. In our case, the measurements were subdivided into two groups: one is for comparison of the chemical shift perturbation on Hck-SH3 surfaces induced by the titrations of $ALIX_{V+PRR}$ or PxxP petides (group 1); the other is for comparison of the chemical shift perturbation on various SH3 domains induced by the titrations of $ALIX_{V+PRR}$ (group 2).

In group 1, mappings of the $ALIX_{V+PRR}$ or PxxP peptide interaction sites on Hck-SH3 surface were performed with a sample containing 50 μM of uniformly ^{15}N-labelled Hck-SH3 diluted in PBS buffer (50 μM PO4 pH 7.0, 150 μM NaCl, 5 mM DTT, 2.5 mM EDTA) containing 10% D_2O. Titrations were monitored by adding up different amouts of unlabeled $ALIX_{V+PRR}$ solution (1.45 mM) or PxxP peptide (3.3 mM) in the same buffer to variedly final molar ratios of $ALIX_{V+PRR}$: Hck (from 0.24

to 4.6) and PxxP peptide : Hck (from 0.5 to 10.2). In group 2, mappings of the $ALIX_{V+PRR}$ interaction sites on various SH3 domains surfaces were performed with a sample containing 80 µM of uniformly ^{15}N-labelled SH3 proteins diluted in buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 2mM *β*-mercaptoethanol) containing 10% D_2O. Unlabled $ALIX_{V+PRR}$ (180 µM) was added up to the final 1:1 molar ratio. At each step, a ^{1}H-^{15}N HSQC spectrum was recorded on a 600 MHz Bruker spectrometer (equipped with a cryoprobe). Heteronuclear assignment was taken from Horita *et al* [71, 94].

6.7 Small Angle X-ray Scattering Data Collection and Analysis

Data used for this SAXS analysis were collected at beamline SWING of the SOLEIL synchrotron, Paris, France, at 10°C, using a wavelength of λ=1Å.03 For measurement of $ALIX_{V+PRR}$ alone, $ALIX_{V+PRR}$ (20 mg/ml) was applied to a HPLC-coupled size exclusion column (S200, PHARMACIA) that output fractions directly into the SAXS measuring cell. The purification by HPLC was performed in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2mM *β*-mercaptoethanol. For $ALIX_{V+PRR}$ in complex with $Fyn_{(R96I)}$-SH3, the complex between $Fyn_{(R96I)}$ and $ALIX_{V+PRR}$ (51 mg/ml) at 1:1 molar ratio was pre-incubated on ice for several hours before data collection. This concentration corresponds to about 100 times the K_D of the interaction, and it can be assumed that > 99% of the molecules is in the complexed state. SAXS data sets for the fractions from the major peaks were checked for consistent radius of gyration (using the Guinier approximation as implemented in PRIMUS), scaled and merged. Data were collected for q-ranges of 0.0066 – 0.6000 $Å^{-1}$. Data collected on buffer alone were subtracted from all collected data sets. Data analysis, *ab initio* shape calculations, and modeling of the flexible C-terminal regions (not included in the ALIX V-domain crystallographic structure of PDB entry 2OJQ [89]) were performed using programs by S. Svergun and colleagues (PRIMUS, GNOM, DAMMIN, CRYSOL, DAMAVER, EOM and BUNCH) (please refer to the EMBL internet site for more details http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html). For BUNCH and EOM

analyses, the molecular structures of the $ALIX_V$ domain (PDB entry 2OJQ) and of the $Fyn_{(R96I)}$ mutant (PDB entry 3HOI) were used. The (canonical) interaction between $Fyn_{(R96I)}$ and HIV-Nef residues 71-78, taken from PDB entry 1EFN [67].

6.8 Crystallization, X-ray data collection and structure determination

Crystallization trials, ($ALIX_{V+PRR}$ alone, $ALIX_{V+PRR}$ in complex with $Fyn_{(R96I)}$, $Fyn_{(R96I)}$ in complex with various PxxP containing peptides derived from ALIX PRR), were all performed by the conventional sitting-drop or hanging–drop vapour diffusion at room temperature or 4 °C. The protein was mixed with an equal volume of the reservoir solution. The protein concentrations of $ALIX_{V+PRR}$ alone or mixed with ligands were varied from 5 mg/ml to 50 mg/ml. $ALIX_{V+PRR}$ was mixed with $Fyn_{(R96I)}$ at 1:1 molar ratio, and peptides were dissolved in 20mM Tris-Hcl pH 7.5, 100mM NaCl and added up to $Fyn_{(R96I)}$ solution at the final molar ratios varying from 1:3 to 1:10. Either home-made or commercial crystallization screen kits were tried extensively.

X-ray diffraction data were collected at 100 K on the APS SERCAT (Advance Photon Source, Southeastern Regional Collaborative Access Team) beamline 22-ID at the Argonne National Laboratory, Chicago, IL, U.S.A, and using a wavelength of λ=1.04 Å. The diffraction data of $Fyn_{(R96I)}$ were indexed and processed using the HKL2000 program (HKL research). Crystal structure of $Fyn_{(R96I)}$ was solved by molecular replacement program MOLREP of the CCP4 package [95, 96], using the coordinates of Fyn (PDB code 1EFN [67]) stripped of HIV-1 Nef as the model. Model building was performed with COOT [97]. Structure were refined with REFMAC, and TLS parameters were incorporated in the final refinement round [98, 99]. All the structure figures were generated using PyMol [100].

Part III Results and Discussion

7.1 DNA Construction

7.1.1 Sumo fusion ALIX recombinants

Previous studies have demonstrated that ALIX has a proline-rich region (PRR) in its C-terminus. This PRR is a likely candidate for the binding of SH3 domains or other cellular proteins and connects ALIX to various signaling processes. PRR is predicted to be mostly flexible, without a compact tertiary fold. Together with the previous pull-down experiments of ALIX-P749A/P752A/P755A with the GST-Src SH3 domain, it was suggested that mutation of the prolines in the Src-binding consensus sequence caused a reduction in the recovery of ALIX, whereas mutation of proline residues outside of this region showed no effect. It was also demonstrated that the $_{752}$PQPPAR$_{757}$ sequence in ALIX C-terminus is the most interesting motif which bound to Src SH3 domain and regulated the EGFR internalization. Based on these results, we decided to make ALIX construction spanning from V-domain to residue 760, or beyond in PRR.

The project began with the extraction of the ALIX cDNA spanning from V domain to PRR from human fetal liver library. It was of great challenge at the beginning of cloning this gene, because the gene of PRR consists of high GC content, which tends to form hair-pin structures. Therefore, three-round PCRs were used to obtain an initial cDNA containing the canonical PxxP motif as a model (Residues 362-774) (Figure 33). Then, standard sticky-end PCR were performed for constructs of $ALIX_V$ (Residues 362-702) and $ALIX_{V+PRR}$ (Residues 362-760) (Figure 33 and 34). The final fragment of $ALIX_{V+PRR}$ still retains the most probable SH3-binding motif according to Schimdt *et al*. Both of PCR products without restriction digestion were directly sub-cloned into the prepared pET11d-Sumostar expression vector (Figure 35). Both clones were verified by sequencing.

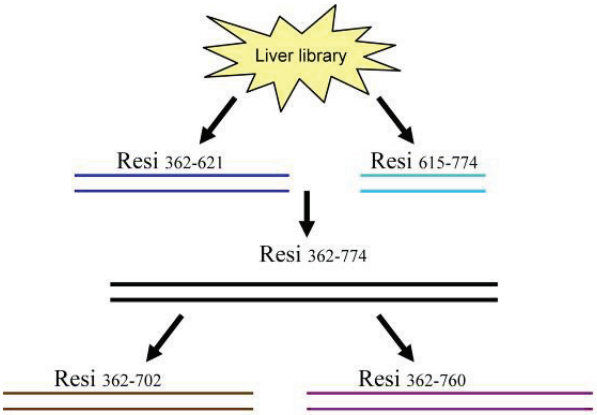


Figure 33 The sketch map of ALIX recombinants PCR

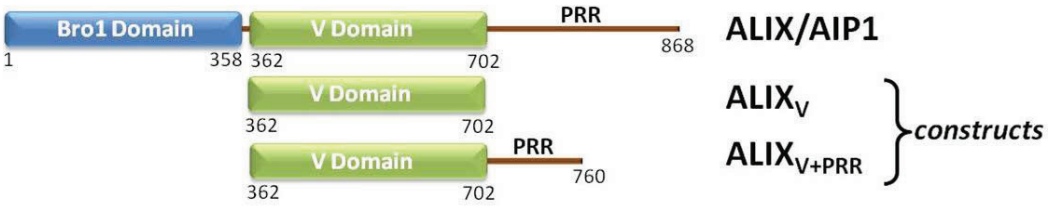


Figure 34 ALIX constructs. ALIX Bro 1 and V domains are respectively shown in blue and green. The different ALIX constructs are depicted in a schematic way.

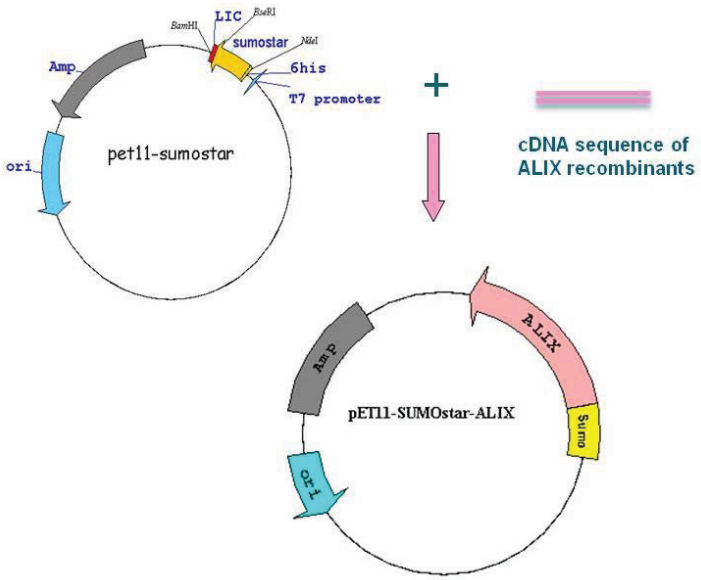


Figure 35 Construction of shuttle expression vector pET11-sumostar-ALIX recombinants

7.1.2 SH3 Recombinants with their C-terminal His_6

pET-42a vector was used to clone and express fusion SH3 proteins (Hck-SH3, $Fyn_{(WT)}$-SH3 and $Fyn_{(R96I)}$-SH3) with a C-terminal His_6 tag (Figure 36). The target gene fragments were sub-cloned between *5'NdeI* and *3'XhoI* restriction sites. Since *5'NdeI* is situated in promoter and contains ATG codon, fusion proteins were directly expressed with N-terminal only an extra Met residues. The C-terminal His_6 tag is not necessary to be removed.

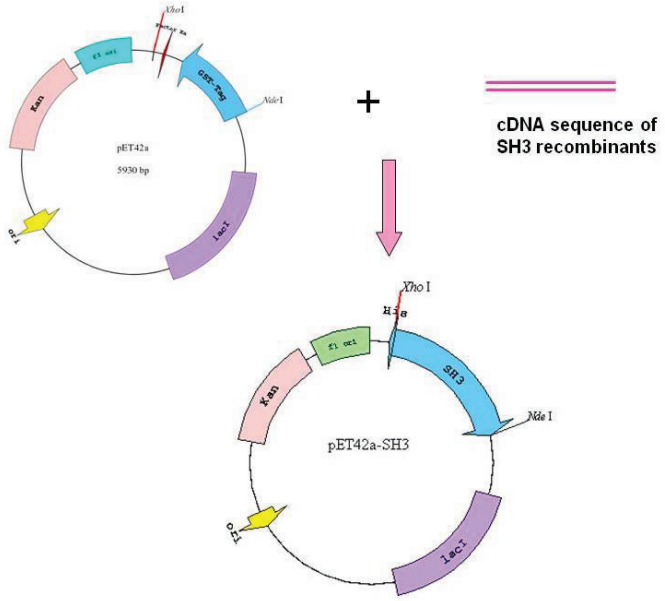


Figure 36 Construction of shuttle expression vector pET42a-SH3 recombinants

7.2 Protein Expression and Purification

7.2.1 SUMO-protease

We generated the SUMO-protease ourself. The SUMO-protease (His_6-Ulp-His_6) was expressed in *E.coli* BL21(DE3) and purified by one-step Ni-NTA column protocol (Figure 37). The protease tended to precipitate at high concentration during its elution from a Ni-NTA column. Increasing flow-rate prevents SUMO-protease from precipitating. Both N- and C- terminal His_6 tags facilitate the protease to be captured in Ni^{2+} resin. Eluted sumo-protease from Ni^{2+} column is preserved in 50% glycerol buffer immediately and stored in -20°C.

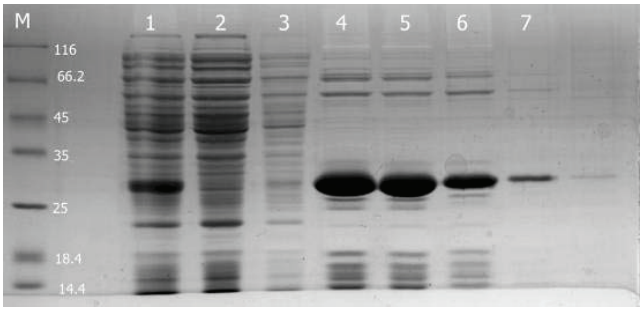


Figure 37 15% SDS–PAGE analysis of expressed SUMO-protease (His_6-Ulp-His_6) captured by Ni-NTA column. M: molecular standard protein; Lane 1: supernatant of lysate; lane2: Flowthrough; lane 3: wash solution; lane 4-7: elution by Ni-NTA buffer containing 200mM imidazole; lane 8: elution by Ni-NTA buffer containing 1.0 M imidazole

7.2.2 SUMO fusion ALIX recombinants

Previous studies reported that the ALIX recombinant with PRR was expressed poorly. In our case, human ALIX recombinant constructs ($ALIX_V$ and $ALIX_{V+PRR}$) are expressed as fusion with SUMO-tag protein in *E.coli* using modified customly-made pET11d-SUMOstar expression vector. This strategy significantly increases the yield of $ALIX_{V+PRR}$ (Figure 38 and 39). The initial fusion proteins are produced with their N-terminal 21 kDa SUMO-tag (His_6-tagged Smt3) shown on SDS-PAGE (Figure37). His_6-Smt3-ALIX recombinants were expressed at about 59kDa ($ALIX_V$) and 65kDa ($ALIX_{V+PRR}$) shown on SDS-PAGE, respectively (Figure 39). On-site cleavage was used to remove the SUMO tag from ALIX recombinants. After cleavage, both of SUMO tag protein (His_6-tagged Smt3) and excess SUMO protease (His_6-Ulp-His_6) were captured by Ni^{2+} resin and only the ALIX recombinant proteins with their fresh N-terminus were released from Ni^{2+} column (Figure 39). The one-step protocol facilitates the purification procedure to get pure and high quantity of ALIX recombinants. ALIX recombinants released from Ni^{2+} column were loaded on Superdex 75 column to remove the aggregated proteins (Figure 41). The final product was concentrated and stored at -80°C for further biochemical or biophysical studies (Figure 41).

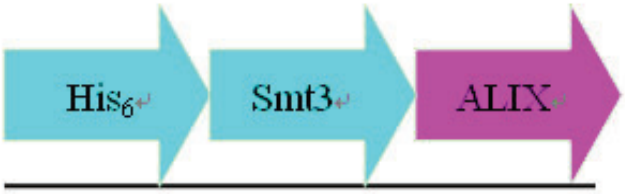


Figure 38.ALIX fusion protein denoted as His_6-Smt3-ALIX.

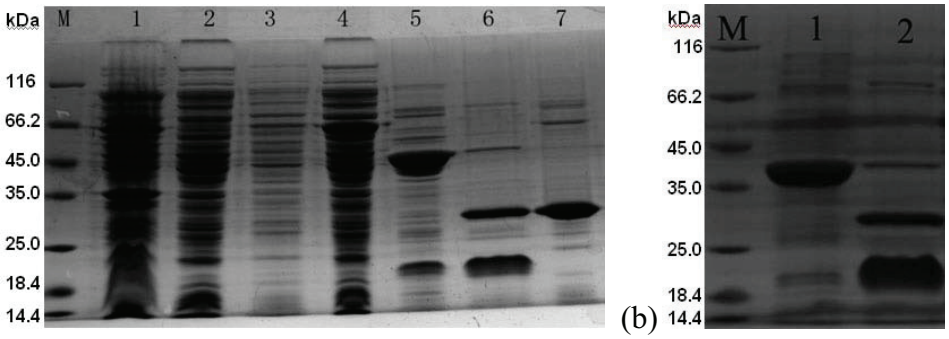


Figure 39.15% SDS–PAGE analysis of expressed ALIX recombinants purified by Ni^{2+} resins. (a) $ALIX_{V+PRR}$ (resi362-760): Lane 1, supernatant of lysate; lane 2, flowthrough; lane 3, wash solution, lane 4: protein after dialysis; lane 5, cleaved protein; Lane 6, elution by buffer comtaining 1M imidazole; lane 7, sumo-protease (b) $ALIX_V$ (resi362-702): lane 1, cleaved protein ; Lane 2: elution by buffer containing 1M imidazole;

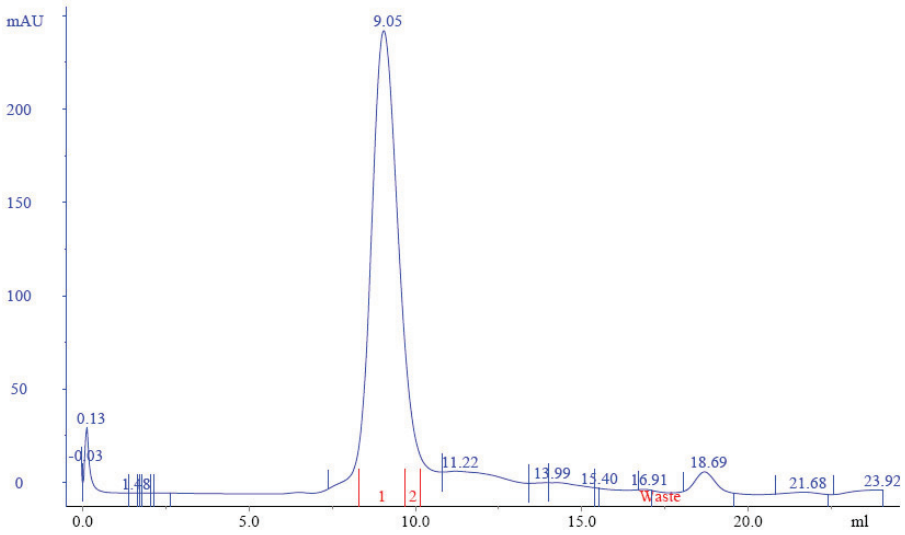


Figure 40.Gel-filtration chromatography analysis of $ALIX_{V+PRR}$ recombinant purified by Ni^{2+} column.

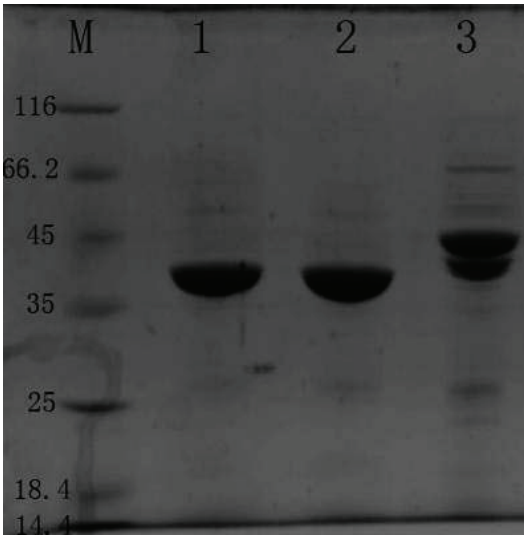


Figure 41.12% SDS-PAGE shows the purified ALIX recombinant proteins. Lane 1-2: $ALIX_V$; Lane 3: $ALIX_{V+PRR}$; $ALIX_{V+PRR}$ shows degradation on gel.

7.2.3 SH3 recombinants with their C-terminal His_6 tag

SH3 recombinants (Hck-SH3, $Fyn_{(WT)}$-SH3 and $Fyn_{(R96I)}$-SH3) with their C-terminal His_6 tag do not affect the interaction with ALIX recombinant and express very well in *E.coli* BL21(DE3) (Figure 42). SH3- His_6 recombinants were purified by one-step affinity column (Ni^{2+}-NTA) to remove most of the impure proteins without cleavage. The final proteins show well homogeneity after gel filtration chromatography Superdex 75 (Figure 43).

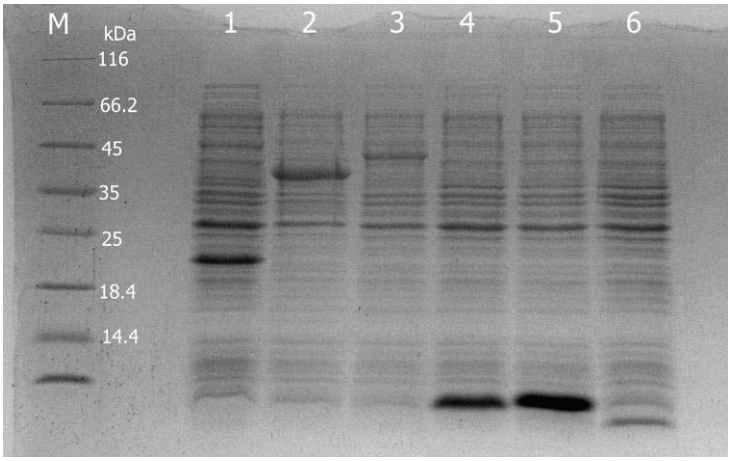


Figure 42.15% SDS–PAGE analysis of His_6-SH3 recombinants induced by 0.5mM ITTG at 37℃ for 3h. Lane 4: $Fyn_{(R96I)}$-SH3; Lane 5: $Fyn_{(WT)}$-SH3; Lane 6: Hck-SH3; M: standard protein molecular weight markers;

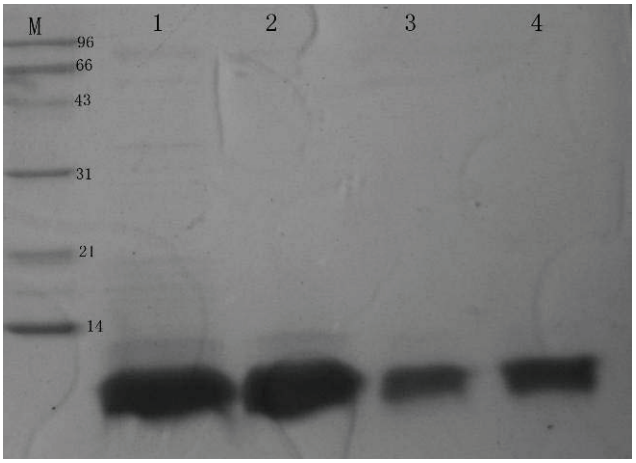


Figure 43.15% SDS–PAGE analysis of final His_6-SH3 recombinants. Lane 1: recombinant of $Fyn_{(R96I)}$-SH3; Lane 2: $Fyn_{(WT)}$-SH3; Lane 3: Hck-SH3; M: standard protein molecular weight markers;

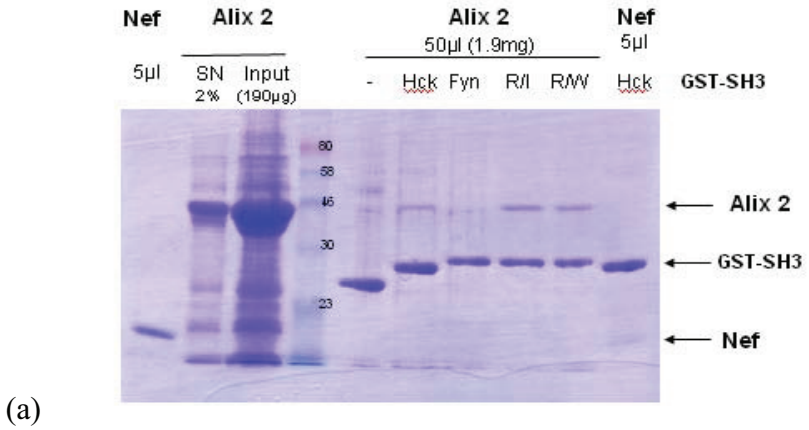
Identification of ALIX as a functional interactor of Hck and selectivity to SH3 domain by residue 96 in SH3 RT loop

Our collaborators have identified that ALIX is a novel potential Hck-SH3 binding protein by yeast two-hybrid screening approach using the Hck SH3 domain as bait (Opi and Shi *et al*, manuscript preparation). To confirm the interaction between full length ALIX and Hck proteins, co-immunoprecipitation and kinase assay *in cellular*, and binding assays *in vitro* were both performed.

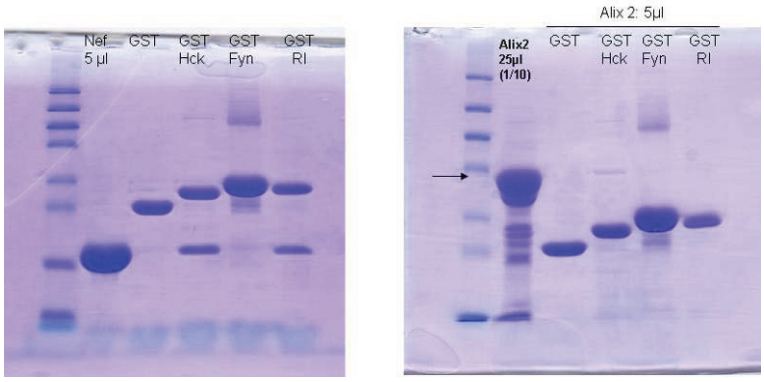
In vivo, $ALIX_{V+PRR}$ –Hck SH3 were verified to co-expressed and interact in intact mammalian cells culture by two-hybrid assay, moreover Hck kinase activity was increased in presence of ALIX [70]. In addition, two independent yeast two hybrid screen were performed in parallel using the human Hck- or $Fyn_{(R96I)}$-SH3 domain as a bait respectively. The results showed that ALIX can interact much better with Ile96 containing SH3 domain of Hck SH3 and $Fyn_{(R96I)}$-SH3 but poorly with Fyn-SH3. It strongly supports the notion that the selectivity of ALIX in cell was strictly dependent on the corresponding RT-loop residue 96 (Opi & Shi *et al* manuscripture). More importantly, the results from the yeast-two hybrid experiments show ALIX is similar to HIV-1 Nef protein in specific and selective interaction with cellular signal proteins SFKs.

7.3 GST pull-down Assay

Isolated GST (Hck, $Fyn_{(R96I)}$ and $Fyn_{(R96W)}$)-SH3 were able to recover $ALIX_{V+PRR}$ (Figure 44) but not $ALIX_V$ (data not shown here) recombinant from lysate of transfected cells. The pull-down experiment corroborates with yeast-two hybrid screen. In addition, the interaction between $ALIX_{V+PRR}$ and Fyn SH3 (an arginine 96 in the RT-loop) is very weak, and the ones between $ALIX_{V+PRR}$ and Hck or Fyn mutants ($Fyn_{(R96I)}$ and $Fyn_{(R96W)}$) are much higher, which has the substitution of this Arg96 by an hydrophobic amino acid. Comparing the interactions between SH3 domains of $ALIX_{V+PRR}$ and that of HIV-1 Nef, it appears that the interaction between HIV-1 Nef and SH3 is much stronger than ALIX's at the almost same concentration. This result is consistent with the potent role of HIV-1 Nef function in interaction with SH3 domains.



(a)



(b)

Figure 44.GST pull-down assays show the interactions between $ALIX_{V+PRR}$ and GST SH3- (Hck, $Fyn_{(WT)}$, $Fyn_{(R96I)}$ and $Fyn_{(R96W)}$), compared with that between HIV-1 Nef and GST SH3 recombinants.

7.4 Isothermal Titration Calorimeter (ITC)

To understand the determinants of specificity and selectivity of different (Hck, $Fyn_{(WT)}$, $Fyn_{(R96I)}$ and $Fyn_{(R96W)}$)-SH3 binding to $ALIX_{V+PRR}$ or peptides, ITC was performed to evaluate thermodynamic parameters (Table 5). The different peptides derived from ALIX PRR were named PI (residues 737-760), PII (residues 748-760) and PIII (residues 737-746), of which PI and PII are canonical PxxP containing sequences (Figure 45). As shown in the top panel of Figure 46, all titration between SH3-$ALIX_{V+PRR}$ or SH3-peptides results in exothermic reaction. Binding curves of SH3-$ALIX_{V+PRR}$ or SH3-peptides are presented as sigmoid plots and could be fitted with simple binding isotherm with its midpoint occurring at a about ratio of 1:1. The dissociation constants reveal that under the used conditions, $Fyn_{(WT)}$-SH3 interacted very weakly with $ALIX_{V+PRR}$ (dissociation constant K_D of 700±240 uM). However, in the case of Fyn mutant ($Fyn_{(R96I)}$-SH3, $Fyn_{(R96W)}$-SH3), the single amino acid mutation at position 96 was enough to promote an interaction with a K_D of 10.3 ±3.2 μM and 13.1±1.4 uM (>50 fold) respectively and comparable to that of Hck-SH3 (K_D of 34.5 ±3.3 μM), which is consistent with published results of SH3 selectively binding to full-length HIV-1 Nef (K_D >20 μM for $Fyn_{(WT)}$-SH3; K_D of 0.25 μM for Hck-SH3; K_D of 0.38 μM for $Fyn_{(R96I)}$-SH3)[39] and the HIV-1 Nefcore (Table 6, unpublished data from Stefan Arode).

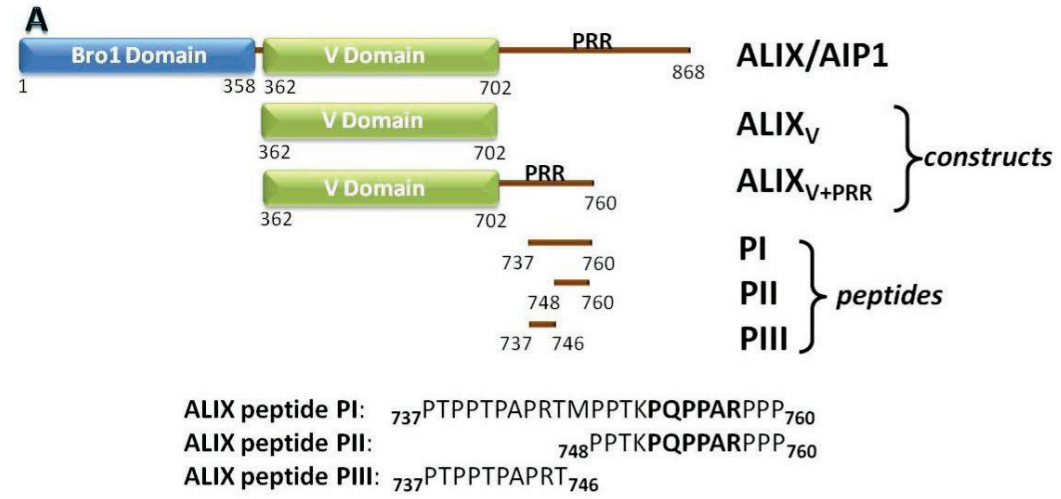
Further analysis of the thermodynamic data from SH3-$ALIX_{V+PRR}$ showed that $Fyn_{(WT)}$-SH3 has an unfavorable ΔS (negative) and the most favorable ΔH (negative), whereas the other three SH3 domains all provide the favorable ΔH (negative) and TΔS (positive) contribution to ΔG. The ITC binding data (Table 5) showed the interactions between $ALIX_{V+PRR}$ and SH3 domain of Hck and Fyn mutants are all entropic (ΔS) driven, which is consistent with most interaction between canonical PxxP and SH3 domains but converse to the association between HIV-1 Nef and various SH3 domains. Conversely, the one mediated by $Fyn_{(WT)}$-SH3 is enthalpy driven. The overall favorable entropic contribution to binding in all cases is likely to arise from the dominant effect of the burial of hydrophobic interface area, since on

forming the complex water molecules which are ordered on this surface are released into the bulk solvent. An entropy-driven association is expected for a basically hydrophobic interaction. However, previous studies have elucidated that the interaction between SH3 domains (Hck- and Src-) and HIV-1 Nef are enthalpy (ΔH) driven and entropic (ΔS) penalty, in spite of the hydrophobic PPIs interface. This means that SH3-$ALIX_{V+PRR}$ interactions thermodynamically differ from HIV-1 Nef-SH3 interactions. In spite of these subtle difference among $Fyn_{(R96I)}$-SH3, $Fyn_{(R96W)}$-SH3 and Hck-SH3, $Fyn_{(R96I)}$/$Fyn_{(R96W)}$-SH3 have the more favorable TΔS (27.4 KJ/mol/23.5 KJ/mol) and the less favorable ΔH (-1.0±0.2 KJ/mol/-4.4±0.2 KJ/mol), while Hck-SH3 has the least of favorable TΔS (10.4 KJ/mol) and the most favorable ΔH (-15.1±1.2 KJ/mol). So based on the favorable TΔS driven contribution to interaction between the two proteins, the interaction between $ALIX_{V+PRR}$ and SH3 domains are $Fyn_{(R96I)}$/$Fyn_{(R96W)}$-SH3～Hck-SH3 > $Fyn_{(WT)}$-SH3, which is consistent with that HIV-1 Nef_{core} and SH3 domains (Table 6).

The PxxP-containing peptides derived from ALIX PRR also bound to Hck-SH3. Moreover, PII, which possesses a canonical class II PxxPx(R/K) core binding motif (Figure 45B), demonstrated a comparable K_D, but altered ΔH and TΔS values as compared to PI (Table 5, Figure 46). The binding affinity of PI is almost identical to the one of Hck-SH3, but its thermodynamic parameters are quite different. The interaction between the SH3 domain from Hck and peptides are all with favorable ΔH contribution to ΔG, but varied ΔS. The interaction between PI and Hck-SH3 is entropic (ΔS) driven, which is consistent with other PxxP-SH3 interaction. Conversely, the interaction with PII is enthalpy (ΔH) driven with entropic (ΔS) penalty, the same as the HIV-1 Nef. PIII does not possess a canonical SH3 binding motif and has very weak interaction with Hck-SH3. The differences between PI and PII might be explained by a more stable PPII-helix of the SH3 binding region in the PI construct. This would be consistent with the observed more favorable TΔS in the PI interaction. However we cannot exclude that additional contact through regions outside PI to Hck-SH3 also contributed to the binding.

Together, these observations demonstrate that ALIX has a similarly selective

association with SH3 domains, but adopts a different recognition mechanism with HIV-1 Nef to various SH3 domains.



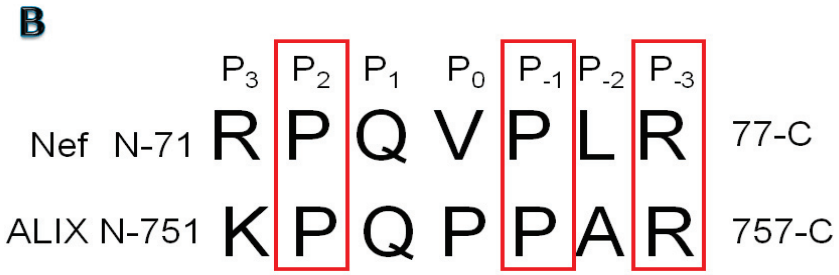


Figure 45.(A) PxxP peptide constructs compared with ALIX sequence. ALIX Bro 1 and V domains are respectively shown in blue and green. The different peptides used in the present study are depicted in a schematic way. (B) PxxP core motif alignment between Nef and ALIX.

Table 5 PxxP peptides binding to Hck-SH3 (Row 1-3) and $ALIX_{V+PRR}$ binding to Fyn--, $Fyn_{(R96I)}$-, $Fyn_{(R96W)}$-and Hck-SH3 (Row 4-7) measured by ITC

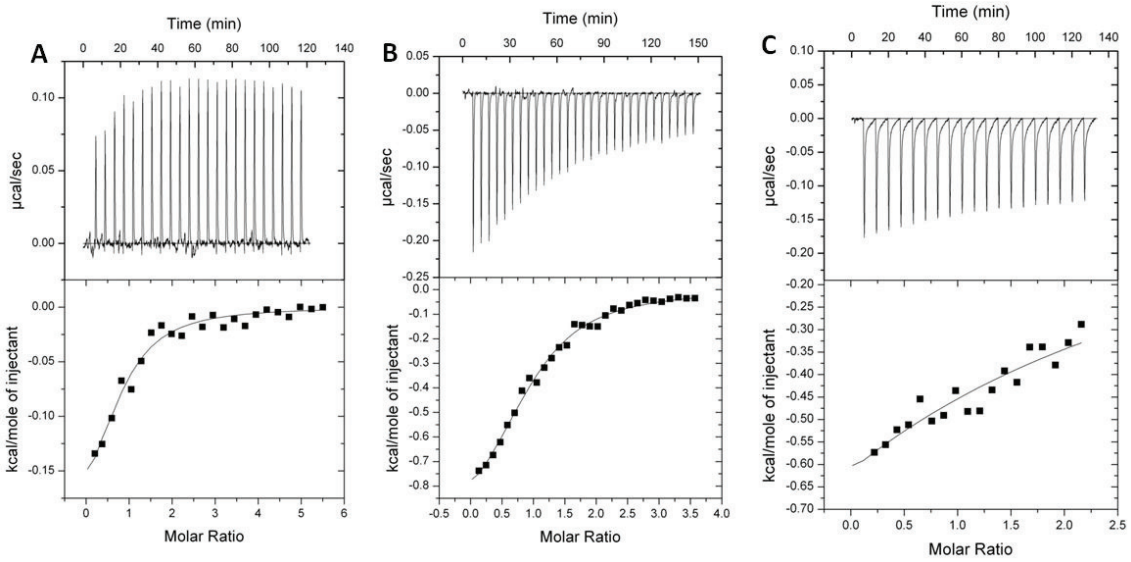
Binding Partner 1	Binding Partner 2	n	K_D (µM)	ΔH (KJ/mol)	TΔS (KJ/mol)	ΔG (KJ/mol)
Hck	PI	1.28±0.02	30.9±3.2	-19.3±0.4	6.46	-25.8
	PII	1.22±0.01	42.2±2.1	-26.5±0.4	-1.56	-24.9
	PIII	1.24±0.02	142±6.24	-26.8±0.7	-4.8	-22
$ALIX_{V+PRR}$	Hck	0.80±0.20	34.5±3.3	-15.1±1.2	10.4	-25.5
	Fyn $_{(R96I)}$	0.80±0.20	10.3±3.2	-1.0±0.2	27.4	-28.4
	Fyn $_{(R96W)}$	0.96±0.1	13.1±1.4	-4.4±0.21	23.5	-27.9
	$Fyn_{(WT)}$	1	700±240	-37.8±11.1	-20	-17.8

n: stiochiometry, $\mathbf{K_D}$: binding constant, **ka**: association constant, **kd**: dissociation constant;

Table 6.Thermodynamic parameters for Nef_{core}-SH3, as measured by ITC.

Binding Partner 1	Binding Partner 2	K_D (µM)	ΔG (KJ/mol)	TΔS (KJ/mol)	ΔH (KJ/mol)
Nef_{core}	$Fyn_{(WT)}$-SH3	ND	-28.4	25.2	-3.23
	Fyn $_{(R96I)}$-SH3	2.7	-31.8	14.5	-17.3
	Fyn $_{(R96W)}$-SH3	1.4	-33.4	-1	-34.4
	Hck-SH3	1.5	-33.3	0.7	34.4

n: stiochiometry, $\mathbf{K_D}$: binding constant, **ka**: association constant, **kd**: dissociation constant;



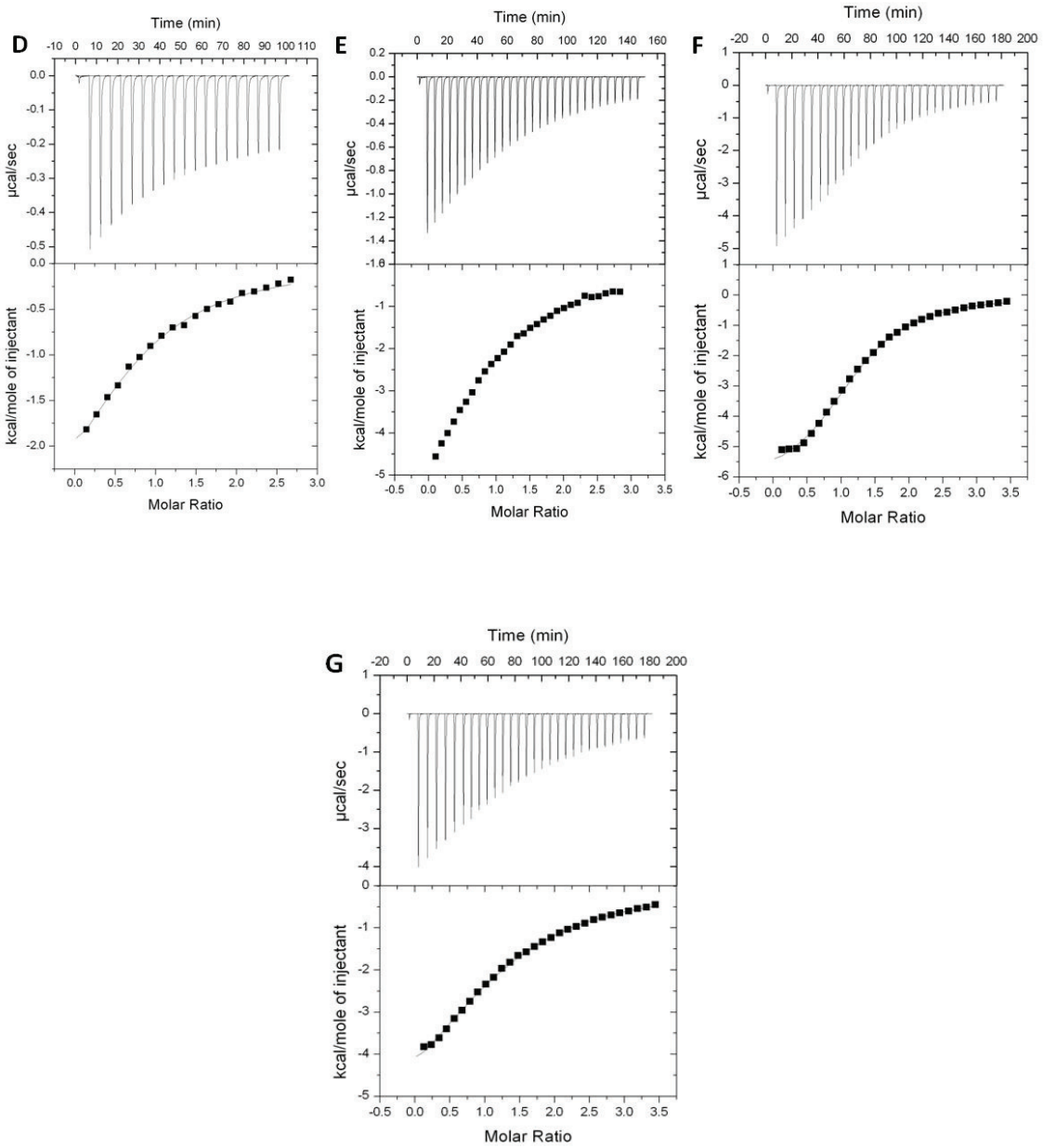


Figure 46. Binding isotherms for Hck-PxxP peptides and $ALIX_{V+PRR}$ -SH3- (Hck, $Fyn_{(WT)}$, $Fyn_{(R96I)}$ and $Fyn_{(R96W)}$) by ITC. Experimental ITC binding curves displayed for the interaction between $ALIX_{V+PRR}$ for (A)$Fyn_{(R96I)}$-SH3, (B) $Fyn_{(R96W)}$-SH3, (C)$Fyn_{(WT)}$-SH3, (D) Hck-SH3, and Hck-SH3 for (E) PI, (F) PII and (G) PIII. Top panel illustrates the raw ITC data from an experiment. Bottom panel illustrates the non-linear least squares fit of the data from the top panel, for each complex.

7.5 Surface Plasmon Resonance (SPR)

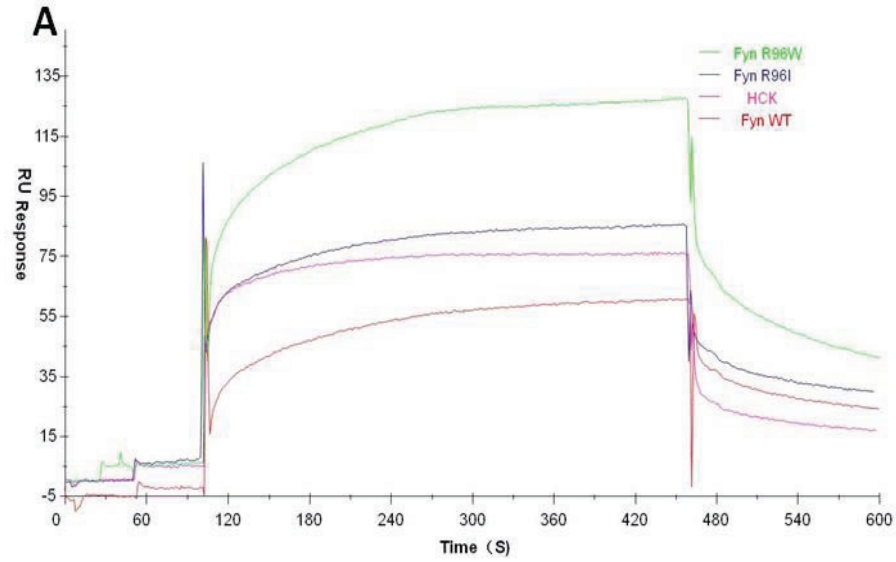
Directly binding between $ALIX_{V+PRR}$ and (Hck, $Fyn_{(WT)}$, $Fyn_{(R96I)}$ and $Fyn_{(R96W)}$)

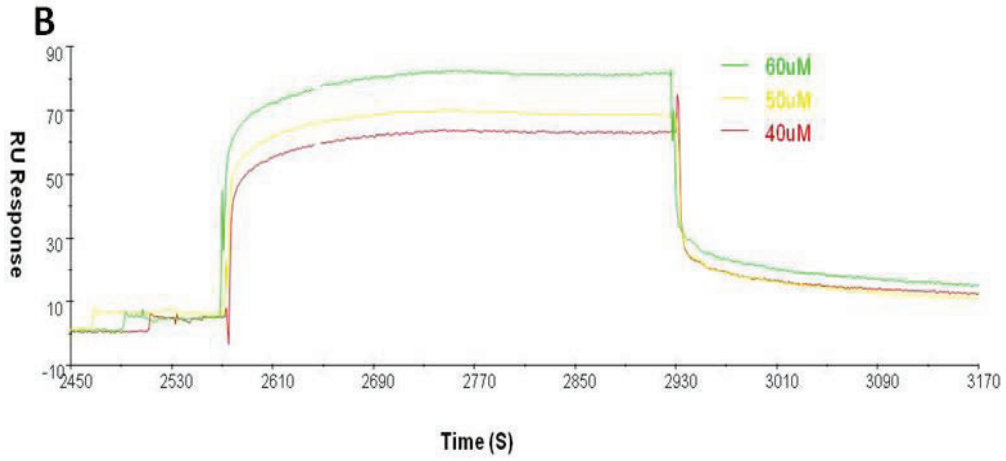
–SH3 was also confirmed using SPR. In SPR, either too low or too high concentration of $ALIX_{V+PRR}$ is unfavorable for the ligand binding, which always increases the nonspecific binding to the sensor chips. These current profiles reflect a specific association between $ALIX_{V+PRR}$ and SH3 proteins. The results of these measurement showed that the equilibrium dissociation constants (K_D) for the interaction between $ALIX_{V+PRR}$ and SH3 proteins were similar to ITC (Table 7) with an $ALIX_{V+PRR}$ / SH3 binding as a simple bimolecular interaction. As seen in Figure 47, association processes exhibited complex kinetic behavior, but the dissociation processes presented a faster off-rate compared with on-rate. The association and dissociation constants were both determined by equilibrium measurements (Table 7). This dissociation constant, as measured by SPR, is comparable with previously reported SH3/polyproline helix (PPII) peptide interaction (Kd 1-200 μM) [22]. Finally, SPR corroborated this selectivity of $ALIX_{V+PRR}$ with apparent K_D values calculated for the interaction between $ALIX_{V+PRR}$ and (Hck, $Fyn_{(R96I)}$ and $Fyn_{(R96W)}$)-SH3, respectively.

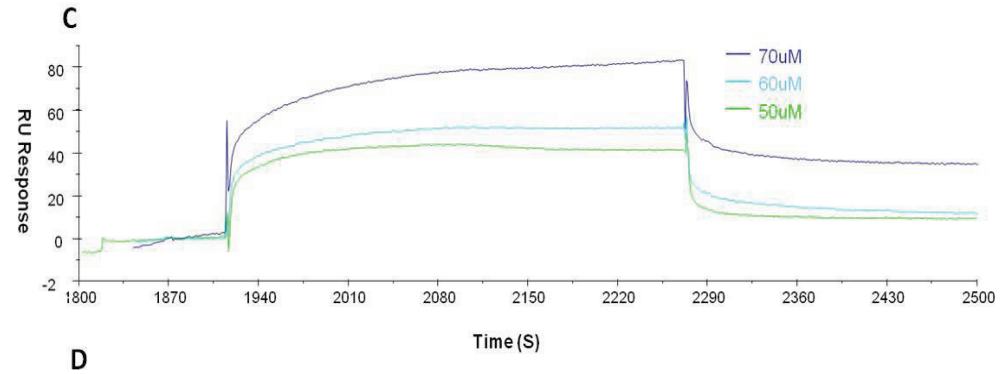
Table 7 Characterization of the SH3 domains binding to $ALIX_{V+PRR}$ by SPR

SH3 protein	K_D(uM)	ka (1/Ms)	kd (1/s)
Fyn RI	19.2	239	4.58E-03
Fyn RW	10	519	5.19E-03
Hck	24.5	336	8.24E-03

ka: association; kd: dissociation; K_D: affinity


A
Fyn R96W
Fyn R96I
HCK
Fyn WT
RU Response
Time（S)


B
60uM
50uM
40uM
RU Response
Time (S)


C
70uM
60uM
50uM
RU Response
Time (S)

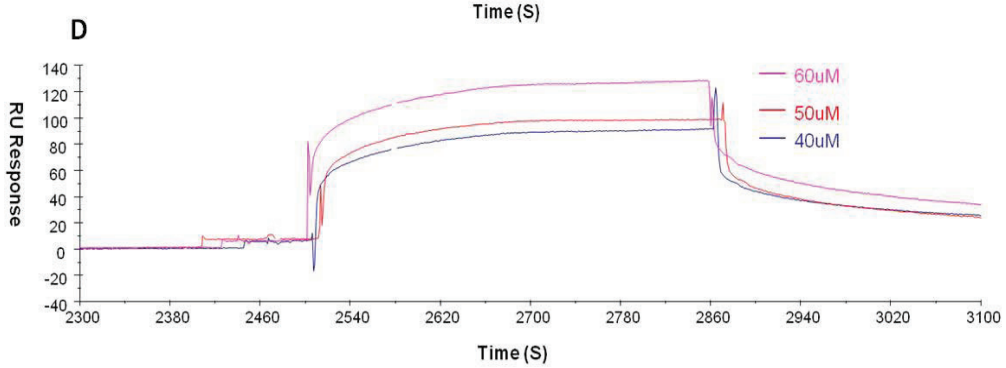

D
60uM
50uM
40uM
RU Response
Time (S)

Figure 47. SPR sensorgram. (A) $ALIX_{V+PRR}$ at 60uM concentration associating with its SH3 binding partners at 20ug/ml, respectively; $ALIX_{V+PRR}$ bound to SH3- (Hck (B), $Fyn_{(R96I)}$ (C) and $Fyn_{(R96W)}$) (D)) at different concentrations.

Structure information to elucidate the interaction between $ALIX_{V+PRR}$ and SH3 domains

The above finding demonstrates that $ALIX_{V+PRR}$ can discriminate SH3 domains according to their amino acids at position 96, a "Nef-like" selectivity, while no tertiary interaction between $ALIX_{V+PRR}$ and SH3 domains seems to exist, conversely to HIV-1 Nef. Sequence analysis and secondary structure prediction showed that ALIX PRR lies within an unstructured region, away from the folded domains (V-domain), in contrast to HIV-1 Nef (Figure 48). In the following, we will present structural characterization to address this issue.

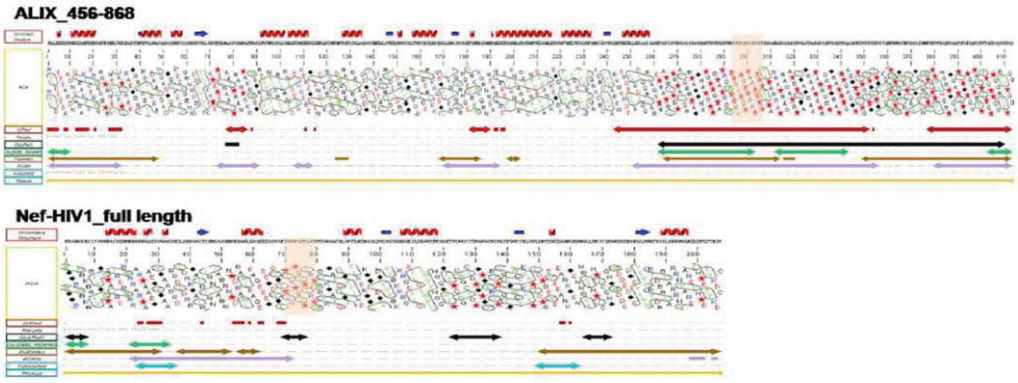


Figure 48.Prediction of disorder for $ALIX_{V+PRR}$ and HIV-1 Nef sequences. $ALIX_{V+PRR}$ and HIV-1 Nef sequences were analyzed using the MeDor Metaserver of Disorder [101]. Putative SH3-binding proline motifs are boxed in light organe.

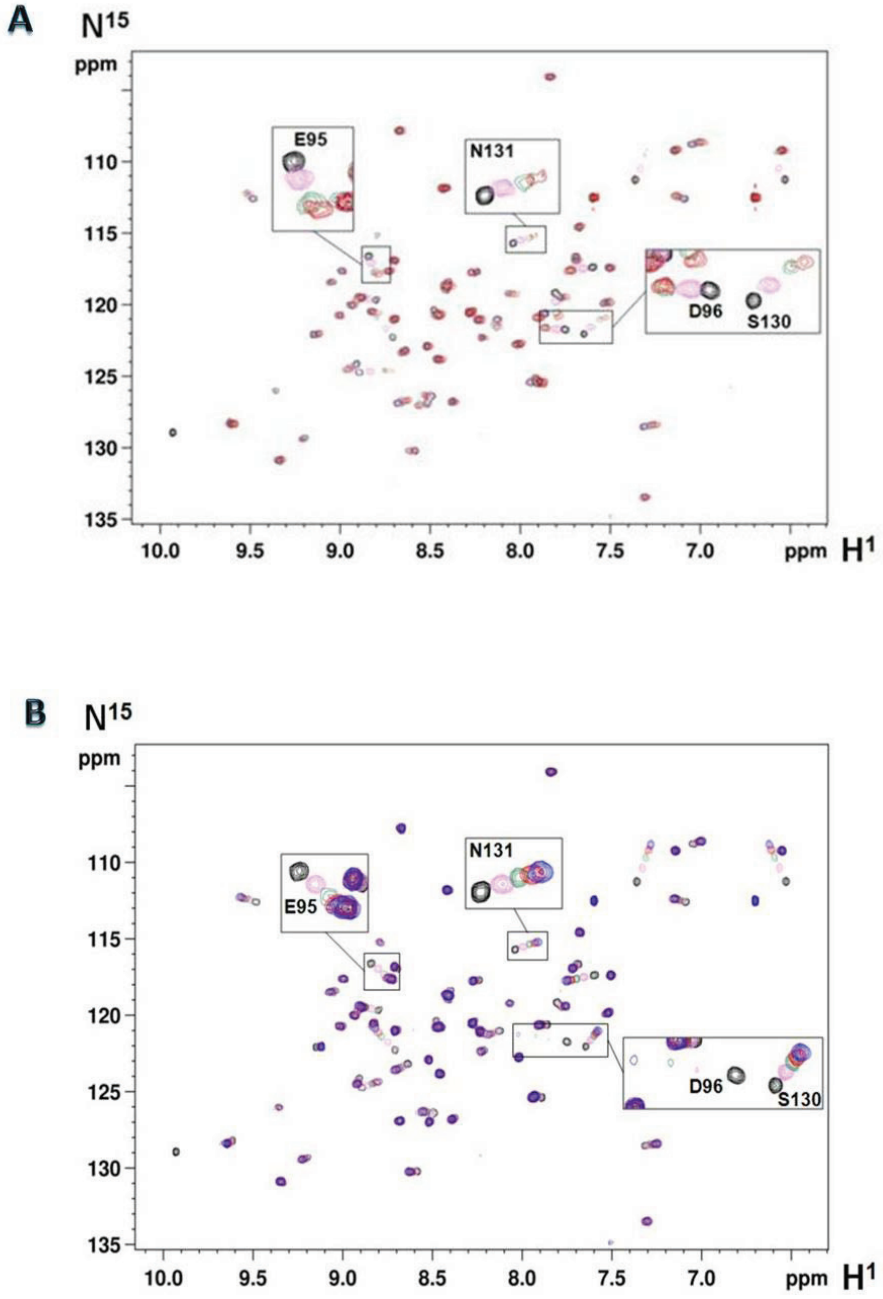
7.6 Nuclear Magnetic Resonance spectroscopy (NMR)

NMR experiments are employed to map binding surfaces of peptides or $ALIX_{V+PRR}$ onto Hck-SH3, $Fyn_{(WT)}$-SH3 and $Fyn_{(R96I)}$-SH3. HSQC spectrum of $Fyn_{(R96W)}$-SH3 is similar to that of $Fyn_{(R96I)}$-SH3 (data not shown here). HSQC spectra in Figure 49 illustrate the binding interaction between Hck-SH3 and $ALIX_{V+PRR}$ or PI. The chemical shift changes induced by $ALIX_{V+PRR}$ (Figure 49A) and PI (Figure 49B)

were measured by superimposing the HSQC spectra of Hck-SH3 domain in the free state and in presence of varying equivalents of either $ALIX_{V+PRR}$ (1.2, 2.3 and 4.6) or PI (1.1, 2.6, 5.2 and 10.2) . The relative chemical shifts of Hck-SH3 in the presence of $ALIX_{V+PRR}$ (4.6) or PI (10.2) pin-pointed that both ligands affected almost the same residues on the Hck-SH3 surface (Figure 49C). And most of the residues involved in the interaction have similar shifts, when comparing $ALIX_{V+PRR}$ titration with PI (which can be due to the similarity observed in K_D). The significantly affected residues were located over three regions, the RT loop, the n-Src loop and the 3_{10}-helix. Standard residues affected by canonical PxxP interaction should affect mainly the 'hydrophobic grooves' formed by residues Tyr87, Glu113, Trp114, Pro129 and Tyr132. The observed patch in the NMR experiment also involved the entire RT-loop and was thus more extended than expected. This observation proposed an extended recognition mode of the Hck-SH3 domain by ALIX. In addition, we have also plotted with the software "Xcrvfit" (from Pr. Bryan Sykes laboratory) the corresponding saturation curves and we have extracted the Kd_{app}'s from the saturation curves (Figure 50). We found a Kd of 52.2 ± 2.5 μM for the PI/SH3-Hck complex and a Kd of 45.2± 4 μM for the ALIX/SH3-Hck complex, both of which are consistent with the observations made by ITC and interferometry experiments (Shi et al, BJ； Table 8).

To explore the RT loop dependent selectivity and specificity, HSQC of $ALIX_{V+PRR}$ onto Hck-SH3, $Fyn_{(WT)}$-SH3 and $Fyn_{(R96I)}$-SH3 were also performed at the same condition. The results showed that the chemical shift changes induced by $ALIX_{V+PRR}$ can be mapped to almost the same residues on the surfaces of all SH3 domains, including the low-affinity wild type $Fyn_{(WT)}$-SH3 (Figure 51). It should be noticed here that the intensity of the chemical shift increased remarkably with the K_D observed (using 1 equivalent of $ALIX_{V+PRR}$ in all experiments) and that even if the global imprint are similar, subtle differences can be observed in term of surface area of the residues involved in the recognition process (Figure 51). Since the overall interaction mode is similar for all SH3 proteins, the differences in K_D must be explained by non-canonical interactions such as supplementary/extended (or tertiary) interactions or the folding of the free ALIX protein. Other cases of extended SH3

recognition sequences have already mentioned the implication of the n-Src and RT-loop in specific contact with non canonical PxxP peptides. These additional interactions can enhance both affinity and selectivity for SH3 domains.



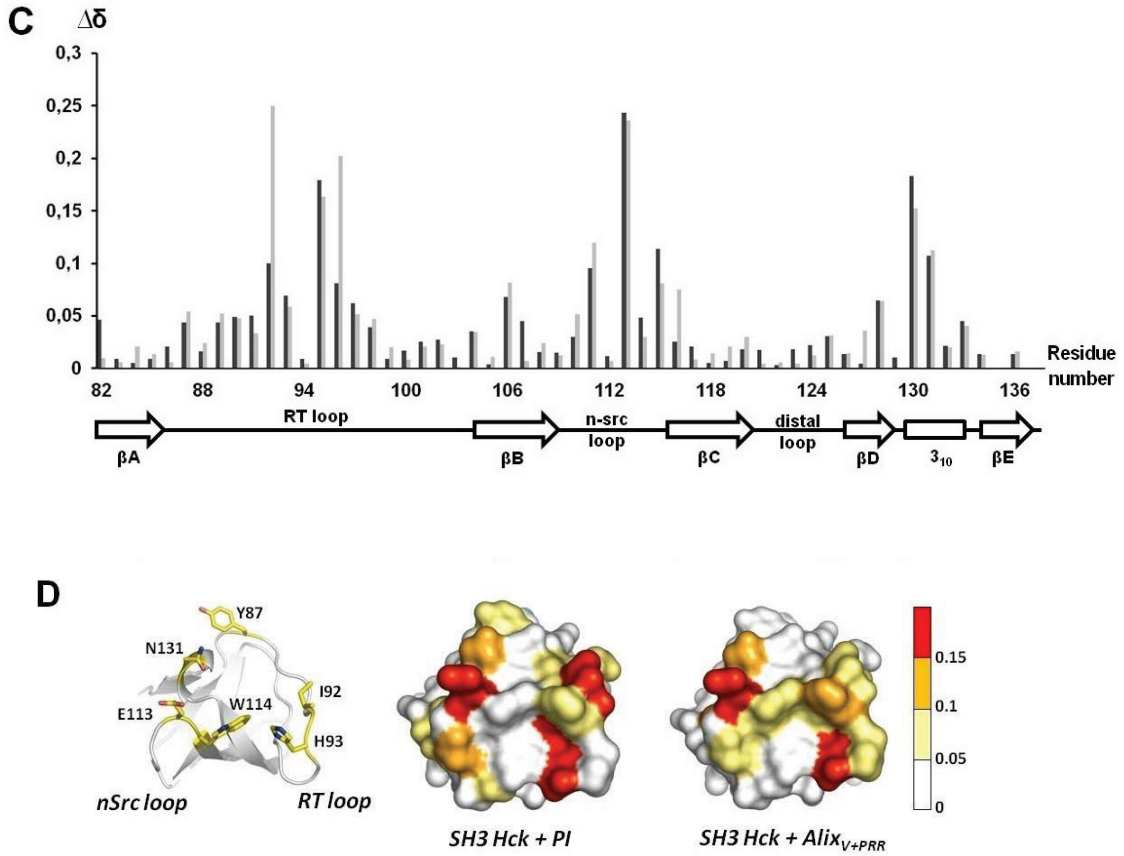
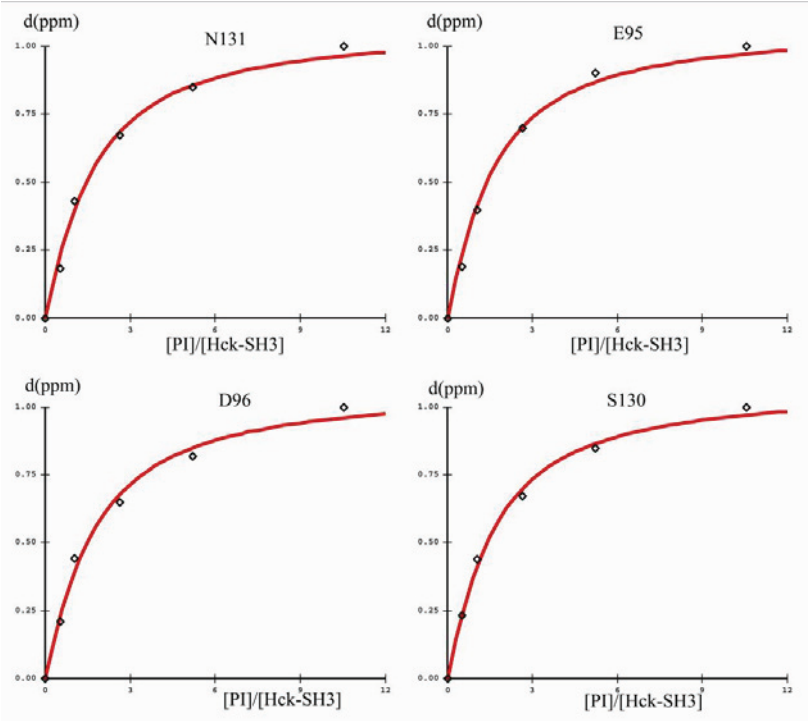


Figure 49**:** Mapping of the atomic recognition surfaces by heteronuclear NMR.

^{1}H-^{15}N NMR HSQC spectra of ^{15}N-labeled SH3-Hck in the free state (black) and in presence of varying equivalents of either $ALIX_{V+PRR}$ (1.2: mauve; 2.3: green; 4.6: red) (A); and PI (1.1: mauve; 2.6: green; 5.2: red and 10.2: blue) (B). (C) The SH3-Hck residues chemical shifts observed in presence of of $ALIX_{V+PRR}$ (4,6 equivalent) and peptide PI (10.2 equivalent) are respectively shown as histograms in dark and light grey. The relative variations of ^{1}H and ^{15}N were calculated according to Grzesiek *et al.*[102] (C) Chemical shift mapping on the surface of Hck-SH3 (PDB code: 1BU1) in presence of $ALIX_{V+PRR}$ and peptide PI. Small chemical shift changes are shaded in yellow, larger shifts colored in orange and the largest shifts are colored in red.

(A)



(B)

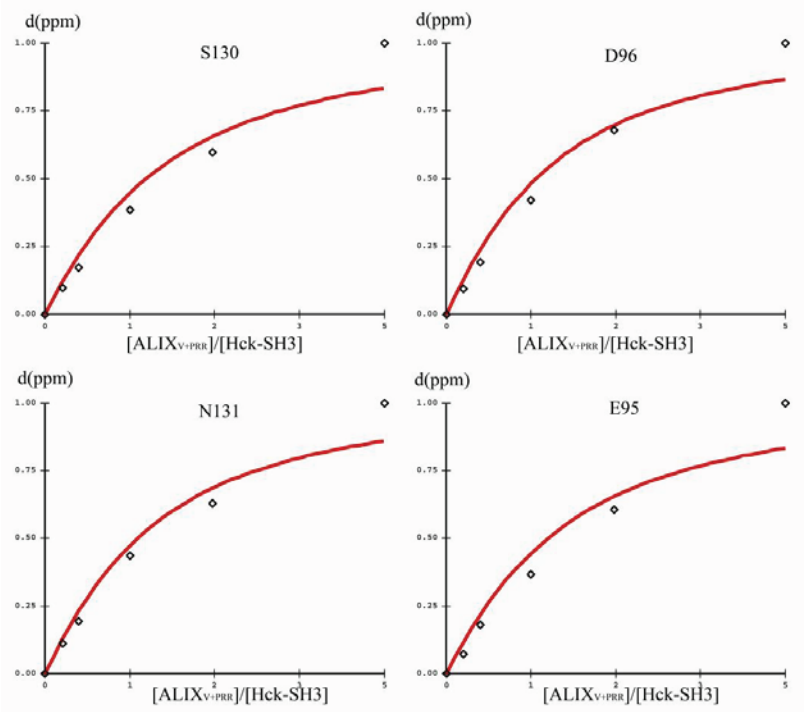


Figure 50. The corresponding saturation curvers for PI (A) and $ALIX_{V+PRR}$ (B) plotted with the software "Xcrvit".

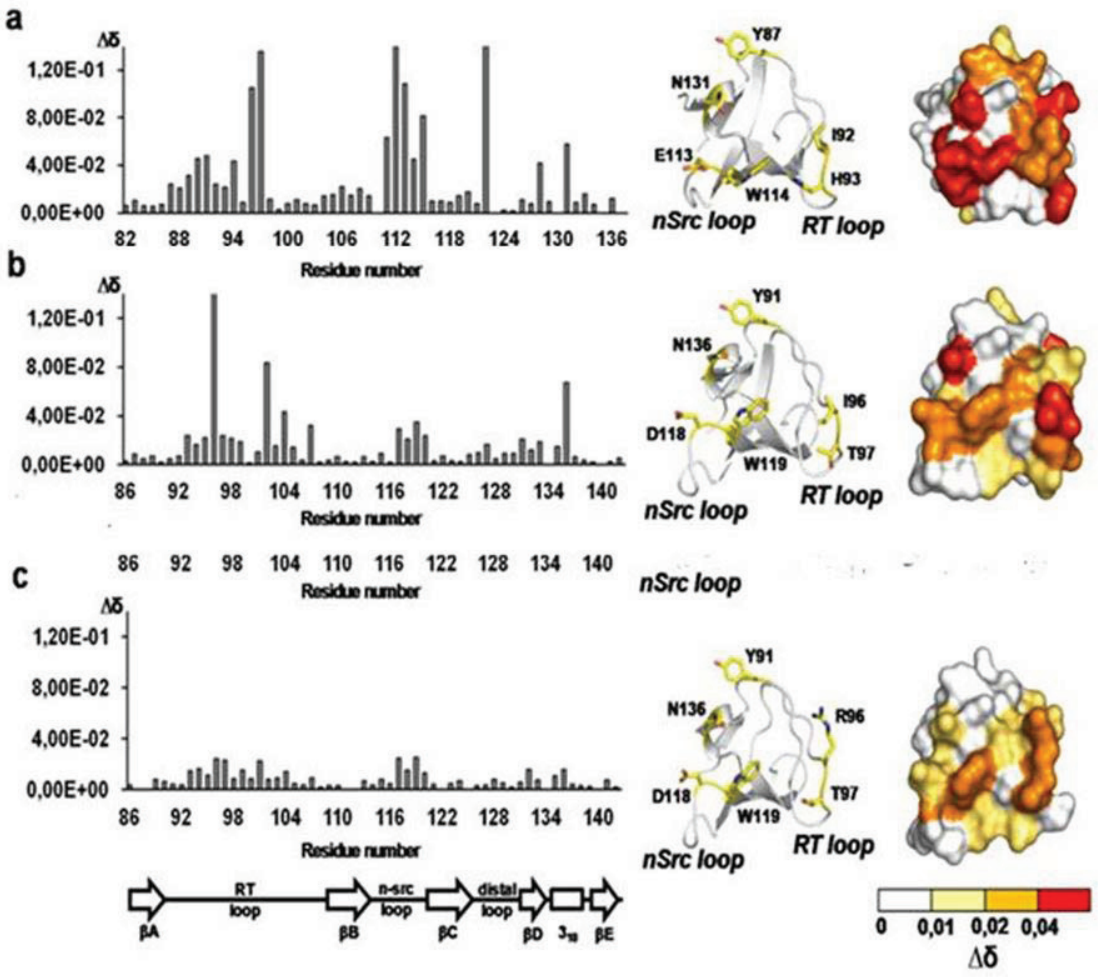


Figure 51 Mapping of the SH3 domains atomic recognition surfaces in the presence of $ALIX_{V+PRR}$ by heteronuclear NMR. (A-C) The SH3 residues chemical shifts observed for (A) Hck-SH3, (B) $Fyn_{(R96I)}$-SH3 and (C) $Fyn_{(WT)}$-SH3 in the presence of one equivalent of $ALIX_{V+PRR}$ are shown as histograms on the left part of the figure. The center part of the figure illustrates a ribbon representation of the different SH3 domains with important interacting residues in yellow for (A) Hck-SH3, (B) Fyn(R96I)-SH3 and (C) $Fyn_{(WT)}$-SH3, (PDB code: 1BU1, 1EFN and 1AVZ respectively). On the right panel of the figure, the corresponding chemical shifts of the SH3 proteins in the presence of 1 equivalent of $ALIX_{V+PRR}$ are shown as a surface representation onto the SH3 domains structures. Small chemical shift changes are shaded in yellow, larger shifts colored in orange and the largest shifts are colored in red.

Table 8 Comparison of Binding affinity from different techniques

	ITC					Interferometery			SPR	RMN
	n	K_d	ΔH	TΔS	ΔG	K_d	X^2	Iterations	K_d	K_d
		(μM)	*(KJ/mol)*	*(KJ/mol)*	*(KJ/mol)*	*(μM)*			*(μM)*	*(μM)*
$ALIX_{V+PRR}$	0.80 ± 0.20	**34.5 ± 3.3**	-15.1 ± 1.2	10.4	-25.5	**45.3**	3 e^{-2}	144	**24.5**	**45.2± 4**
PI	1.28 ± 0.25	**30.9 ± 3.2**	-19.3 ± 0.4	6.5	-25.8	**79.4**	1.1 e^{-3}	79		**52.2± 2.5**
PII	1.22 ± 0.25	**42.2 ± 2.1**	-26.5 ± 0.4	-1.6	-24.9	**169.7**	9.7e^{-4}	100		
PIII	1.24 ± 0.25	**142 ± 6.2**	-26.8 ± 0.7	-4.8	-22.0	**1351**	1.5 e^{-3}	84		

7.7 Crystallization

We did not get the crystals of SH3 domains in complexes with either PxxP-containing peptides or $ALIX_{V+PRR}$ despite extensive efforts. Instead, Apo-$Fyn_{(R96I)}$ and $Fyn_{(R96W)}$ both crystallized out of the mixture solution in different crystallization conditions and in different morphology (Figure 52 a-f). Most of the SH3 crystals diffracted very well (Table 9). Moreover, we also tried to crystallize $ALIX_{V+PRR}$ alone, and obtained micro-crystals (Figure 52 g and h). However, these crystals were quite small and not suitable for diffraction (Figure 52 g and h).

I determined the crystal structures of Apo-$Fyn_{(R96I)}$ and $Fyn_{(R96W)}$. The residues in RT-loop, especially residue D100 in all final $Fyn_{(R96I)}$ SH3 atomic models, adopt a pseudo bound form, where it engages a intermolecular ionic interaction with a crystallographic symmetric-related molecule, similar to the SH3 D100-Nef R77 salt bridge, previously observed in SH3-Nef structures (PDB entry3HOH and 1EFN) [67]. By superposition of structure models of free-state $Fyn_{(R96I)}$ SH3 in different crystal forms (bound and unbound forms) (final atomic model:88lcl_2; PDB code: 3HOH and 3HOI; respectively), the main difference stems from the conformation of the RT and nSrc loops (Figure 53). In order to explain why the Ile96 in free-state $Fyn_{(R96I)}$ SH3 also

adapt the bound form that is performed in complex structures, we analyzed the crystal packing in alternative forms. In the bound form, the crystal lattice is more tightly packed than the one in unbound form, which maycontribute to the variation of RT and nSrc loops and alternative conformations of I96 in $Fyn_{(R96I)}$ SH3. The tightly packed lattice of bound form prevents the strategy to soak the binding partners into the canonical PxxP binding pocket and forming complexes (Figure 54). This tight packing and strong intermolecular association may explain the difficult to get the complex crystals. Moreover, in our $Fyn_{(R96I)}$ structure, the side chains of residues T97 and D99 in the RT loop and D118 in nSrc loop are partly disordered according to the electronic density map, suggesting high flexibility of these residues.

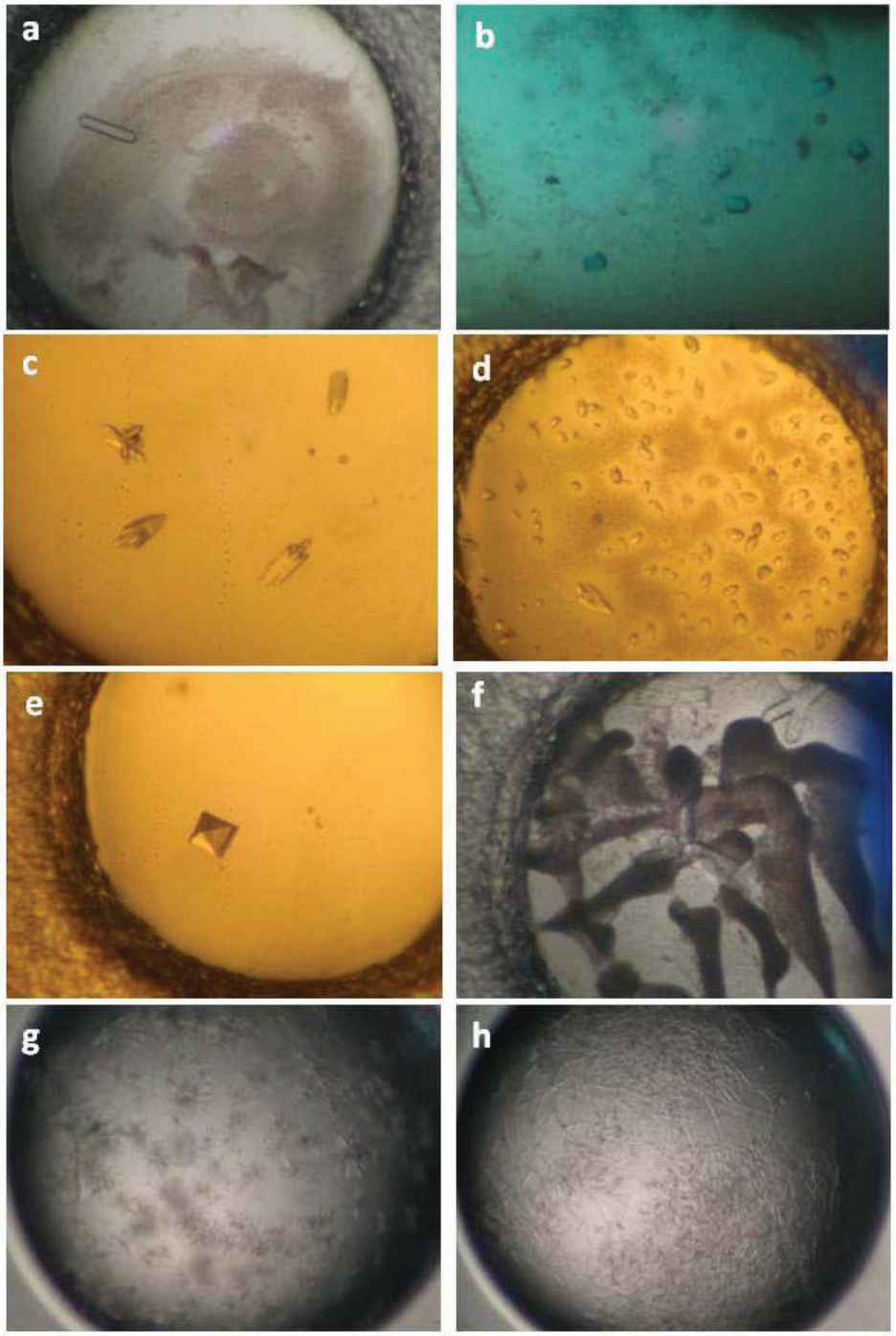


Figure 52. $Fyn_{(R96I)}$SH3 crystals appearing in different condition.(a) 30%PEG3350, 0.4N NaNO3 (b) 30%PEG3350, pH5.5 (c)-(d) 30% PEG 2000 MME, 0.05M Na Cacodylate pH6.0, 0.05M MgCl2 (e) 30% PEG 550 MME, 0,1M Bis-tris Ph7.0, 0.05M MgCl2; (f) $Fyn_{(R96W)}$ SH3 crystals : 25% PEG 5000MME, 0,1M Tris-HCl pH7.5; 17% PEG 10000, 0,1M Bis-tris pH5.5, 0.1M NH4-acetate; 25% PEG 3350, 0,1M Bis-tris pH5.5, 0.1M NaCl; 25% PEG 3350, 0,1M Hepes pH7.5, 0.2M NaCl. (g)-(h): $ALIX_{V+PRR}$ crystals

Table 9 Summary of X-ray diffraction data collection for $Fyn_{(R96I)}SH3$ crystal

Data collection	88lcl_2	w9t
X-ray wavelength (Å)	1.04	1.04
Space group	$P4_1$	$P4_1$
Cell dimensions (Å)	41.442, 41.442, 32.786 α=β=γ=90º	41.641, 41.641, 32.713 α=β=γ=90º
Independent reflections	18,660	14,689
Resolution range (Å)	50.0-1.5 (1.56-1.5) [a]	50.0-1.28 (1.31-1.28) [a]
Completeness (%)	98.4 (95.92) [a]	97.72 (86.42) [a]
Multiplicity	9.6 (5.5)[a]	6.2 (3.9)[a]
Rsym(%)[b]	12.9 (34.3) [a]	5.7 (21) [a]
I/σ	23.03 (2.8) [a]	33.14 (2.51) [a]
Refinement statistics		
Rwork(%)[c]	16.39 (17.5)	18.51 (18.1)
Rfree(%)[d]	21.38 (33.2)	21.27 (24.5)
Root mean square deviation from ideal values		
Bond lengths (Å) Bond angles (°)	0.026 1.897	0.025 2.147
Percent of residues with most favoured, additional allowed, and generally allowed dihedral angles	98.39, 1.61, 0	95.38, 4.62, 0

a. Values in parentheses are for the outermost resolution shells

b. Rsym = Σ| (Ihkl) – (I)|/Σ (Ihkl) where Ihkl is the weighted mean intensity of a given reflection.

c. Rwork= Σhkl |Fobs -Fcalc|/ΣhklFobs, where Fobs and Fcalc are the observed and calculated structure factors, respectively.

d. Rfree is the Rwork calculated using a randomly selected 5% sample of reflection data omitted from the refinement.

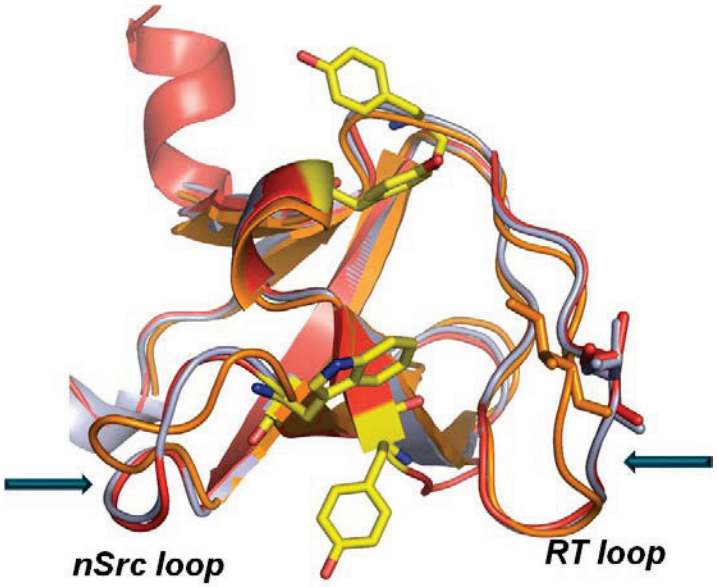


Figure 53 Superposition of $Fyn_{(R96I)}$ SH3 structures ribbon diagrams. Red: final atomic model (88lcl_2); Grey: bound $Fyn_{(R96I)}$ SH3 (PDB entry 3HOH); Orange: unbound $Fyn_{(R96I)}$ SH3 (PDB entry 3HOI)

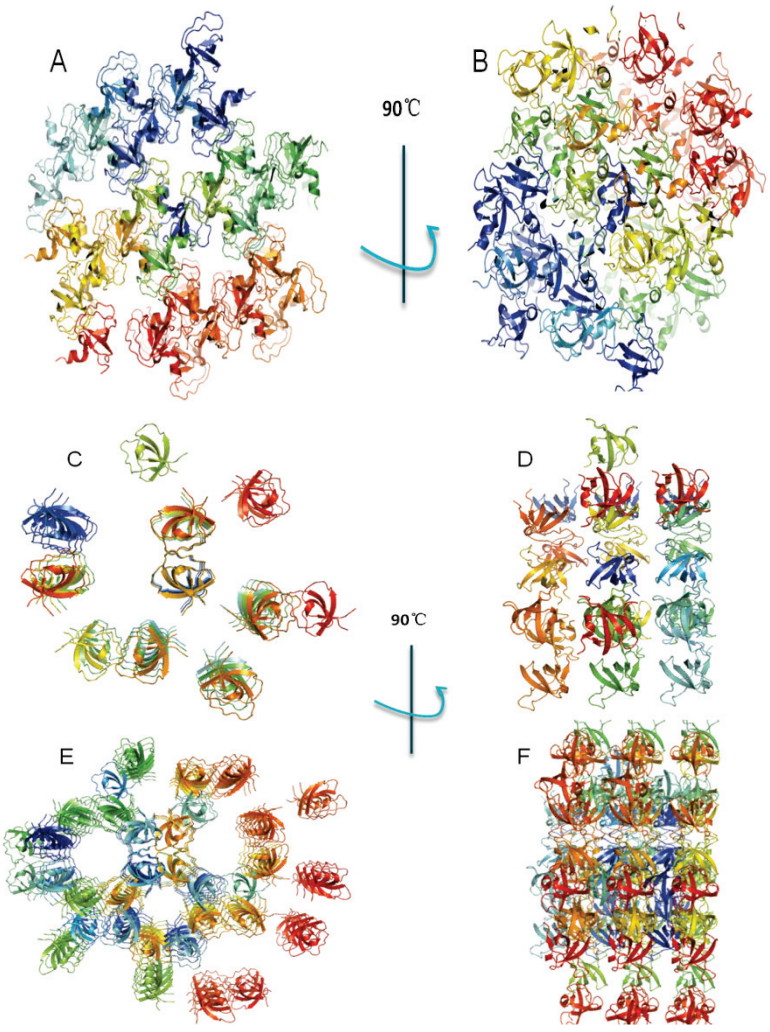


Figure 54 Comparison of the crystal packing of $Fyn_{(R96I)}$ SH3 structures. (A-B) bound $Fyn_{(R96I)}$ SH3 (88lcl_2); (C-F) unliganded $Fyn_{(R96I)}$ SH3 (PDB entry 3HOI): C-D is derived from the one molecule in a ASU, and E-F is from two molecules in the asymmetric unit. Different conformations in crystal lattice are shown in different colors.

RT loop flexibility related to the selectivity of HIV-1Nef to SH3 domain

It has been well demonstrated that the RT loop contributes to the specific interaction between SH3 domain and HIV-1 Nef outside of the canonical PxxP motif. Crystal structures of free SH3 domains and SH3 domains in complex with HIV-1 Nef proteins were analyzed and corroborated that Ile96 is one of the specific determinants that contributes to the flexibility of RT loop.

The specific amino acid of Arg77 in the HIV-1 Nef PxxPxR motif greatly contributes to the integrity of the interface between SH3 domains and HIV-1 Nef. Upon binding to a SH3 domain, Arg77 at the PxxP motif established a salt bridge with SH3 Asp100, and in doing so, disrupts the hydrogen-bond network that stablizes the SH3 domain RT-loop in its unliganded form [66, 67]. For example, previous studies show that SH3 Asp100 plays a pivotal role in maintaining a network of hydrogen-bonds in either unbound (12 H-bonds) or bound (8 H-bonds) $Fyn_{(WT)}$SH3 RT loop. In unbound $Fyn_{(WT)}$SH3, Asp100 interacts with the side chain of Tyr93 and Thr97 as well as the main chains of Thr97 and Arg96 (Figure 55A). Additional hydrogen bonds are made between the side chain of Arg96 and the proximal main chain of Tyr132 of the underlying strand βD, which stabilize the RT loop. Upon complex formation, Asp100 at $Fyn_{(WT)}$SH3 bends away from its center position to form a salt bridge with Arg77 in HIV-1 Nef, which disrupts the hydrogen-bond network of the free RT-loop (Figure 55B). The disruption of Asp 100 interaction network confers to the RT loop an increase of flexibility, and allows the interaction with the tertiary component of the HIV-1 Nef surface. Arg96 and Tyr97 are conserved in the Src SH3 domain, resulting in restriction of mobility of the RT loop observed in $Fyn_{(WT)}$SH3. On the one hand, the substitution of Arg96Ile precludes the hydrogen

bonding network seen in the $Fyn_{(WT)}$SH3 and increase the flexibility of RT loop (Figure 55C)[22, 66, 67]; on the other hand, the substitution is more compatible with the hydrophobic pocket on the surface of HIV-1 Nef. For instance, Compared with unbound $Fyn_{(WT)}$SH3, Asp100 in unbound Hck SH3 forms less hydrogen bonds (7 H-bonds) in the network because two substitutions of (Arg96Ile, Thr97His) preclude the hydrogen bonding network. Modellisation studies suggested that Asp100 in the complex between Hck SH3 and HIV-1 Nef adopts the same orientation with that in $Fyn_{(WT)}$SH3 complex with HIV-1 Nef and this complexation also requires breaking of less hydrogen bonds (3 H-bonds) (Figure 55C) [22, 66]. Consequently, in the free state, the RT loop in Hck SH3 is more mobile than that in Fyn or Src SH3. Moreover, though it is also seen in the Arg96Ile mutant of $Fyn_{(WT)}$SH3 that the mutant Ile96 mimicks Ile96 in Hck SH3, there is extra a bound water molecule in the model of HIV-1 Nef-$Fyn_{(R96I)}$SH3 occupying the Arg96 guanidino group position in $Fyn_{(WT)}$-SH3 bound with HIV-1 Nef (Figure 55D) [66]. Comparison with the crystal structure of HIV-1 Nef in complex with another Fyn SH3 mutant ($Fyn_{(R96W)}$SH3) which forms 8 H-bonds and 4 H-bonds in free and complex states respectively, the conformation of residue Trp96 is consistent with Ile96 except that it is observed to insert more deeply into the conserved aromatic pocket at the interface forming the complex (Arold & Dumas, unpublished results). Based on this, the R96-containing RT loop has more flexibility than the one containing I96 (Figure 56). Above all, the RT loop inflexibility and hydrophobicity of amino acid 96 in RT loop enhances the specificity and selectivity of SH3 domain of SFKs for HIV-1 Nef.

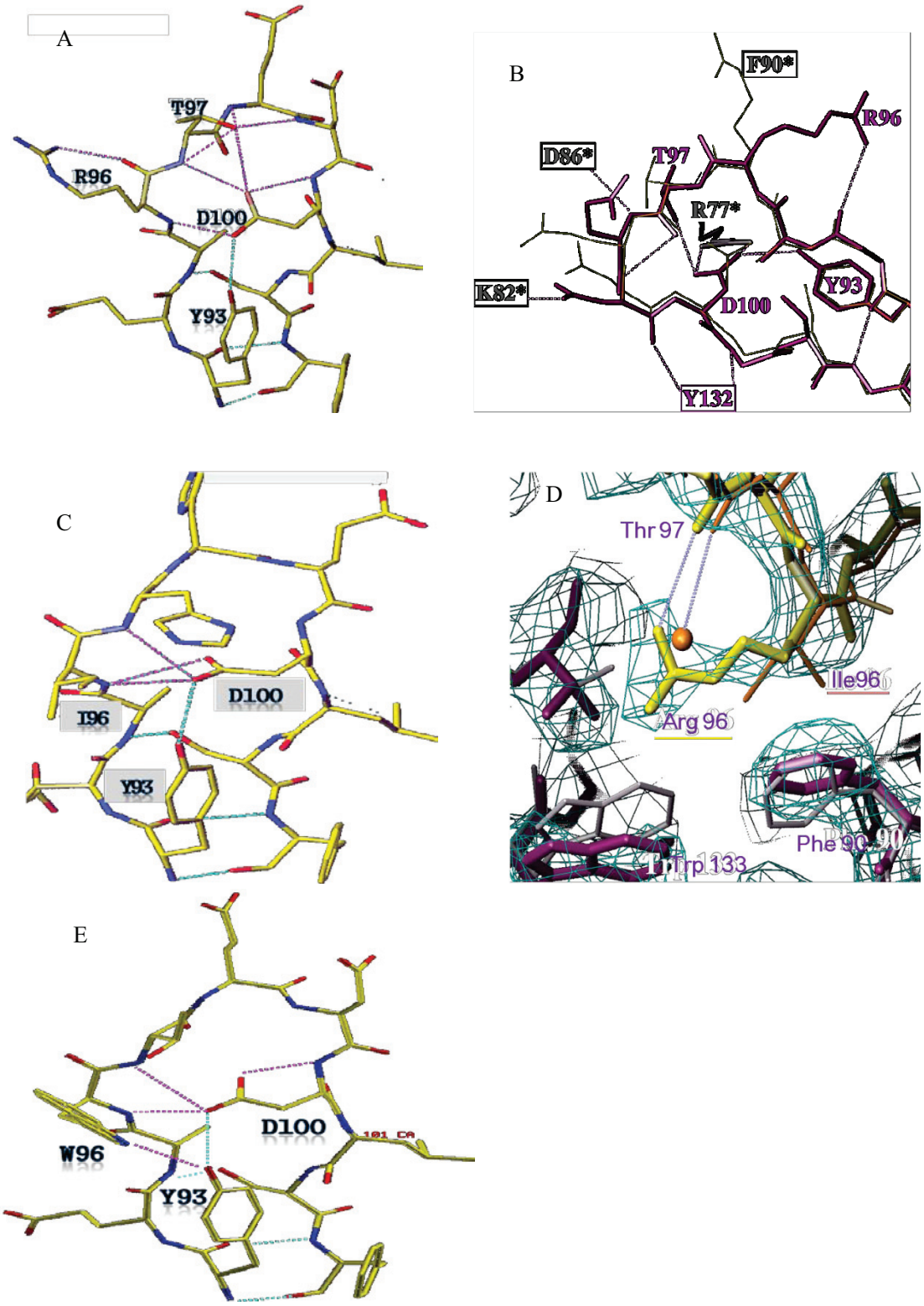


Figure 55.Difference between the D100-mediated hydrogen-bond network. (A) unbound $Fyn_{(WT)}$-SH3 (PDB entry 1SHF); (B) superimposition of bound (PDB entry 1AVZ) and unbound $Fyn_{(WT)}$-SH3; (C) unbound Hck-SH3 (PDB entry 1BU1); (D) superimposition of bound $Fyn_{(WT)}$-SH3 and $Fyn_{(R96I)}$-SH3 (PDB entry 1EFN). A water is presented as orange [22]; (E) unbound $Fyn_{(R96W)}$-SH3 (PDB entry 3HOF); The dot lines presents as hydrogen-bonds. The blue

dot lines in figure C & D are the ones disrupted upon complex formation with HIV-1 Nef.

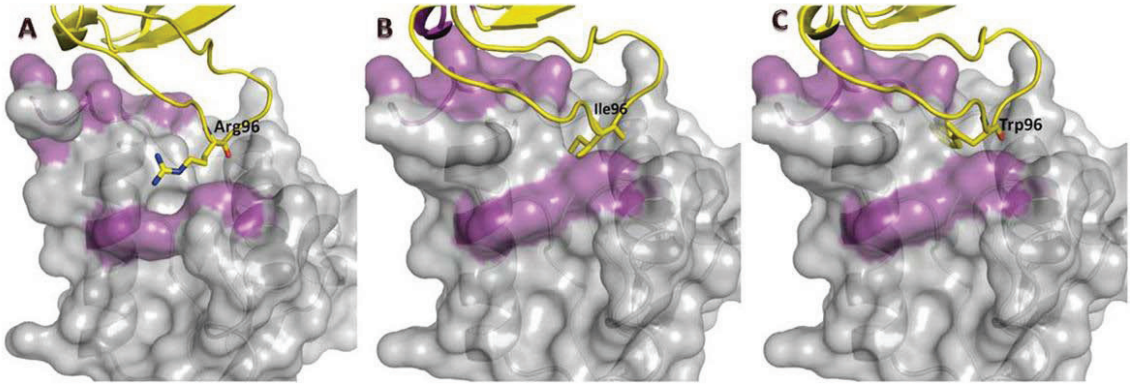


Figure 56 Conformational comparison of residues 96 of various SFKs SH3 in complex with HIV-1 Nef. (A): $Fyn_{(WT)}$-SH3 (PDB entry 1AVZ); (B): $Fyn_{(R96I)}$-SH3 (PDB entry 1EFN); (C): $Fyn_{(R96W)}$-SH3 (PDB entry 3HOF); SFKs SH3 are showed in yellow. HIV-1 Nef is presented in grey, of which the PxxP region and the interface of the aromatic pocket are labeled out in purple.

7.8 Small Angle X-ray Scattering (SAXS) Data Collection and Analysis

Fisher *et al*. suggested that the PRR region of ALIX is disordered [91] and no structural details could be obtained regarding this highly disordered region. This large flexible region may hamper crystallization of $ALIX_{V+PRR}$ alone or its complexes, since we were unable to obtain suitable crystals despite extensive screening. SAXS was used to access this structural recognition process by the analysis of apo-$ALIX_{V+PRR}$ and $Fyn_{(R96I)}$SH3-bound $ALIX_{V+PRR}$.

The *ab initio* envelopes calculated for $ALIX_{V+PRR}$ SAXS data were consistent with a model where residues 703 to 760 were protruding freely in the solvent (Figure 59A, B), without forming a folded domain, and without substantial contacts with the V-domain. This was also supported by an analysis using the program EOM. EOM allows to select an ensemble of models that best fit the data, from an initially generated pool of about 10,000 randomly generated structures (in our case, we used the ALIX-V domain of PDB entry 2OJQ as fixed structure, and only varied the flexible residues 703 to 760. EOM-selected ensemble models showed a high flexibility and extended structure of residues 703 to 760 (Figure 59A, B). So the analysis of $ALIX_{V+PRR}$ alone supports that the structural flexibility of the PRR in

absence of binding partners, in agreement with secondary structure predictions for this region. If tertiary contacts were to occur between $Fyn_{(R96I)}$-SH3 and the folded core (the V domain) of $ALIX_{V+PRR}$, then SAXS should reveal a complex with the SH3 domain positioned close to the V domain. Because the SH3 binding motif is situated at the C-terminal end of $ALIX_{V+PRR}$, 50 residues away from the V domain, a tertiary SH3 binding mode should also result in bringing this segment close to the V domain core, and hence yield a reduced maximum diameter (Dm) for the SAXS data. However the Dm of both apo and $Fyn_{(R96I)}$-SH3-bound $ALIX_{V+PRR}$ were highly similar (185 ± 10 Å), and EOM reconstructions showed both a high flexibility and large and variable distance of the PRR-bound $Fyn_{(R96I)}$-SH3 from the V domain core (Figure 59C, D). It can be noticed here that the aggregation observed on the diffusion curve is certainly due to the high concentration of the solution (50 mg/ml) needed to be sure to observe more than 99% of the bound proteins, given the affinity of the SH3 interaction.

These data do not support a HIV-1 Nef-like tertiary SH3 recognition mode of $ALIX_{V+PRR}$ towards Hck-SH3 binding, and corroborate that a HIV-1 Nef-like selectivity can be achieved by a linear mode of interaction, since ALIX PRR lies within an unstructured region, away from folded domains capable of providing the support for a tertiary interface. Thus, this result is consistent with our previous analysis that there is no extended interaction on the $ALIX_{V+PRR}$ besides the canonical PxxP consensus. This analysis supports the notion that the structural flexibility of the PRR domain allows both SH2 and SH3 interaction to occur simultaneously possibly leading to the described phosphorylation of the Brol domain. We therefore conclude that the SH2 and SH3 domains contribute differently to binding, which resulted in functional differences of the complexes formed by ALIX and Src kinases.

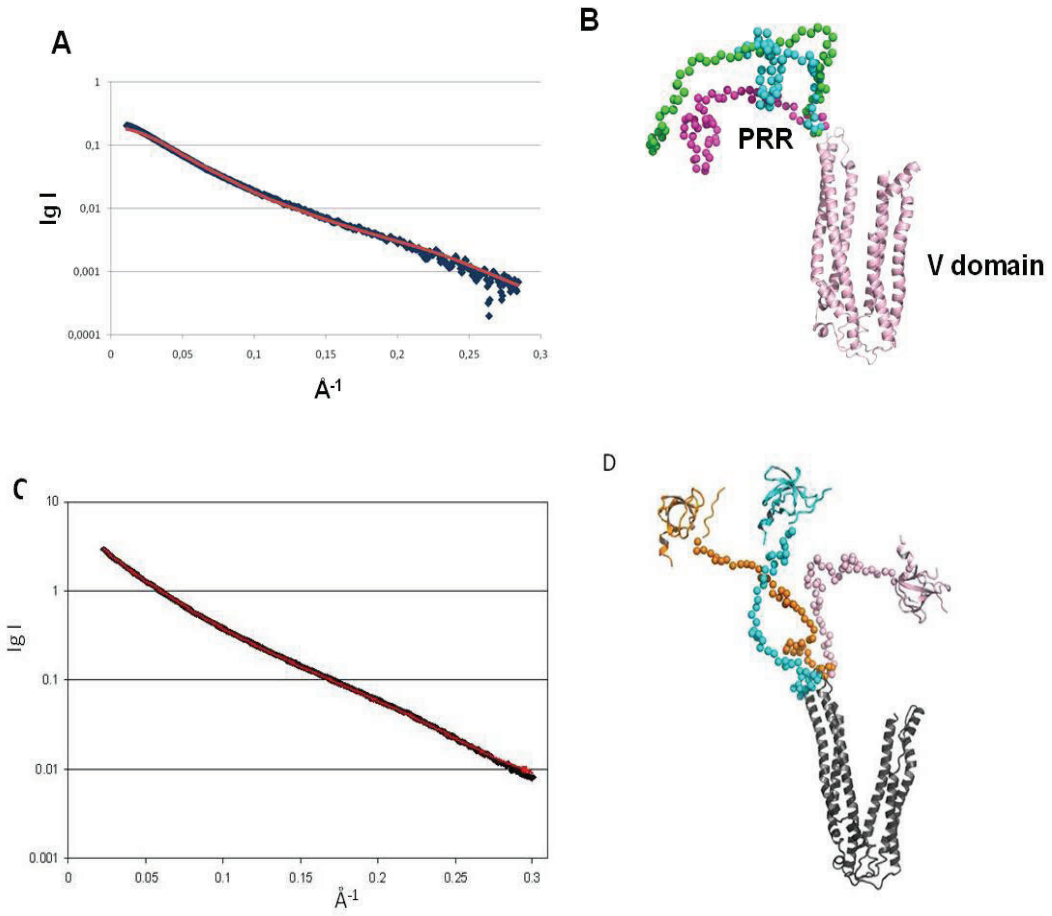


Figure 59: SAXS model of apo-$ALIX_{V+PRR}$ (A-B) and $Fyn_{(R96I)}$-SH3-bound $ALIX_{V+PRR}$ (C-D) (A,C) Fit of EOM SAXS ensemble model (red) to data (black). (B-D) Three representative structural models produced by EOM based on SAXS data were superimposed on their ALIX-V domain. Pseudo-residues placed by EOM (spheres) are colored differently for each individual model.

Part IV Conclusions

ALIX has been well recognized as a functionally cellular protein since it was discovery in 1999. The combined studies of cell biology and structural biology studies have elucidated the mode of membrane-related trafficking or virus budding via N-terminal Bro 1 and middle V domains. The C-terminal PRR domain is comprised several polyproline regions and was predicted as non-structural, however, there is no detailed information on the ALIX functions in various cellular interacting partners via its polyproline sequences and its relationship with the other domains. Previous studies has shown that ALIX was an SH3-binding partner of endophilin, as evidenced by yeast two-hybrid analysis [103], and a functional binding partner of the SH2 and SH3 domains of the *Xenopus* and mammalian Src and Fyn orthologs [104, 105]. Moreover, the PRR was demonstrated to play a key role in regulating ALIX activity, which keep ALIX in an auto-inhibited conformation through intra-molecular interactions and activate ALIX by taking part in inter-molecule interactions. Indeed, some degrees of flexibility between the two arms of the V-shape domain and tripeptide connecting Bro 1 and V domains can bring the N-terminal Bro 1 domain and C-terminal PRR into spatial proximity allowing additional interactions between phosphotyrosine in patch 2 from the Bro1 domain and Src family SH2 domains. This proximity could then mediate an intra molecular interaction regulated by the PRR, which would allowed to mask the binding sites in both Bro1 and V domain and thus inhibit the ALIX activity. In the presence of Src Kinase enzymes, this conformational locked protein could then facilitate the Src kinase to interact with Bro1 and PR region simultaneously through its SH2 and SH3 domains, thus releasing the auto-inhibited ALIX conformation into an open and active protein conformation. However, the molecular details of these interactions remained elusive. Based on the Y2H screening carried out by our collaborators, it is presumable that the segment containing canonical PxxP motif and its upstream sequence of V domain is responsible for the interaction between ALIX and SH3 domain of SFKs Hck. Here, we combine several methodologies to study the interaction mechanism mediated by ALIX canonical PxxP region and SFR SH3 and to

answer the following important questions

Firstly, does $ALIX_{V+PRR}$ interact with SFKs Hck in vivo? Checkmate experiment was performed. In this assay, the $ALIX_{V+PRR}$ construct induced transcription of the luciferase reporter gene upon co-expression with the wild type but not with a mutated, non functional, Hck-SH3 domain. So $ALIX_{V+PRR}$ is a Hck interactor.

Secondly, is the interaction between ALIX and Hck is biological relevant? To study this, we verified that full length ALIX and Hck proteins can interact in mammalian cells HEK293 cell by co-IP and kinase assay. We thus concluded that Hck not only interacted with ALIX but also displayed a significantly increased catalytic activity upon co-expression with ALIX in HEK293 cell. Collectively, these experiments show that ALIX-Hck interaction is biologically relevant as the two proteins can structurally and functionally interact in mammalian cells.

Most importantly, the *in vivo* cellular assays show ALIX can functionally interact with Hck and $ALIX_{V+PRR}$ construct, which seems to be similar to HIV-1 Nef protein in specific and selective interaction with cellular signal proteins SFKs SH3 domain. The cellular roles of ALIX overlap with those of Hck, contributing to regulation of cellular trafficking [106] and viral propagation [107].

Thirdly, does $ALIX_{V+PRR}$ interact with SFKs Hck in vitro and is ALIX a HIV-1 Nef like protein? The GST pull-down, ITC, SPR and interferometry experiments corroborated with the in vivo assay that $ALIX_{V+PRR}$ was able to interact Hck, Fyn mutant ($Fyn_{(R96I)}$ and $Fyn_{(R96W)}$), but not $Fyn_{(WT)}$. Our analysis also identified amino acid residues 737–760 as an ALIX proline-rich motif involving in Hck-SH3 domain binding. These results are consistent with the potent role of HIV-1 Nef function in interaction with SKFs SH3 domains. Taken the observation from ITC, it seemed that $ALIX_{V+PRR}$ adopts the different recognition mechanism with HIV-1 Nef to various SH3 domains.

Fourthly, is ALIX a HIV-1 Nef like protein preforming the 3D recognition mechanism toward SFKs SH3 domains, and what is the structural information about PRR? Combination of sequence analysis, secondary structure prediction and HSQC experiments, PxxP motif mainly contributes to the interaction between ALIX and

SFKs SH3 domain and the overall interaction mode is similar for all SH3 proteins. The differences in K_D must be explained by non-canonical interactions such as supplementary/extended (or tertiary) interactions or pre-structuring of the free ALIX protein. In our case, the failure of the crystallization seems due to the flexibilily of RT and nSrc loop in SH3 domain. SAXS was then used to access this structural recognition process by analysis of apo-$ALIX_{V+PRR}$ and $Fyn_{(R96I)}$-SH3-bound $ALIX_{V+PRR}$. Previous studies have demonstrated that the PRR keeps ALIX in an auto-inhibited conformation [108]. Here, Our SAXS analysis supports this proposition since the observed structural flexibility of the PRR could enable it to mediate an intra-molecular interaction inhibiting the activity of ALIX and also to facilitate the interaction of Src kinase via its SH2 and SH3 domains with the ALIX N-terminal Bro1 domain and C-terminal PRR, respectively. We also demonstrated that ALIX residues outside of the canonical PxxP motif enhanced affinity and also possibly specificity for SH3 domains through a 'linear' binding mode which appear to be different with the unique tertiary SH3 interface displayed by HIV-1 Nef. The structural framework of the SH3–ALIX interaction demonstrated here will help to elucidate how SH3 domains of SFKs, such as Hck, and ALIX assist in virus budding and regulation of cell surface receptors, which might be particularly relevant in distribution tissues that specifically express SFKs and that constitute an important reservoir for HIV spreading.

Future direction

Our results show that the specific hydrophobic pocket on HIV-1 Nef can be a promising drug target. Given the need for selectivity toward cellular SH3-ligand interactions, it is surprising that the tertiary binding mode evolved by the HIV-1 Nef protein has no (or only marginally few) cellular counterparts. This may suggest that the 3D-binding mode of HIV-1 Nef might not be suited for signaling by SH3 binding cellular proteins. Indeed, the tertiary recognition mode of HIV-1 Nef results in SH3 binding that is so strong, that it not only activates Src kinases that contain a favored isoleucine at position 96 (such as Hck or Lyn), but also perturbs the function of those

with less favorable residues (such as Lck, with a serine, or Fyn and Src with an arginine) [109-112]. Moreover, it does not even restrict selection to Src kinases only, since HIV-1 Nef also interacts with the C-terminal Vav1 SH3 domain (containing an arginine in the equivalent to position 96 in Vav1 SH3 RT-loop) [113]. The HIV-1 Nef SH3 domain binding mode may therefore be too strong to be widely used in cells.

Concerning HIV pathogenesis, it is unclear whether this mode of (strong) binding was selected by HIV-1 Nef allowing it to perturb a wide range of cellular SH3 components in different cells and/or subcellular locations, or if the binding to several SH3 domains with variable affinity only illustrates “bystander” consequences of this unique binding mode. More exploration are needed to understand protein-protein interactions related to HIV-1 Nef–SH3 recognitions. Besides ALIX, the Y2H experiment has screened out the other five probable HIV-1 Nef-similar proteins. It is necessary to uncover the recognition mode between these cellular proteins and SH3 domain, and to discriminate the interface of these proteins and HIV-1 Nef onto various SH3 domains.

Reference

1. Arkin, M.R. and J.A. Wells, *Small-molecule inhibitors of protein-protein interactions: progressing towards the dream.* Nat Rev Drug Discov, 2004. **3**(4): p. 301-17.
2. Nooren, I.M. and J.M. Thornton, *Diversity of protein-protein interactions.* EMBO J, 2003. **22**(14): p. 3486-92.
3. Jones, S. and J.M. Thornton, *Principles of protein-protein interactions.* Proc Natl Acad Sci U S A, 1996. **93**(1): p. 13-20.
4. Gadek, T.R. and J.B. Nicholas, *Small molecule antagonists of proteins.* Biochem Pharmacol, 2003. **65**(1): p. 1-8.
5. Reynolds, C., D. Damerell, and S. Jones, *ProtorP: a protein-protein interaction analysis server.* Bioinformatics, 2009. **25**(3): p. 413-4.
6. Bahadur, R.P., et al., *A dissection of specific and non-specific protein-protein interfaces.* J Mol Biol, 2004. **336**(4): p. 943-55.
7. Lo Conte, L., C. Chothia, and J. Janin, *The atomic structure of protein-protein recognition sites.* J Mol Biol, 1999. **285**(5): p. 2177-98.
8. Bogan, A.A. and K.S. Thorn, *Anatomy of hot spots in protein interfaces.* J Mol Biol, 1998. **280**(1): p. 1-9.
9. DeLano, W.L., *Unraveling hot spots in binding interfaces: progress and challenges.* Curr Opin Struct Biol, 2002. **12**(1): p. 14-20.
10. Jackson, R.M., *Comparison of protein-protein interactions in serine protease-inhibitor and antibody-antigen complexes: implications for the protein docking problem.* Protein Sci, 1999. **8**(3): p. 603-13.
11. Pagliaro, L., et al., *Emerging classes of protein-protein interaction inhibitors and new tools for their development.* Curr Opin Chem Biol, 2004. **8**(4): p. 442-9.
12. Baltimore, D., et al., *A nuclear tyrosine kinase becomes a cytoplasmic oncogene.* Ann N Y Acad Sci, 1995. **758**: p. 339-44.
13. Mayer, B.J., M. Hamaguchi, and H. Hanafusa, *A novel viral oncogene with structural similarity to phospholipase C.* Nature, 1988. **332**(6161): p. 272-5.
14. Larson, S.M. and A.R. Davidson, *The identification of conserved interactions within the SH3 domain by alignment of sequences and structures.* Protein Sci, 2000. **9**(11): p. 2170-80.
15. STEFAN M. LARSON1 and ALAN R. DAVIDSON1, *The identification of conserved interactions within the SH3 domain by alignment of sequences and structures.* Protein Sci, 2000. **9**: p. *11*.
16. Mayer, B.J., *SH3 domains: complexity in moderation.* J Cell Sci, 2001. **114**(Pt 7): p. 1253-63.
17. Li, S.S., *Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction.* Biochem J, 2005. **390**(Pt 3): p. 641-53.

18. Musacchio, A., et al., *SH3--an abundant protein domain in search of a function.* FEBS Lett, 1992. **307**(1): p. 55-61.
19. Musacchio, A., et al., *Crystal structure of a Src-homology 3 (SH3) domain.* Nature, 1992. **359**(6398): p. 851-5.
20. Yu, H., et al., *Solution structure of the SH3 domain of Src and identification of its ligand-binding site.* Science, 1992. **258**(5088): p. 1665-8.
21. Noble, M.E., et al., *Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin.* EMBO J, 1993. **12**(7): p. 2617-24.
22. Arold, S., et al., *RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef.* Biochemistry, 1998. **37**(42): p. 14683-91.
23. Ren, R., et al., *Identification of a ten-amino acid proline-rich SH3 binding site.* Science, 1993. **259**(5098): p. 1157-61.
24. Rickles, R.J., et al., *Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries.* EMBO J, 1994. **13**(23): p. 5598-604.
25. Weng, Z., et al., *Identification of Src, Fyn, and Lyn SH3-binding proteins: implications for a function of SH3 domains.* Mol Cell Biol, 1994. **14**(7): p. 4509-21.
26. Alexandropoulos, K., G. Cheng, and D. Baltimore, *Proline-rich sequences that bind to Src homology 3 domains with individual specificities.* Proc Natl Acad Sci U S A, 1995. **92**(8): p. 3110-4.
27. Feng, S., et al., *Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands.* Proc Natl Acad Sci U S A, 1995. **92**(26): p. 12408-15.
28. Cicchetti, P., et al., *Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho.* Science, 1992. **257**(5071): p. 803-6.
29. Rickles, R.J., et al., *Phage display selection of ligand residues important for Src homology 3 domain binding specificity.* Proc Natl Acad Sci U S A, 1995. **92**(24): p. 10909-13.
30. Feng, S., et al., *Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions.* Science, 1994. **266**(5188): p. 1241-7.
31. Yu, H., et al., *Structural basis for the binding of proline-rich peptides to SH3 domains.* Cell, 1994. **76**(5): p. 933-45.
32. Musacchio, A., M. Saraste, and M. Wilmanns, *High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides.* Nat Struct Biol, 1994. **1**(8): p. 546-51.
33. Kaneko, T., L. Li, and S.S. Li, *The SH3 domain--a family of versatile peptide- and protein-recognition module.* Front Biosci, 2008. **13**: p. 4938-52.
34. Nguyen, J.T., et al., *Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors.* Science, 1998. **282**(5396): p. 2088-92.
35. Sparks, A.B., et al., *Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2.* Proc Natl

Acad Sci U S A, 1996. **93**(4): p. 1540-4.
36. Santiveri, C.M., et al., *Interaction between the N-terminal SH3 domain of Nck-alpha and CD3-epsilon-derived peptides: non-canonical and canonical recognition motifs.* Biochim Biophys Acta, 2009. **1794**(1): p. 110-7.
37. St-Arnaud, R. and I. Quelo, *Transcriptional coactivators potentiating AP-1 function in bone.* Front Biosci, 1998. **3**: p. d838-48.
38. Zhao, Z.S., E. Manser, and L. Lim, *Interaction between PAK and nck: a template for Nck targets and role of PAK autophosphorylation.* Mol Cell Biol, 2000. **20**(11): p. 3906-17.
39. Lee, C.H., Leung, B., Lemmon, M. A., Zheng, J., Cowburn, D., Kuriyan, J., Saksela, K., *A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein.* EMBO J, 1995. **14**(20): p. 5006-15.
40. Ladbury, J.E. and S. Arold, *Searching for specificity in SH domains.* Chem Biol, 2000. **7**(1): p. R3-8.
41. Kowanetz, K., et al., *CIN85 associates with multiple effectors controlling intracellular trafficking of epidermal growth factor receptors.* Mol Biol Cell, 2004. **15**(7): p. 3155-66.
42. Agrawal, V. and K.V. Kishan, *Promiscuous binding nature of SH3 domains to their target proteins.* Protein Pept Lett, 2002. **9**(3): p. 185-93.
43. Lewitzky, M., et al., *Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix.* J Biol Chem, 2004. **279**(27): p. 28724-32.
44. Seet, B.T., et al., *Efficient T-cell receptor signaling requires a high-affinity interaction between the Gads C-SH3 domain and the SLP-76 RxxK motif.* EMBO J, 2007. **26**(3): p. 678-89.
45. Liu, Q., et al., *Structural basis for specific binding of the Gads SH3 domain to an RxxK motif-containing SLP-76 peptide: a novel mode of peptide recognition.* Mol Cell, 2003. **11**(2): p. 471-81.
46. He, Y., L. Hicke, and I. Radhakrishnan, *Structural basis for ubiquitin recognition by SH3 domains.* J Mol Biol, 2007. **373**(1): p. 190-6.
47. Nishida, M., et al., *Novel recognition mode between Vav and Grb2 SH3 domains.* EMBO J, 2001. **20**(12): p. 2995-3007.
48. Felder, S., et al., *Kinase activity controls the sorting of the epidermal growth factor receptor within the multivesicular body.* Cell, 1990. **61**(4): p. 623-34.
49. Hubbard, S.R., *Structural analysis of receptor tyrosine kinases.* Prog Biophys Mol Biol, 1999. **71**(3-4): p. 343-58.
50. Hubbard, S.R. and J.H. Till, *Protein tyrosine kinase structure and function.* Annu Rev Biochem, 2000. **69**: p. 373-98.
51. Erpel, T. and S.A. Courtneidge, *Src family protein tyrosine kinases and cellular signal transduction pathways.* Curr Opin Cell Biol, 1995. **7**(2): p. 176-82.
52. Garrett, T.P., et al., *Crystal structure of the first three domains of the type-1*

insulin-like growth factor receptor. Nature, 1998. **394**(6691): p. 395-9.
53. Leppanen, V.M., et al., *Structural determinants of growth factor binding and specificity by VEGF receptor 2.* Proc Natl Acad Sci U S A. **107**(6): p. 2425-30.
54. Wiesmann, C., et al., *Crystal structure at 1.7 A resolution of VEGF in complex with domain 2 of the Flt-1 receptor.* Cell, 1997. **91**(5): p. 695-704.
55. Hubbard, S.R., et al., *Crystal structure of the tyrosine kinase domain of the human insulin receptor.* Nature, 1994. **372**(6508): p. 746-54.
56. Taylor, S.S. and E. Radzio-Andzelm, *Three protein kinase structures define a common motif.* Structure, 1994. **2**(5): p. 345-55.
57. Hunter, T. and B.M. Sefton, *Transforming gene product of Rous sarcoma virus phosphorylates tyrosine.* Proc Natl Acad Sci U S A, 1980. **77**(3): p. 1311-5.
58. Sicheri, F. and J. Kuriyan, *Structures of Src-family tyrosine kinases.* Curr Opin Struct Biol, 1997. **7**(6): p. 777-85.
59. Yeatman, T.J., *A renaissance for SRC.* Nat Rev Cancer, 2004. **4**(6): p. 470-80.
60. Superti-Furga, G. and S.A. Courtneidge, *Structure-function relationships in Src family and related protein tyrosine kinases.* Bioessays, 1995. **17**(4): p. 321-30.
61. Okada, M., et al., *CSK: a protein-tyrosine kinase involved in regulation of src family kinases.* J Biol Chem, 1991. **266**(36): p. 24249-52.
62. Ogawa, A., et al., *Structure of the carboxyl-terminal Src kinase, Csk.* J Biol Chem, 2002. **277**(17): p. 14351-4.
63. Roskoski, R., Jr., *Src protein-tyrosine kinase structure and regulation.* Biochem Biophys Res Commun, 2004. **324**(4): p. 1155-64.
64. Xu, W., S.C. Harrison, and M.J. Eck, *Three-dimensional structure of the tyrosine kinase c-Src.* Nature, 1997. **385**(6617): p. 595-602.
65. Sicheri, F., I. Moarefi, and J. Kuriyan, *Crystal structure of the Src family tyrosine kinase Hck.* Nature, 1997. **385**(6617): p. 602-9.
66. Stefan Arold1†, P.F., Marie-Paule Strub1, Francois Hoh1, Serge Benichou2, Richard Benarous2 and Christian Dumas1*, *The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling.* Structure 1997. **5**: p. 12.
67. Chi-Hon Lee, K.S., * Urooj A. Mirza,* and a.J.K. Brian T. Chait, *Crystal Structure of the Conserved Core of HIV-1 Nef Complexed with a Src Family SH3 Domain.* Cell, , 1996. **Vol. 85,**.
68. Collette, Y., et al., *HIV-2 and SIV nef proteins target different Src family SH3 domains than does HIV-1 Nef because of a triple amino acid substitution.* J Biol Chem, 2000. **275**(6): p. 4171-6.
69. Choi, H.J. and T.E. Smithgall, *Conserved residues in the HIV-1 Nef hydrophobic pocket are essential for recruitment and activation of the Hck tyrosine kinase.* J Mol Biol, 2004. **343**(5): p. 1255-68.
70. Betzi, S., et al., *Protein protein interaction inhibition (2P2I) combining high throughput and virtual screening: Application to the HIV-1 Nef protein.* Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19256-61.
71. Horita, D.A., et al., *Solution structure of the human Hck SH3 domain and*

identification of its ligand binding site. J Mol Biol, 1998. **278**(1): p. 253-65.
72. Deacon, N.J., et al., *Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients.* Science, 1995. **270**(5238): p. 988-91.
73. Fackler, O.T. and A.S. Baur, *Live and let die: Nef functions beyond HIV replication.* Immunity, 2002. **16**(4): p. 493-7.
74. Kestler, H.W., 3rd, et al., *Importance of the nef gene for maintenance of high virus loads and for development of AIDS.* cell, 1991. **65**(4): p. 651-62.
75. Missotten, M., et al., *Alix, a novel mouse protein undergoing calcium-dependent interaction with the apoptosis-linked-gene 2 (ALG-2) protein.* Cell Death Differ, 1999. **6**(2): p. 124-9.
76. Vito, P., et al., *Cloning of AIP1, a novel protein that associates with the apoptosis-linked gene ALG-2 in a Ca2+-dependent reaction.* J Biol Chem, 1999. **274**(3): p. 1533-40.
77. Chatellard-Causse, C., et al., *Alix (ALG-2-interacting protein X), a protein involved in apoptosis, binds to endophilins and induces cytoplasmic vacuolization.* J Biol Chem, 2002. **277**(32): p. 29108-15.
78. Chen, B., et al., *The glioma-associated protein SETA interacts with AIP1/Alix and ALG-2 and modulates apoptosis in astrocytes.* J Biol Chem, 2000. **275**(25): p. 19275-81.
79. Katoh, K., et al., *The ALG-2-interacting protein Alix associates with CHMP4b, a human homologue of yeast Snf7 that is involved in multivesicular body sorting.* J Biol Chem, 2003. **278**(40): p. 39104-13.
80. Strack, B., et al., *AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding.* cell, 2003. **114**(6): p. 689-99.
81. Vincent, O., et al., *YPXL/I is a protein interaction motif recognized by aspergillus PalA and its human homologue, AIP1/Alix.* Mol Cell Biol, 2003. **23**(5): p. 1647-55.
82. Zhai, Q., et al., *Structural and functional studies of ALIX interactions with YPX(n)L late domains of HIV-1 and EIAV.* Nat Struct Mol Biol, 2008. **15**(1): p. 43-9.
83. Zhou, X., et al., *The CHMP4b- and Src-docking sites in the Bro1 domain are autoinhibited in the native state of Alix.* Biochem J, 2009. **418**(2): p. 277-84.
84. Mirko H. H. Schmidt, I.D., and Oliver Bo¨ gler, *Src Phosphorylation of Alix/AIP1 Modulates Its Interaction with Binding Partners and Antagonizes Its Activities*.* JBC, 2005. **280,**(No. 5, Issue of February 4,): p. 12.
85. Matsuo, H., et al., *Role of LBPA and Alix in multivesicular liposome formation and endosome organization.* Science, 2004. **303**(5657): p. 531-4.
86. Schmidt, M.H., et al., *Alix/AIP1 antagonizes epidermal growth factor receptor downregulation by the Cbl-SETA/CIN85 complex.* Mol Cell Biol, 2004. **24**(20): p. 8981-93.
87. Odorizzi, G., *The multiple personalities of Alix.* J Cell Sci, 2006. **119**(Pt 15): p. 3025-32.
88. Robert D. Fisher, 3 Hyo-Young Chung,1,3 Qianting Zhai,1 Howard

Robinson,2 Wesley I. Sundquist,1,* and and Christopher P. Hill1, *Structural and Biochemical Studies of ALIX/AIP1 and Its Role in Retrovirus Budding.* cell, 2007. **128**: p. 12.
89. Lee, S., et al., *Structural basis for viral late-domain binding to Alix.* Nat Struct Mol Biol, 2007. **14**(3): p. 194-9.
90. McCullough, J., et al., *ALIX-CHMP4 interactions in the human ESCRT pathway.* Proc Natl Acad Sci U S A, 2008. **105**(22): p. 7687-91.
91. Fisher, R.D., et al., *Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding.* cell, 2007. **128**(5): p. 841-52.
92. Munshi, U.M., et al., *An Alix fragment potently inhibits HIV-1 budding: characterization of binding to retroviral YPXL late domains.* J Biol Chem, 2007. **282**(6): p. 3847-55.
93. Pires, R., et al., *A crescent-shaped ALIX dimer targets ESCRT-III CHMP4 filaments.* Structure, 2009. **17**(6): p. 843-56.
94. Morton, C.J., et al., *Solution structure and peptide binding of the SH3 domain from human Fyn.* Structure, 1996. **4**(6): p. 705-14.
95. Potterton, E., et al., *The CCP4 molecular-graphics project.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 11): p. 1955-7.
96. Vagin, A. and A. Teplyakov, *Molecular replacement with MOLREP.* Acta Crystallogr D Biol Crystallogr. **66**(Pt 1): p. 22-5.
97. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
98. Winn, M.D., M.N. Isupov, and G.N. Murshudov, *Use of TLS parameters to model anisotropic displacements in macromolecular refinement.* Acta Crystallogr D Biol Crystallogr, 2001. **57**(Pt 1): p. 122-33.
99. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions.* Methods Enzymol, 2003. **374**: p. 300-21.
100. DeLano, W.L., *The PyMOL Molecular Graphics System and Useers Manual.* San Carlos, CA, DeLano Scientific, 2002.
101. Lieutaud, P., B. Canard, and S. Longhi, *MeDor: a metaserver for predicting protein disorder.* BMC Genomics, 2008. **9 Suppl 2**: p. S25.
102. Grzesiek, S., et al., *The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase.* Nat Struct Biol, 1996. **3**(4): p. 340-5.
103. Chatellard-Causse, C., et al., *Alix (ALG-2-interacting protein X), a protein involved in apoptosis, binds to endophilins and induces cytoplasmic vacuolization.* J Biol Chem, 2002. **277**(32): p. 29108-15.
104. Che, S., et al., *Identification and cloning of xp95, a putative signal transduction protein in Xenopus oocytes.* J Biol Chem, 1999. **274**(9): p. 5522-31.
105. Schmidt, M., I. Dikic, and O. Bögler, *Src phosphorylation of Alix/AIP1 modulates its interaction with binding partners and antagonizes its activities.* J Biol Chem, 2005. **280**(5): p. 3414-25.

106. Guiet, R., et al., *Hematopoietic cell kinase (Hck) isoforms and phagocyte duties - from signaling and actin reorganization to migration and phagocytosis.* Eur J Cell Biol, 2008. **87**(8-9): p. 527-42.
107. Hassaïne, G., et al., *The tyrosine kinase Hck is an inhibitor of HIV-1 replication counteracted by the viral vif protein.* J Biol Chem, 2001. **276**(20): p. 16885-93.
108. Zhou, X., et al., *The CHMP4b- and Src-docking sites in the Bro1 domain are autoinhibited in the native state of Alix.* Biochem J, 2009. **418**(2): p. 277-84.
109. Briggs, S.D., E.C. Lerner, and T.E. Smithgall, *Affinity of Src family kinase SH3 domains for HIV Nef in vitro does not predict kinase activation by Nef in vivo.* Biochemistry, 2000. **39**(3): p. 489-95.
110. Collette, Y., et al., *Physical and functional interaction of Nef with Lck. HIV-1 Nef-induced T-cell signaling defects.* J Biol Chem, 1996. **271**(11): p. 6333-41.
111. Greenway, A.L., et al., *Simian immunodeficiency virus and human immunodeficiency virus type 1 nef proteins show distinct patterns and mechanisms of Src kinase activation.* J Virol, 1999. **73**(7): p. 6152-8.
112. Trible, R.P., L. Emert-Sedlak, and T.E. Smithgall, *HIV-1 Nef selectively activates Src family kinases Hck, Lyn, and c-Src through direct SH3 domain interaction.* J Biol Chem, 2006. **281**(37): p. 27029-38.
113. Fackler, O.T., et al., *Activation of Vav by Nef induces cytoskeletal rearrangements and downstream effector functions.* Mol Cell, 1999. **3**(6): p. 729-39.

Acknowledgments

I have great time during my past four years of PhD study. I am grateful to all the people who help me for my PhD project

First and foremost, I would like to thank my supervisors, Dr. Xavier Morelli and Prof. Mingdong Huang, for their constructiveadvices on my PhD project. I am deeply impressed by their profound professionism. They always inspired me and helped me greatly when I encoutered difficulties in my project. In addition, they provided me various opportunities, including study in France to expand my experience not only for science and but also for my life. Especially, I would like to thank Dr. Xavier Morelliand his PhD student Adrien Lugari for their help during my stay in France.

I am also grateful to our collaborators, Dr. Yves Collette , Dr Sandrine Opi and Audrey Restouin (INSERM, Marseille, France). They performed the cellular biology experiments. Their successful cellular biology experiment results facilitate my research article and my thesis. I also would like to express my heartfelt gratitude to Dr. Stefan T. Arold (INSERM, Montpellier, France) for his guidance onthe ITC and SAXS measurements. In addition, I would like to thank our other collaborators Dr. Pascale Zimmermann (Leuven, Belgium) and Dr. Javier Perez(Synchrotron SOLEIL – SWING, Paris, France), who provided me access to Biacore and synchrotron, and Dr. Isabelle Parrot (INSERM, Montpellier, France) for peptide synthesis, and Dr. Jacques Corbeil (Québec, Canada) for the use of interferometry. I also owe a special debt of gratitude to all the professors and collegues in FJIRSM and IMR for their kindly help and supports during my PhD life.

Finally, I am indebted to my family and friends for their understandings and supports on me in completing this thesis.

List of publications

Underlined are the research articles published during studying toward the PhD degree

1. **Xiaoli Shi**, Sandrine Opi, Adrien Lugari, Mingdong Huang, Stefan T. Arold, Yves Collette and Xavier Morelli; (2010), *Identification and biophysical assessment of the molecular recognition mechanisms between the human haemopoietic cell kinase Src homology domain 3 and ALG-2-interacting protein X.* Biochem J. 431(1): p. 93-102.
2. Shihui GUO, **Xiaoli SHI**, Feng YANG and Mingdong HUANG, (2009), *Structural basis of transport of lysophospholipids by human serum albumin*, Biochem.J. 423, 23-40.
3. **SHI Xiao-Li**, ZENG Tu, HUANG Ming-Dong, (2008), *Crystal Structure of the Fab Fragment of an Anti-factor IX Antibody 10C12*, *Chinese J. Struct. Chem.* Vol. 27, No.8.
4. **Xiaoli Shi**, Yongdong Li, Chuanbing Bian, Gengxiang Zhao and Mingdong Huang. Crystallization of an anti-factor IX antibody and its complex, 2005, Acta Cryst. D61, 701-703
5. **Shi Xiao-li,** Zhu Li-li, Yuan Cai, Zhang Li-mei, Chen Jin-can, Huang Ming-dong, Functions and Structure of Tumstatin: a Tumor Growth Suppressor, 2006, Chinese J. Cancer Research on Prevention and Treatment, 33, 377-379
6. Yongdong Li, **Xiaoli Shi**, GrahamParry, Liqing Chem, Jennifer A. Callahan, Andrew P. Mazar and Mingdong Huang. Optimization of crystals of an inhibitory antibody of urokinase plasminogen activator receptor (uPAR)with Hydrogen peroxide and low protein concentration, 2005, Protein and peptide letters, **12, 655-658**
7. CHEN Ming-Huang, **SHI Xiao-Li,** YE Xiao-Ming,XIE Jie-Ming,ZHAO Rong,RAO Ping-Fang and WANG Zi-Ren, Study on the secondary structure and bioactivities of Luffaculin 1, 2004, Chinese J. Struct. Chem, 23 , 232-235

8. **Xiaoli Shi**, Edward X. Zhou, YE Xiao-Ming, Edward J. Meehan, Minghuang Chen, and Liqin Chen. Molecular Replacement Studies Of Cucurmosin from Cucurbita Moschata: Structure Homology with Trichosanthin, 2003, Chinese J. Struct. Chem., 22, 165-168
9. Zhao GX , Yuan C , Bian CB , Hou XM , **Shi XL** , Ye XM , Huang ZX , Huang MD Protein expression and preliminary crystallographic analysis of amino-terminal fragment of urokinase-type plasminogen activator PROTEIN EXPRESSION AND PURIFICATION 49 (1): 71-77 2006
10. Minghuang Chen, **Xiaoli Shi,** Xiao-Ming Ye, Li-Qing Chen, Structure-function studies of Cucurmosin, 2003, Toxicology, 191 (1) 48
11. Minghuang Chen, Xiaoming Ye, **Xiaoli shi**, Ziren Wang, Luffaculins, ribosome-inactivating protein from the seeds of Luffa acutangula, 2003, Toxicology 191 (1) 48
12. ChuanBing Bian, Cai Yuan, Lin Lin, **XiaoLi Shi**, XiaoNing Ye, ZiXiang Huang and MingDong Huang, Purification and preliminary crystallographic analysis of a Penicillium expansum lipase, 2005, Biochimica et Biophysica Acta, 1752, 99-102
13. LIN Jun-kai, CHEN Ming-Huang, XIE Jie-Ming , ZHAO Rong, YE Xiao-Ming, **SHI Xiao-Li** and WANG Zi-Ren,Purication and Characterization of Two Luffaculins, Ribosome-inactivating Proteins from Seeds of Luffa acutangula, 2002, Chinese J. Biochemistry and Molecular Biology, 18, 609-613

Under preparation:

1. **Xiaoli Shi**, Sandrine Opi, Adrien Lugari, Mingdong Huang, Stefan T. Arold, Yves Collette and Xavier Morelli; *Searching for HIV Nef-like tertiary SH3 recognition by combining proteomic and structural biology analysis;* submitted to JBC,
2. Chunyan Li,Guanyu Ruan , Yu Xiao, Shengxiong Dong, **Xiaoli Shi**, Jincan Chen, Mingdong Huang ; Direct interaction of photosensitizer pentalysine β-carbonyl-phthalocyanine zinc with human serum albumin

Curriculum Vitae

Name: Xiaoli Shi

Sex: Female

Date of Birth: 30/06/1978

Address : Structural biology group

State Key Laboratory of Structural Chemistry

Fujian Institute of Research on the Structure of Matter

Chinese Academy of Sciences

155 Yan Qiao Xi Lu, Fuzhou, Fujian 350002 P. R. China

Tel: 86-591-83705814 (O)

Email: sharly78@gmail.com ; sxl_sharly@fjirsm.ac.cn

Education

2007.3-2010.11 : Ph.D.

Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, FuZhou, P. R. China

Major : Molecular biology and Biochemistry

2008.3-2011.2 : Ph.D.

Institut de Microbiologie de la Méditérannée (IMM), CNRS, France

Major : Molecular biology and Biochemistry

2002.1-2005.12: Master

Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, FuZhou, P. R. China

Major : Physical Chemistry

1997.9-2001.7 : B.S.

Department of Biology, Xiamen University, Xiamen, P. R. China

Major : Biology

Research experience

07/2001-current : Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, FuZhou, P. R. China

Project: Cloning, Expression, Purification and Crystallization of several proteins related with cancer