MARITA HIIPAKKA

Ligand Recognition in SH3-mediated Protein Interactions

■

ACADEMIC DISSERTATION

To be presented, with the permission of
the Faculty of Medicine of the University of Tampere,
for public discussion in the auditorium of Finn-Medi 1,
Biokatu 6, Tampere, on April 23rd, 2005, at 12 o'clock.

Acta Universitatis Tamperensis 1070

ACADEMIC DISSERTATION
University of Tampere, Institute of Medical Technology
Finland

Supervised by
Professor Kalle Saksela
University of Helsinki

Reviewed by
Professor Kari Keinänen
University of Helsinki
Docent Jari Ylänne
University of Oulu

Distribution
Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 3 215 6055
Fax +358 3 215 7685
taju@uta.fi
www.uta.fi/taju
http://granum.uta.fi

Cover design by
Juha Siro

Printed dissertation
Acta Universitatis Tamperensis 1070
ISBN 951-44-6262-9
ISSN 1455-1616

Electronic dissertation
Acta Electronica Universitatis Tamperensis 424
ISBN 951-44-6263-7
ISSN 1456-954X
http://acta.uta.fi

Tampereen Yliopistopaino Oy – Juvenes Print
Tampere 2005

To my family

CONTENTS

1. ABSTRACT

Src homology 3 (SH3) domains are small (typically comprising of 50-70 amino acids) that were first identified as regions of homology among the Src family tyrosine kinases. They are modular protein units, thus capable of functioning on their own, found in various cellular proteins including enzymes, adaptors and structural proteins. These domains mediate inter- and intramolecular protein interactions by binding to ligands that contain a region with a secondary structure called polyproline type II (PPII) helix. Sequence variation of the PPII helical region of the ligand can affect the specificity of ligand recognition by SH3 domains. Moreover, studies on HIV-1 Nef and later on other proteins have shown that additional contacts outside the SH3 binding interface also contribute to the binding affinity and specificity. Especially an SH3 region called the RT-loop can provide additional determinants for ligand binding by an SH3 domain.

In this study we have examined the role of residues of Nef implicated by structural studies to be involved in SH3 binding or other protein interactions in the capacity of Nef bind to the SH3 domain of the tyrosine kinase Hck *in vitro* and to associate with the serine/threonine kinase PAK2 in cells.

The strength of binding to Hck SH3 agreed well with predictions based on structural data, but also correlated with association of Nef with PAK2. Since PAK2 does not contain an SH3 domain, an SH3-containing adaptor protein therefore seems critical for bridging Nef and PAK2.

To further examine the role of the RT-loop region in binding of Hck SH3 to Nef and to evaluate RT-loop's importance as a general affinity/specificity determinant in SH3-mediated protein interactions a large library of artificial SH3 domains (RRT-SH3) was generated by replacing six variable residues in the "tip" of the RT-loop of Hck SH3 with random amino acids.

RRT-SH3 domains with highest affinity to Nef were selected using phage display and shown to bind HIV-1 Nef more than 40-fold better than unmodified Hck SH3. In some but interestingly not all cases the increased affinity provided by the engineered RT-loop residues depended on the same surface determinants in Nef contacted by the RT-loop of the wild-type Hck SH3. RRT-SH3 domains binding tightly to Nef protein variants that normally bind poorly if at all to Hck SH3 (SIV239mac Nef and F90R mutant of HIV-1 Nef) could also be generated, suggesting that this approach might be generally useful for targeting diverse SH3 ligand proteins with high-affinity. Unpublished data on several non-Nef proteins presented in this thesis provide support for this hypothesis.

We studied the potential of the Nef-targeted RRT-SH3 domains to function as cellular inhibitors of Nef. Our results showed that the RRT-SH3 domains efficiently associated with Nef in cells and could inhibit critical cellular functions of Nef. Most of the natural Hck SH3 substrates did not associate with these RRT-SH3 domains in cells implicating that the modified RT-loops ensure the specific binding to Nef with a concomitant loss to other natural partners of Hck SH3. In conclusion, these ligand-tailored SH3 domains could serve as tools both in research and in therapeutic applications as specific intracellular inhibitors of SH3-mediated interactions.

2. LIST OF ORIGINAL COMMUNICATIONS

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals. The original publications I-III are reprinted with permission from Elsevier and the original publication IV with permission from the Society for General Microbiology.

I Manninen A[1], Hiipakka M[1], Vihinen M, Lu W, Mayer BJ, Saksela K. (1998) SH3-Domain binding function of HIV-1 Nef is required for association with a PAK-related kinase. Virology 250:273-82.

II Hiipakka M, Poikonen K, Saksela K. (1999) SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef. J Mol Biol. 293:1097-106.

III Hiipakka M, Huotari P, Manninen A, Renkema GH and Saksela K (2001) Inhibition of cellular functions of HIV-1 Nef by artificial SH3 domains. Virology 286:152-159.

IV Hiipakka M and Saksela K (2002) Capacity of simian immunodeficiency virus strain mac Nef for high-affinity Src homology 3 (SH3) binding revealed by ligand-tailored SH3 domains. J Gen Virol. 83:3147-3152.

[1] These authors contributed equally to this work.

3. ABBREVIATIONS

The standard one-letter abbreviations are used for amino acids. Point mutations are labelled like P69A, in which proline 69 is mutated to alanine.

AIDS	Acquired immunodeficiency syndrome
AP-1	Activator protein-1
AP	Alkaline phosphatase
ARRE2	Antigen receptor response element of the interleukin-2 gene
AMP	Adenosine 5'-monophosphate
ATP	Adenosine 5'-triphosphate
BSA	Bovine serum albumin
cDNA	Complementary DNA
DTT	Dithiotreitol
ECL	Enhanced chemiluminescence
EVH1	Ena/VASP homology 1
FACS	Fluorescence activated cell sorter
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
GST	Glutathione-S-transferase
Hck	Hemopoietic cell kinase
HIV	Human immunodeficiency virus
IL-2	Interleukin-2
IVKA	*In vitro* kinase assay
MBP	Maltose binding protein
MHC	Major histocombatibility complex
NBT/BCIP	Nitro blue tetrazolium chloride/5-Bromo-4-chloro-3-indolyl phosphate toluidine salt, a substrate for AP
NFAT	Nuclear factor of activated T-cells
NMR	Nuclear magnetic resonance
PAK	p21 activated kinase
PE	Phycoerythrin
PH	Pleckstrin homology
PMA	Phorbol 12-myristate, 13-acetate
PMSF	Phenylmethylsulfonyl fluoride
PPII	Polyproline type two (helix)
PRR	Proline-rich region
PTB	Phosphotyrosine binding (domain)
RRT	Randomised RT-loop
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH2	Src homology 2 (domain)
SH3	Src homology 3 (domain)
SIV	Simian immunodeficiency virus
SPR	Surface plasmon resonance
TCR	T-cell receptor
UEV	Ubiquitin E2 variant

4. INTRODUCTION

Protein-protein interactions govern cellular physiology by assembling multiprotein complexes and pathways that are critical in coordinating specific biochemical functions. Protein complexes often involve diverse peptide recognition modules that have evolved to recognise ligands containing a core structural motif. SH3 domains were among the first protein modules to be identified and are now found from a large number of proteins from yeasts to mammals (Cohen 1995, Sudol 1998). They often serve functions in regulation of cell growth or differentiation thus being involved in pathogenesis of diseases such as cancer. Moreover, many microbial pathogens, like HIV-1, exploit the cellular SH3-mediated processes (Saksela et al. 1995).

SH3 domains are small (50-70 amino acids long) conserved protein folds with three variable loop structures. They bind ligands that contain regions that adopt a polyproline type II (PPII) helical structure. Most of the natural SH3 interactions are weak, (K_d values on the micromolar range) but sequence variation in the PPII helical region of the ligand has been shown to improve the specificity (and affinity) of the SH3/ligand complex formation (Mayer 2001; Musacchio 2002). However, secondary structures as well as molecular contacts outside the SH3-binding interface have been shown to contribute to the specificity even more. Especially in the case of the HIV-1 Nef/Hck SH3 interaction in addition to the canonical PPII helical motif, relatively distant non-continuous region is shown to bind an SH3 structure known as the RT-loop with unusually high affinity and specificity (Lee et al. 1995, Lee et al. 1996).

The purpose of this work was to study the SH3 binding capacity of HIV-1 Nef and see the possible correlation of this to the ability of Nef to associate with a cellular serine/threonine kinase PAK2. Another goal was to exploit the central role of the SH3 RT-loop in ligand recognition by selecting phage display libraries of artificial SH3 domains with Nef proteins and further assess their role as intracellular inhibitors of Nef.

5. REVIEW OF THE LITERATURE

5.1. Signal transduction in eukaryotic cells

A multicellular organism is critically dependent on inter- and intracellular communication. For example, cell growth, differentiation, movement and apoptosis are events that are governed by signals that cells receive from their environment either by direct cell-cell contacts or by a variety of extracellular stimuli, including cytokines, hormones and growth factors (Touw et al. 2000). In many cases these factors exert their effects by binding to specific cell surface receptors that transmit the signal inside the cell (or by traversing the plasma membrane and directly binding to internal receptors as is the case with steroids; Pawson and Nash 2000). Several types of cell surface receptors have been identified. They recognise various signalling molecules by their extracellular domains. Binding of the ligand activates a cascade of events that typically ends in the nucleus where the transcription of specific genes takes place. The sum of these actions determines the ultimate response of the cell to a specific stimulus (Pawson and Nash 2000).

As many of the signals are transmitted via transmembrane receptors the plasma membrane serves as a meeting point for different signalling molecules. Many of them are associated with the plasma membrane either directly or indirectly. A direct membrane association of a cytoplasmic protein can occur, for example, via posttranslational modification with fatty acids that include an irreversible N-myristoylation (Farazi et al. 2001) or reversible palmitoylation (Smotrys and Linder 2004). Indirect membrane localisation, on the other hand, can be achieved via specific protein-protein interactions that can be regulated via a covalent modification like protein (or lipid) phosphorylation by specific enzymes called kinases. Phosphorylation, together with dephosphorylation (by phosphatases) plays an important role in regulating many of the signal transduction events like transcription or protein-protein interactions where, for example phosphorylated tyrosine residues serve as docking sites for specific protein domains (Fischer 1999; Holmberg et al. 2002, Pawson 2004). In eukaryotes, protein kinases phosphorylate specific serine, threonine or tyrosine residues within the protein. It has been estimated that the human genome contains 518 putative protein kinase genes representing almost 2 % of all human genes (Manning et al. 2002).

5.2. Small domains in signal transduction

Inside the cells the signals can be transduced, for example, by small molecules (second messengers like cyclic AMP or Ca^{2+}) or by complex networks of interacting proteins (Cohen et al. 1995; Hunter 1997; Sauro and Kholodenko 2004). These signalling proteins are often composed of series of independent modules, domains, which define the localisation or activity of a protein bringing the interacting partners together (Pawson and Nash 2003). Signalling domains are compact units that are able to function and maintain their structures also in isolation (Cohen et al. 1995). They are not restricted to any particular kind of proteins but are found, for example, in protein and lipid kinases, protein phosphatases and various adaptor proteins. Location and number of the domains between different proteins may vary (Cohen et al. 1995). The SMART database (http://smart.embl-heidelberg.de/; Letunic et al. 2004) currently (Mar -05) lists 153 different signalling domains. Some of them recognise, for example, phosphorylated tyrosine residues (SH2 and PTB domains) whereas others mediate interactions with

proline rich sequences (like SH3 and WW domains). The use of small homology domains has allowed the nature to generate similar signalling pathways for different proteins. It also ensures crosstalk between signalling molecules allowing them to act in concert to form cell regulatory systems and signalling networks (Ponting and Russell 2002).

5.3. Proline-rich binding motifs

Proline-rich regions (PRRs) are found in both prokaryotic and eukaryotic proteins. They can be classified to repetitive short PRRs that are involved in binding and structural processes, tandemly repeated PRRs like salivary PRRs having polyphenol binding properties and nonrepetitive PRRs that are involved in a wide variety of cellular processes (Williamson 1994) that include, for example, signal transduction (Kay et al. 2000), cellular immune response (Jardetzky et al. 1996), cell motility (Mahoney et al. 1999) or transcription (Sudol et al. 2001).

5.3.1. Properties of proline and polyproline sequences

Proline has unique features that make it special among the twenty natural amino acids. It is the only imino acid in mammals meaning that the side chain is cyclised onto the backbone amide position forming a cyclic pyrrolidine ring by N-substitution (Williamson 1994). This makes the conformation of proline very restricted; the Φ torsion angle is constrained to -65° (±15°). The bulkiness of the N-substituent also restricts the conformation of the residue preceding the proline. In contrast to natural all-*trans* peptide bonds, proline is also able to form stabile *cis* peptide bonds, occurring to more than 5 % in the globular proteins (Williamson 1994).

The left-handed polyproline II (PPII) helix similar to fibrous collagen is believed to be the dominant conformation in polyproline structures (Adzhubei and Sternberg 1993). However, the PPII helices are not necessarily formed exclusively by proline (Pisabarro and Serrano 1996), and the structure can be adopted also by sequences that totally lack proline (Adzhubei and Sternberg 1993). However, proline predominates in these structures and glutamine is the other significantly favoured amino acid, most probably because of its ability to form a side chain hydrogen bond to a backbone oxygen (Kelly et al. 2001; Chellgren and Creamer 2004). Glycine and nonpolar amino acids are disfavoured probably because of the preference for turns and high surface exposure of the PPII helices, respectively (Stapley and Creamer 1999). In addition to the collagen triple helix structure, PPII helices are common in globular proteins (Adzhubei and Sternberg 1993; Williamson 1994), peptides bound to the major histocombatibility complex II (Jardetzky et al. 1996), and they may also have significance in protein unfolded states (Pappu and Rose 2002).

PPII helix is an extended structure with three residues per turn and approximately triangular in cross-section (see Figure 3.). PPII helices are short, most just four residues long. They may serve as a bridge connecting two different or similar secondary structures (Adzhubei and Sternberg 1993; Stapley and Creamer 1999). Structural information of the long PRRs in proteins is scarce. However, in many cases small proline-rich peptides that are ligands for SH3 domains are relatively unstructured before binding to their substrates (Viguera et al. 1994; Dong et al. 2000). On the other hand, a functionally important region in the HIV-1 Nef protein adopts a PPII helical structure also when not being complexed to the SH3 ligand (Grzesiek et al. 1996). When

occurring in proteins, the PPII helices do not concentrate on certain domains but are distributed evenly in the protein structure, located mostly in surface areas (Adzhubei and Sternberg 1993). Since the proline-rich regions are exposed, the on and off rates of binding can be very fast. This makes them perfect mediators of interactions that require rapid recruitment of interchange of different proteins. Indeed, in addition to the aforementioned SH3 and WW domains recognising proline rich sequences, several other domains exerting similar preference for proline motifs have evolved, including Ena/VASP homology 1 (EVH1; reviewed in Ball et al. 2002), GYF (Nishizawa et al. 1998; Freund et al. 1999), Ubiquitin E2 variant (UEV; Pornillos et al. 2002) domains and profilin proteins (Mahoney et al. 1999).

In the PPII helix structure prolines form a continuous hydrophobic strip around the surface of the helix. Proline cannot act as a hydrogen bond donor but the backbone carbonyls can serve as intermolecular hydrogen bonding sites (acceptors) being both conformationally restricted and electron rich. (Kay et al. 2000; Zarrinpar et al. 2003a). Because of the restricted mobility (and therefore a relatively low entropy) of PPII polypeptides the binding of a ligand causes only a small drop in entropy compared to a more flexible peptide. Another feature apart from the entropy advantage that makes proline as a good ligand is the large flat hydrophobic surface that easily binds other hydrophobic surfaces such as aromatic rings (Williamson 1994) where high complementary of binding surfaces are not needed (Kay et al. 2000).

5.4. SH3 domains

5.4.1. Discovery and occurrence of SH3 domains

Src homology 3 (SH3) domains are the first characterised proline recognition modules (reviewed in Mayer 2001; Musacchio 2002). They are conserved from yeasts to mammals suggesting for a specific role in higher eukaryotes although there are reports of distantly related SH3 domains from bacteria as well (Whisstock and Lesk 1999; Bakal and Davies 2000). SH3 domains are one of the most common modules in eukaryotic genomes. For example, total of 63 SH3 domains were found in the fruit fly *D. melanogaster* genome (Rubin et al. 2000), 26 for the yeast *S cerevisiae* (The SMART database; http://smart.embl-heidelberg.de/; Letunic et al. 2004) and around 300 SH3 domains in the human genome (Saksela et al. unpublished).

SH3 domains are relatively short, comprising about 50-70 amino acid residues. They were first found as independent regions of similarity to the N-terminus of Src nonreceptor tyrosine kinase when the phospholipase C γ (PLCγ) and Crk oncogene were cloned (Mayer et al. 1988; Stahl et al. 1988). Since they were recognised in different types of proteins without an apparent enzymatic activity, it became clear that the found homology region was modular, thus able to function independently in different protein complexes (phospholipase versus tyrosine kinase). It was also evident that some Src mutations in this region were transforming, implicating a role for the SH3 in regulating the Src kinase activity (Kato et al. 1986). Soon after the first reports, other SH3 containing proteins were discovered, for example from proteins involved in cellular signalling pathways or from cytoskeletal proteins (Lehto et al. 1988; Otsu et al. 1991). They also have been found in different proteins involved in cytoskeletal dynamics (Hing et al. 1999; Buday et al. 2002) and endocytosis (reviewed in McPherson 1999). Examination, for example, of the SMART database (http://smart.embl-heidelberg.de/) reveals that many of the proteins have more than one SH3 domain and the SH3 domains

are mostly flanked by other protein modules, like SH2 (Src homology 2) or PH (pleckstrin homology) domains, but never with WW domains (also recognising proline rich regions), see Figure 1. One exception of this may be the yeast open reading frame of 103 amino acids that is a protein consisting almost entirely of one SH3 domain (accession number Q6B0T5 in SMART and AY693345 in GenBank database). In many instances the SH3 domains are accompanying catalytic domains in proteins such as the kinase domain of the Src family of non-receptor tyrosine kinases. Many of the SH3 domains are found in adaptor proteins that consist only of modular domains, like Grb2 that contains two SH3 domains and one SH2 domain and SETA and its related proteins (Borinstein et al. 2000) that have three SH3 domains. In addition, many SH3 domains are also found in proteins containing domains that regulate the activity of GTP-binding proteins, such as βPIX containing a Rho GEF (guanine-nucleotide exchange factor) domain and phosphatidylinositol-3' kinase (PI3K) p85 subunit containing a Rho GAP (GTPase activating protein) domain. The abundance of SH3 domains may have something to do with the protein recognition and evolution of the complex signalling pathways. It might also provide a regulatory mechanism by competition of certain binding sites between different SH3 domains.

5.4.2. Examples of SH3 function

Src family kinases are kept in an inactive state via intermolecular interactions involving the SH3 domain among others. The activation can be achieved by SH3 domain displacement. HIV-1 Nef protein is a high affinity ligand for Hck SH3 and its binding to the SH3 domain was shown to increase the Hck catalytic activity (Moarefi et al. 1997). Also Sin (Src interacting protein, a.k.a. Efs) was shown to bind to Src SH3 domain and thereby increase the Src activity (Alexandropoulos and Baltimore 1996) by SH3 domain displacement (Pellicena and Miller 2001).

SH3 domain binding may affect the catalytic activity of interacting proteins. Engagement of antigen receptor complexes results in activation of Src family kinases and association of PI3K. It was shown that Src family members Lyn and Fyn were capable of SH3 domain-mediated binding to a proline-rich region of p85 subunit of the PI3K and stimulated the PI3K activity 5 to 7-fold. The interaction could be blocked by peptides binding to SH3 domains (Pleiman et al. 1994). In another study Src family SH3 domains as well as a subset of other SH3 domains like the ones from p85 and Grb2 were able to bind to GTPase dynamin and stimulate the GTPase activity *in vitro* (Gout et al. 1993).

SH3 domain has been described to contribute to localisation of Src tyrosine kinase. In the inactive form Src kinase was found associated only with endosomal membranes in detergent soluble fraction. Activation resulted in increase in kinase activity as well as the translocation of Src to a detergent insoluble fraction, particularly in focal adhesions (Kaplan et al. 1994; Kaplan et al. 1995). Mutational studies showed that variants with alteration or removal of the regulatory Y527 residue (that promotes the active “open” conformation of Src) were translocated in focal adhesions. Mutation on the kinase ATP-binding site combined with the Y527 mutation resulting in a variant having an open conformation but inactive kinase also translocated in focal adhesions suggesting this phenomenon to be independent on the kinase activity. Likewise, a variant containing only the amino terminal half of the molecule translocated in focal adhesions. An intact SH3 domain was required for the association suggesting that protein-protein interactions of the SH3 domain induce Src to associate with focal adhesions.

The SH3 domain of the p85α regulatory subunit of PI3K has been shown to partially unfold in acidic pH and slowly aggregate into amyloid fibrils (Guijarro et al. 1998). However, this protein has not been associated with a disease involving amyloid formation, like Alzheimer's disease, but was rather suggesting a more general role of globular proteins to be involved in amyloidogenesis and providing a study model of such a phenomenon.

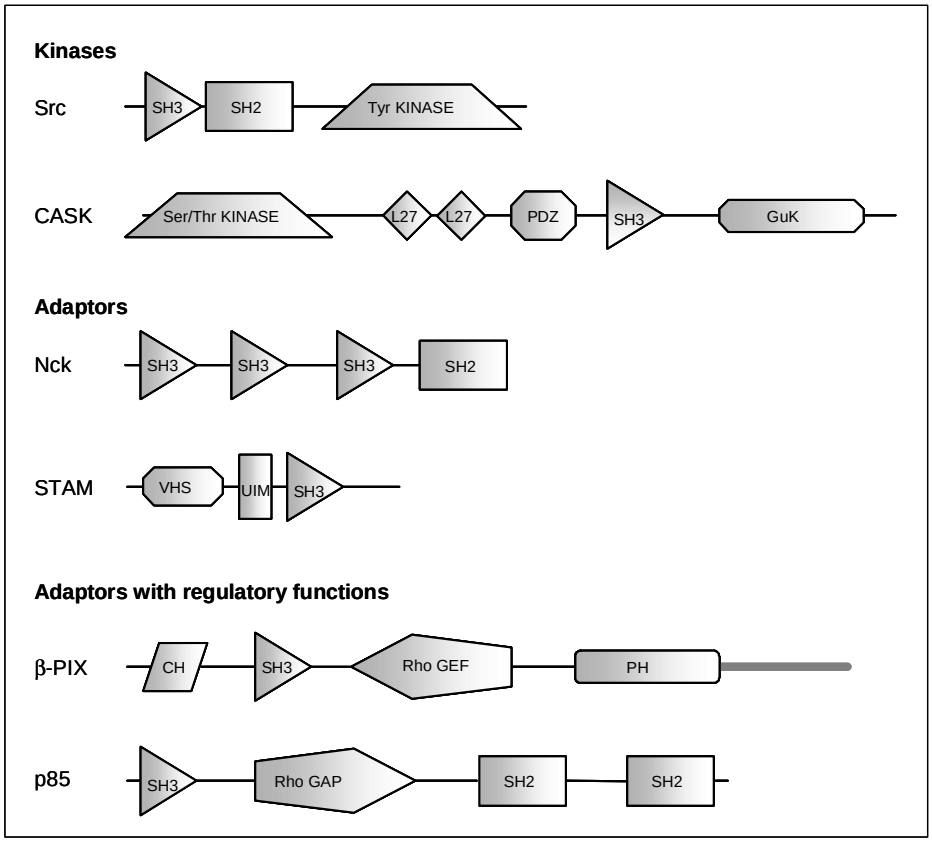


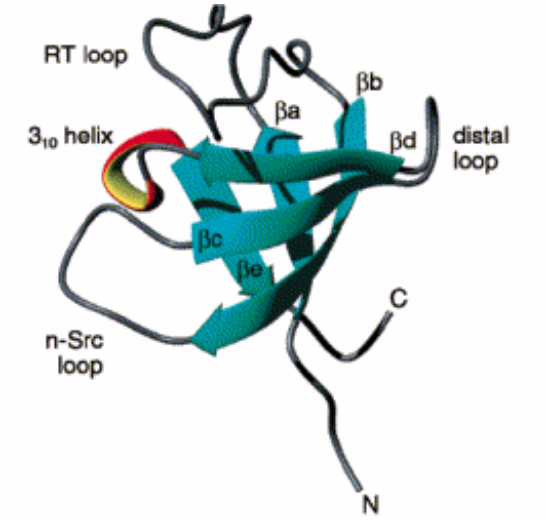
Figure 1.
Examples of SH3 domain-containing proteins. Abbreviations: SH3, Src homology 3; SH2; Src homology 2, Tyr, tyrosine; Ser, serine; Thr, threonine, L27, domain in receptor targeting proteins Lin-2 and Lin-7; PDZ, domain present in PSD-95, Dlg and ZO-1/2; GuK, guanylate kinase; VHS, domain present in VPS-27, Hrs and STAM; UIM, ubiquitin-interacting motif; CH, calponin homology; PH, pleckstrin homology, Rho GEF, guanine nucleotide exchange factor for Rho/Rac/cdc42-like GTPases; Rho GAP, GTPase-activator protein for Rho-like GTPases.

5.4.3. SH3 structure

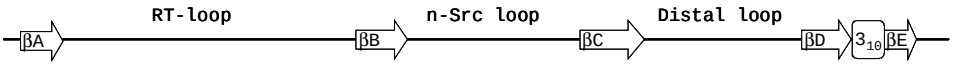
The first SH3 structures were solved in the beginning of 1990's (Musacchio et al. 1992; Yu et al. 1992; Noble et al. 1993) showing a conserved fold (Figure 2a.). To date structures of more than 30 SH3 domains, either non-liganded or complexed with a substrate, have been solved (for some of them several structures are available). These structures show that the SH3 domain fold is composed of five β strands (β_A-β_E) and a single turn of 3_{10} helix. The five β strands are arranged to two sheets that are in right angles relative to one another. The first sheet is composed of antiparallel runs of β_A, one half of β_B and β_E. The second sheet is composed of the remaining of β_B, β_C and β_D. The strands β_A and β_B are connected by the RT-loop (named after the mutations of R95 and T96 residues in Src that cause a deregulated kinase activity and transforming ability), β_B

and β_C are separated by the n-Src loop (named after the six amino acid insertion in the neuronal isoforms of Src) and the β_C and β_D by the distal loop (that is distal, on the opposite site of the SH3 ligand binding site). The strands β_D and β_E are separated by a 3_{10} helix.

a.



b.




```
KYVRALYDYEA------REDDELSFKKGDIQIVLEKS---------DDGWWKGRLN--------DTGREGLFPSNYVEEID  Consensus

TTFVALYDYES-------RTETDLSFKKGERLQIVNNT---------EGDWWLAHSL--------STGQTGYIPSNYVAPSD  Src
TTFVALYDYES-------RTETDLSFKKGERLQIVNNT---RKVDVREGDWWLAHSL--------STGQTGYIPSNYVAPSD  N-Src
IIVVALYDYEA-------IHHEDLSFQKGDQMVVLEES----------GEWWKARSL--------ATRKEGYIPSNYVARVD  Hck
DIVVALYPYDG-------IHPDDLSFKKGEKMKVLEEH----------GEWWKAKSL--------LTKKEGFIPSNYVAKLN  Lyn
TLFVALYDYEA-------RTEDDLSFHKGEKFQILNSS---------EGDWWEARSL--------TTGETGYIPSNYVAPVD  Fyn
NLVIALHSYEP-------SHDGDLGFEKGEQLRILEQS----------GEWWKAQSL--------TTGQEGFIPFNFVAKAN  Lck
NLFVALYDFVA-------SGDNTLSITKGEKLRVLGYN--------HNGEWCEAQTK----------NGQGWVPSNYITPVN  Abl
KKVVALYDYMP-------MNANDLQLRKGDEYFILEES---------NLPWWRARDK---------NGQEGYIPSNYVTEAE  Btk
IYVRAQFEYDPAKDDLIPCKEAGIRFRVGDIIQIISKD---------DHNWWQGKLE------NSKNGTAGLIPSSELQEWR  Cask
CAVKALFDYKA-------QREDELTFIKSAIIQNVEKQ---------EGGWWRGDYG---------GKKQLWFPSNYVEEMV  PLC-γ1
YQYRALYDYKK-------EREEDIDLHLGDILTVNKGSSDGQEARPEEIGWLNGYNE--------TTGERGDFPGTYVEYIG  p85A
EYVRALFDFNG-------NDEEDLPFKKGDILRIRDKP---------EEQWWNAEDS---------EGKRGMIPVPYVEKYR  Crk #1
MPAYVKFNYMA-------EREDELSLIKGTKVIVMEKC---------SDGWWRGSYN----------GQVGWFPSNYVTEEG  Nck1 #2
QLARALYDNTA-------ESPQELSFRRGDVLRVLQRE-----GAGGLDGWCLCSLH----------GQQGIVPANRVKLLP  Efs/Sin
RKVRAIYDFEA-------AEDNELTFKAGEIITVLDDS---------DPNWWKGETH----------QGIGLFPSNFVTADL  STAM1
YKVETLHDFEA-------ANSDELTLQRGDVVLVVPSD----SEADQDAGWLVGVKESDWLQYRDLATYKGLFPENFTRRLD  Amphiphysin
LVVRAKFNFQQ-------TNEDELSFSKGDVIHVTRVE---------EGGWWEGTLN----------GRTGWFPSNYVREVK  β-PIX
RWARALYDFEA-------LEDDELGFHSGEVVEVLDSS---------NPSWWTGRLH----------NKLGLFPANYVAPMT  GADS #2
TYVQALFDFDP-------QEDGELGFRRGDFIHVMDNS---------DPNWWKGACH----------GQTGMFPRNYVTPVN  Grb2 #2
PKMEVFQEYYG-LPPPPGAIGPFLRLNPGDIVELTKAE--------AEQNWWEGRN--------TSTNEIGWFPCNRVKPYV  Vav1 #1
KYAKSKYDFVA-------RNNSELSVLKDDILEILDDR----------KQWWKVRNA---------SGDSGFVPNNILDIVR  Eps8
GVIYALWDYEP-------QNDDELPMKEGDCMTIIHRE------DEDEIEWWWARLN----------DKEGYVPRNLLGLYP  53BP2
SYYQGLWDCHG-------DQPDELSFQRGDLIRILSKE-------YNMYGWWVGELN----------SLVGIVPKEYLTTAF  SKAP55
VEAIVEFDYQA-------QHDDELTISVGEIITNIRKE---------DGGWWEGQIN----------GRRGLFPDNFVREIK  CIN85 #1
SQVEALFSYEA-------TQPEDLEFQEGDIILVLSKV---------NEEWLEGECK----------GKVGIFPKVFVEDCA  p67phox #2
VVARAEYDFAA-------VSEEEISFRAGDMLNLALKE-----QQPKVRGWLLASLD---------GQTTGLIPANYVKILG  Pex13p
```

Figure 2.

a. Structure of Hck SH3 domain. Reprinted from (Horita et al. 1998), with permission from Elsevier.

b. Alignment of some SH3 sequences discussed in text. On top the secondary structure elements and the consensus SH3 domain sequence (adopted from Larson and Davidson, 2000) are indicated. Hydrophobic core positions are shaded in grey and the conserved residues involved in peptide binding are boxed. Dashes indicate gaps to maximise the alignment. If multiple SH3 domains are present in a protein, the number after the SH3 domain indicates the SH3 domain in question (for instance, Crk #1 means the first SH3 domain of Crk).

The N- and C-termini are close to each other. This is a characteristic feature of protein domains in general. In evolution the close proximity of the termini has ensured that small domains have fairly easily become parts of various proteins without disrupting the general structure of the protein in question (Pawson et al. 2002).

Analysis of 266 nonredundant SH3 domains showed that the variability seen in the length of different SH3 domains arises mainly from insertions and deletions in the n-Src and distal loop regions whereas less variation was observed in the RT-loop region (Larson and Davidson 2000). Thus, the loop structures are the most variable regions on the otherwise conserved SH3 domain. The most conserved residues in SH3 domains are the ones that are involved in ligand binding (see later in Section 5.4.5). In addition to the sequence conservation seen in the residues involved in ligand binding, the so-called hydrophobic core, comprising residues that are virtually never accessible to solvent, are also well conserved (see Figure 2b.). For example, the alanine residue flanking the RT-loop is conserved and was suggested to play also a role in maintaining the RT-loop in an exactly right conformation for the ligand binding (Larson and Davidson 2000).

5.4.4. Typical SH3 ligands

Deletion mapping of one of the clones (3BP-1) selected by the Abelson tyrosine kinase (Abl) SH3 domain suggested a proline rich-sequence as a putative SH3 binding site (Cicchetti et al. 1992). Subsequent alanine scanning mutagenesis of the peptides identified the core PxxP motif and revealed the specific residues for the binding (Ren et al. 1993). Proline rich regions in proteins were somewhat underappreciated at that time and in fact, only by discovering the missing link of the activation of Ras signalling pathway, namely the Grb2 (Growth factor receptor bound 2) adaptor protein able to bind to the proline rich region of Sos (Son of sevenless) via an SH3 domain leading to the membrane recruitment of the complex and to subsequent activation of Ras by Sos (McCormick 1993), the significance of proline-rich regions was widely accepted to be important in mediating cellular communication.

After discovering the first evidence of SH3 binding to proline rich sequences (Cicchetti et al. 1992; Ren et al. 1993) numerous studies using yeast two-hybrid screens, phage display and combinatorial chemistry have been undertaken to exploit the SH3 ligand recognition (Cheadle et al. 1994; Ren et al. 1994; Rickles et al. 1994; Sparks et al. 1994; Yu et al. 1994). These studies revealed that SH3 domains recognise linear sequences consisting of two xP dipeptides where the x was usually a hydrophobic residue (marked by Φ from now on), connected via a scaffolding residue to form the core <u>ΦP</u>x<u>ΦP</u> motif of the SH3 ligands. Structural studies of SH3/ligand complexes (Feng et al. 1994; Lim et al. 1994; Musacchio et al. 1994; Yu et al. 1994) showed that a typical SH3 ligand adopts an extended PPII helical conformation with exactly three residues per one turn (Figure 3.). This structure places the two prolines in the <u>ΦP</u>x<u>ΦP</u> core adjacent to each other on one side of the helix and the hydrophobic residues (Φ) preceding the prolines on the other side (Figure 3.). Thus the dipeptides ΦP form the face of the PPII helix that makes contact with the SH3 domain. The scaffolding residue, on the other hand, points away from the SH3 domain (Figure 3.). This residue is often a proline as well and it has been suggested to give more stability to the PPII helical conformation (Yu et al. 1994).

In these initial studies it was also shown (Feng et al. 1994; Lim et al. 1994) that SH3 domains can recognise residues outside the ΦPxΦP core to position the ligands in two possible orientations. This is due to twofold rotational pseudosymmetry of the PPII

helix meaning that side chains and backbone carbonyl oxygens are displayed with similar spacing in either of the two N- to C-terminal orientations. The appreciated consensus sequences for SH3 ligands were named as Class I (plus orientation) R/KxΦPxΦP and Class II (minus orientation) ΦPxΦPxR/K. The positively charged arginine (or lysine) binds the same RT-loop residues of the SH3 domains in both binding orientations thus determining the orientation of the ligand (see Figure 3. and discussion below).

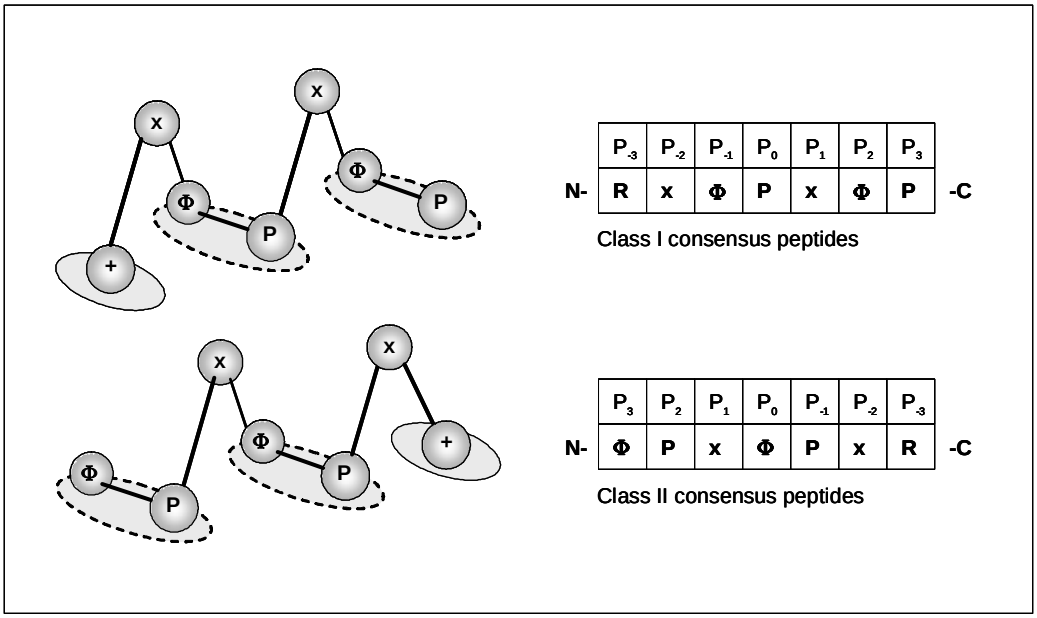


Figure 3.
Binding of Class I and Class II ligands to SH3 domains. Ligand residues are numbered according to Lim et al. (1994). The ΦP dipeptides contact the two binding pockets (circled with dashed line) and the positively charged residue contacts the specificity pocket (circled with solid line) on the SH3 surface. Depending on the binding orientation the PxxP defining prolines occupy the positions P_0 and P_3 (Class I) or P_2 and P_{-1} (Class II). The positively charged residue inherently occupies the position P_{-3}.

5.4.5. SH3 ligand binding site

The SH3 surface for ligand binding is rather shallow and is defined by side chains of well conserved residues (the boxed residues in Figure 2b; Musacchio et al. 1992; Yu et al. 1992; Musacchio et al. 1994; Yu et al. 1994). Among the most conserved ones are the first tryptophan of the generally conserved aromatic WW dipeptide in the beginning of the β_C, and a proline in the end of β_D just before the 3_{10} helix, as well as asparagine and tyrosine residues in the 3_{10} helix. The two tyrosine residues in the beginning of the RT-loop (often seen to generate an ALYDY-sequence motif) also show strong conservation among different SH3 domains. These residues form a relatively hydrophobic surface where the side chains are closely packed against each other. This hydrophobic patch is bordered by RT- and n-Src loops, which contain a set of conserved acidic residues that also contribute to the specificity of the ligand binding (see Figure 4.). The ligand binding site is formed by these conserved residues (Yu et al. 1994). In determining the orientation of the Class I or Class II ligand binding an acidic residue in the RT-loop (for example, Asp 99 in chicken c-Src; Weng et al. 1995) has been shown to form a salt bridge with the ligand orientation defining arginine (or lysine) residue.

Some SH3 domains like the one in Src kinase and PI3K are capable of binding to peptides that are either in plus or in minus orientation (Feng et al. 1994; Viguera et al. 1994; Yu et al. 1994). On the other hand, some, like Abl SH3 domain may have preference for only one orientation (Musacchio et al. 1994; Pisabarro et al. 1998) although the biological function of this is not clear. Of note, in the inactive conformation of Src the SH3 complexes with the linker region between the SH2 and kinase domains. This linker adopts a left-handed polyproline type II helical conformation only having a single proline residue in the "PxxP" defining region (Xu et al. 1997). One reason for a differential ability or recognition for both orientations may be the capacity to accommodate different binding partners when assembled to multiprotein complexes. The sub-optimal binding of the Src SH3 in the kinase regulation, on the other hand, may be involved in very dynamic nature of these interactions (Xu et al. 1997).

It was recently shown that the highly conserved tryptophan residue (the first W of the conserved WW-dipeptide, see Figure 2b.) on the SH3 binding surface is able to direct whether it can bind to Class I, Class II or ligands from both classes by moving the side-chain indol ring towards or away from the conserved proline residue in the ligand. In addition, upon ligand binding also RT- and n-Src loops move to aid the association (Fernandez-Ballester et al. 2004).

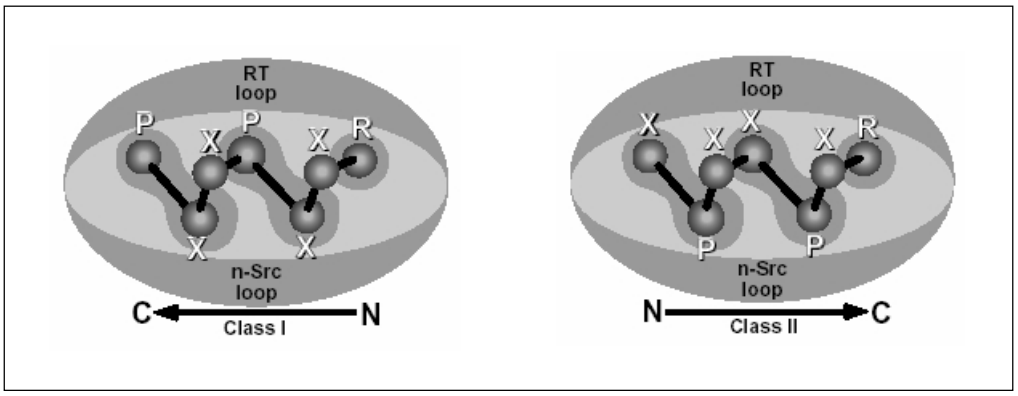


Figure 4.
Schematic view of Class I (left) and Class II (right) peptide ligands binding to an SH3 domain. Reproduced from Mayer (2001) with permission from The Company of Biologists Ltd.

5.4.6. Atypical SH3 ligands and ligand recognition

Recently, sequences lacking a typical PxxP motif have been reported to bind to a variety of SH3 domains. These involve cases where the peptide does not adopt the classical PxxP consensus when occupying the SH3 binding surface. In other types of interactions the ligand contains a defined motif but it does not accommodate the canonical PxxP binding surface on the SH3 domain. The complex binding modes that involve, for example, SH3 dimerisation to take place represent still another strategy for SH3/ligand recognition. It seems that in many cases the atypical ligand recognition provides a novel regulatory function for an SH3 domain.

5.4.6.1. Atypical PxxP-like ligands

Many recent reports have demonstrated that SH3 domains have evolved specificity to peptides that adopt a PxxP-like binding motif. For example, mapping of the adaptor protein amphiphysin SH3 domain binding site in dynamin GTPase revealed a specific sequence PSRPNR as a motif for interaction. Selection of a combinatorial peptide library with amphiphysin SH3 domain confirmed the novel PxxP-like consensus motif (Grabs et al. 1997). A similar consensus sequence was also present in synaptojanin that was later also found as an amphiphysin SH3 binding ligand by a phage display approach (Cestra et al. 1999). The novel consensus motif ϕXRPXR where ϕ represents a hydrophobic amino acid, deviates from the conventional Class II consensus (ΦPxΦPxR/K) having an arginine residue preceding the second proline residue of the core. In the classical consensus a hydrophobic residue normally occupies this site.

As a monomer Eps8 (epidermal growth factor receptor pathway substrate 8) was shown to bind the consensus PxxDY from random peptide libraries and subsequently shown to recognise a similar motif in the natural ligands such as Abi-1 and RN-Tre (Kishan et al. 1997; Mongiovi et al. 1999).

The adaptor protein CIN85/Ruk/SETA SH3 domains selected a specific PX(P/A)XXR sequence from the random peptide library (Kurakin et al. 2003). The consensus sequence could be mapped to a known binding partner, like Cbl-b protein as well as predicting novel CIN85 interaction partners by the database searches. Interestingly, the SH3 domains of α and βPIX proteins bind to PAK (Manser et al. 1998) in a sequence that was shown to be highly similar to the novel CIN85 consensus motif suggesting similar strategies in ligand recognition.

The adaptor protein SKAP55 (Src kinase-associated protein of 55 kDa) is able to bind FYB (Fyn T binding protein) by its novel proline independent motif RKxxYxxY (Kang et al. 2000). Fyn and Lck SH3 domains (Class I binders) were also able to bind this motif suggesting that the recognition of the RKxxYxxY motif could be extended to SH3 domains that are able to bind to Class I consensus sequences thus the motif being a remnant of the Class I consensus sequence (R/KxΦPxΦP). This motif represents an alternative strategy of connecting the SKAP55 to cellular signalling proteins that would otherwise mediate the conventional Class I proline-dependent interactions (Kang et al. 2000).

An atypical SH3-binding consensus sequence PxxxRxxKP was identified originally in a T-cell receptor associated adaptor protein SLP-76 as a high-affinity binding motif for the C-terminal SH3 domain of adaptor protein Grb2 (Lewitzky et al. 2001 and GADS (Grb2-related adaptor downstream of Shc, also known as Mona, Grf40, GRID, GrpL, Grap-2, GrbX, GrbLG; Berry et al. 2002). The similar region of homology was also found from Gab1 (Grb2-associated binding protein 1), Gab2, Gab3, BLNK as well as from the signal transduction protein AMSH (associated molecule with the SH3 domain of STAM) and deubiquitinating enzyme UBPY. UBPY was already earlier reported to associate with Hrs binding protein (Hbp/STAM2) via a similar consensus motif (Kato et al. 2000). In addition, this motif seems to be evolutionary conserved since two proteins involved in the development of R7 photoreceptor cell, Drk and Dos represent the Drosophila homologues of this interaction. Drk (homologue of Grb2 and GADS) is capable of binding two sequence motifs in Dos (Daughter of sevenless) that highly

resemble the SLP-76 binding site of Grb2 and GADS (Feller et al. 2002). Structural studies of the C-terminal SH3 of GADS/PxxxRxxKP (Harkiolaki et al. 2003; Liu et al. 2003) and the STAM2 SH3/PxxxRxxKP (Kaneko et al. 2003) complexes revealed specific binding surfaces on these SH3 domains resulting in atypical ligand selection. In a recent study the C-terminal SH3 domain of GADS was shown to bind to HPK1 (hematopoietic progenitor kinase 1) with a mixed consensus sequence (PxxPxRxxK; Lewitzky et al. 2004).

5.4.6.2. Atypical motif-containing ligands docking to atypical SH3 binding surface

An example of an atypical SH3 interaction that involves an unconventional binding surface on the SH3 domain comes from the yeast *S. cerevisiae* peroxisome associated protein Pex13p that is able to interact with peroxisomal proteins Pex14p and Pex5p via its SH3 domain. The canonical Class II site mediated the interaction with the Pex14p whereas the interaction with the Pex5p was not PxxP-dependent (Bottger et al. 2000). Instead, a short 25 amino acids long α-helical structure in N-terminal part of the Pex5p was found both necessary and sufficient for this interaction to take place (Barnett et al. 2000). Structural studies revealed the canonical Pex14p PxxP ligand and the non-conventional Pex5p α-helical structure to be distinct and in different locations, the Pex5p binding site being on the opposite side of the SH3 domain compared to the conventional PxxP binding site (Pires et al. 2003).

5.4.6.3. Other atypical complex SH3 interactions

The atypical complex SH3 recognition can be achieved by various mechanisms that may involve the canonical PxxP contacting residues on the SH3 surface, but differ from the conventional ligand recognition. For example, in the complex of p53-binding protein 2 (53BP2) with p53 (Gorina and Pavletich 1996) the 53BP2 uses one of its four ankyrin repeats together with the SH3 domain to occupy two non-contiguous binding sites on the p53. The SH3 domain surface that typically occupies the PxxP peptide, docks the L3 loop structure of the p53 in a non-PxxP manner.

Eps8 SH3 domain was shown to exist as a monomer and a three dimensional (3D) domain-swapped dimer in equilibrium (Kishan et al. 1997; Mongiovi et al. 1999; Kishan et al. 2001). The dimerisation was mediated by a beta strand exchange between two Eps8 SH3 domains where the n-Src loops were functioning as hinges to connect the monomers. Similarly, MAGUK (membrane associated guanylate kinase) proteins show intra- and intermolecular interactions between SH3 and GK (guanylate kinase like) domains (Tavares et al. 2001) and might assemble via 3D-domain swapping where the core SH3 domain from one protein would interact with the SH3-GK intervening subdomain from another protein (McGee et al. 2001). The multiprotein assembly of the Grb2 and Vav is mediated by the SH3 domain dimerisation of these proteins (Nishida et al. 2001; Ogura et al. 2002).

5.5. Specificity of the SH3 ligand binding

As discussed (see Sections 5.4.4.-5.4.5.), despite having distinct specificity pockets, many SH3 seem to have overlapping recognition profiles. On the other hand, the short ligand PxxP motif restricts the variability that can be used for generating binding specificity. These facts raise an important question of achieving the specificity within SH3-mediated protein-protein interactions. On the SH3 side, the third binding pocket of

the SH3 surface formed by the RT- and n-Src loops is a key determinant of the specificity of ligand binding. Most of the sequence variation is concentrated on the loop regions in otherwise very similar SH3 domains. Large deviations were seen in the RT- and n-Src loop structures by examining the structural data (Feng et al. 1995; Arold et al. 1998). The length of the loops can also vary considerably human SH3 n-Src loops having in average 6 residues and RT-loops having in average 18 residues. In the case of the RT-loop the conserved acidic residues are located near the neighbouring β sheets leaving the central part of the loop more variable. These conserved residues provide the critical structural elements for the third binding pocket on the SH3 surface (see Section 5.4.5) whereas the more variable, non-conserved residues may contribute additional specificity to binding via more unique interactions with the ligand protein.

Considering that most of the binding energy to the SH3/minimal ligand interaction is provided by the two ΦP dipeptides and the positively charged residue, any short PxxP peptide would be expected to be relatively promiscuous in binding to any SH3 domain. This could be seen, for instance, by an early study with various natural peptide ligands interacting with several SH3 domains (Viguera et al. 1994). However, sequence variation of either the consensus or non-consensus residues in the PPII helix region has been shown to affect the specificity of ligand binding. These studies have used peptide ligands that contain the Class I (R/KxΦPxΦP) or Class II (ΦPxΦPxR/K) consensus sequences embedded within less than ten flanking amino acid residues. For example, specific targets that contained atypical PxxP consensus motifs have been revealed from studies of Abl SH3 ligand selection (Feng et al. 1994; Yu et al. 1994; Weng et al. 1995) and from the analysis of CrkN-SH3/C3G peptide complex (Knudsen et al. 1995; Wu et al. 1995). The influence of the sequence variation of the non-consensus residues has been studied by selection of phage display or chemically synthesised peptide libraries (Sparks et al. 1994; Yu et al. 1994). Despite the evidence of some specificity, the SH3 binding surface for a peptide is relatively small and with limited hydrogen-bonding interactions (Lee et al. 1996) and consequently the binding affinities are weak (K_d values typically in micromolar range). The recognition of the ΦP dipeptides by the SH3 hydrophobic surface is based on the notion that the dipeptide has a unique backbone substitution pattern where a C-substituted residue is followed by the proline, the only naturally available N-substituted residue (Nguyen et al. 1998). In the SH3 binding the proline is not recognised by the "lock and key" model where the majority of the proline ring would be in contact with the SH3 surface. Instead, in the proline residue the primary recognition element is the δ carbon (the carbon atom that is cyclised to main-chain nitrogen atom; see Figure 5.).

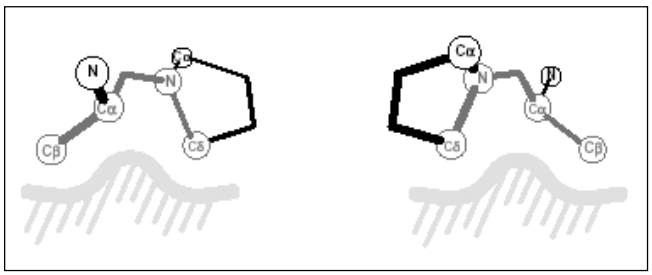


Figure 5.
Schematic view of proline recognition in Class I (left) and Class II (right) peptide ligands binding to an SH3 domain. The primary recognition element is the δ carbon cyclised to main chain nitrogen atom. Reproduced from Mayer (2001) with permission from The Company of Biologists Ltd.

In other words, the N-substitution is the main feature recognised by the SH3 domains. The N-substitution takes advantage of relatively low binding affinities but is still very discriminatory when compared to all other natural occurring amino acids (Nguyen et al. 1998; Zarrinpar and Lim 2000).

Several studies have shown that additional molecular contacts outside of (and in some cases rather distant from) the PxxP motif can improve the specificity and affinity of the SH3 binding. Selecting biased phage display libraries where the Class I or Class II consensus sequences were embedded within a short random sequence resulted in target peptides that showed approximately 20-fold differences in their affinities compared to the ligand core (Rickles et al. 1994; Rickles et al. 1995). Subsequent structural studies revealed the binding of the two dodecapeptides to Src SH3 domain displaying enhanced specificity and affinity (K_d values of 0.45 μM and 1.2 μM) compared to the core attached to a single arginine residue (K_d value of 22 μM) and that the flanking residues contacting RT- and n-Src loops were responsible of the improved affinity binding (Feng et al. 1995). In another study (Pisabarro and Serrano 1996) a rational design of peptide ligands was utilised to obtain high affinity and specificity target for Abl SH3 domain. Affinity of the peptide to its cognate SH3 domain was increased (~0.4 μM compared to the wild-type 3BP1 peptide 34 μM; Viguera et al. 1994) and more importantly, the affinity for the Fyn SH3 was decreased approximately ten times compared to the wild-type peptide, increasing the specificity by about 1000 times (Pisabarro and Serrano 1996). Yet another study showed the Abl SH3 domain complexed to a peptide with a high affinity compared to the FynSH3 domain binding to the same peptide (K_d values of 1.5 μM compared to the 273 μM, respectively; Pisabarro et al. 1998). In all cases the changes in the specificity could be explained by molecular contacts in Abl SH3 RT- and n-Src loops.

Another example of the ligand residues outside the PxxP core sequence comes from the mapping of the peptide binding sites of the second SH3 domain of the adaptor protein Nck (Nck SH3-2). In this case a strong requirement for a serine residue located C-terminally to the Class II consensus sequence in the 18-mer PAK1 peptide was shown (Zhao et al. 2000). Additionally, many other proteins were reported to bind the Nck SH3-2 were shown to contain this motif.

Secondary structures next to the PPII helical core may contribute the ligand binding specificity, as well. A recent report showed a novel interaction where the C-terminal Src-kinase (Csk) SH3 domain binds with a high affinity (K_d 0.8 μM) and specificity to the proline-enriched phosphatase (PEP) (Gregorieff et al. 1998). The solution structure of the complex revealed that in addition to a characteristic Class II PPII helical structure, the peptide also adopted a 3_{10} helix after the PxxP core. The PxxP motif of the PEP peptide, although needed, was not sufficient for the peptide to bind Csk. Instead, the very 3_{10} helix structure was providing the complex with the high affinity and specificity of binding by positioning the hydrophobic positions of the peptide onto a hydrophobic patch on the SH3 binding surface. The critical isoleucine and valine residues C-terminal to the PxxP core occupied the three residues in the end of the β_B and in the beginning of the n-Src loop of the Csk SH3. Thus, this case showed that high specificity of an SH3 domain towards its natural ligand could rise from the tertiary interactions outside the PxxP motif (Ghose et al. 2001).

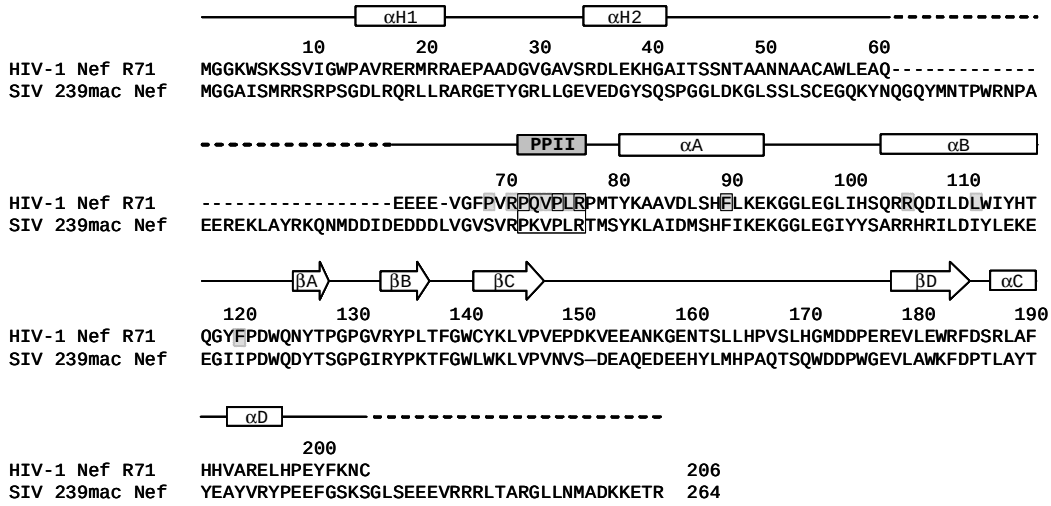
The C-terminal SH3 domain of the phagocyte NADPH oxidase cytosolic activator $p67^{phox}$ has been shown to interact with the C-terminal PxxP sequence of another cytosolic component $p47^{phox}$ (Finan et al. 1994). Recently tryptophan fluorescence binding studies showed that the PxxP motif of the $p47^{phox}$ containing flanking sequences bound to $p67^{phox}$ C-terminal SH3 with very high affinity (K_d 0.024 μM where as the PxxP peptide only had very modest binding affinity (K_d 20 μM; Kami et al. 2002). Subsequent structural studies showed that the high affinity binding required 22 additional amino acids C-terminal to the $p47^{phox}$ PxxP motif. These remaining residues adopted a helix-turn-helix (HTH) secondary structure consisting of two closely packed antiparallel α-helices that bound the specificity pocket between the RT- and n-Src loops of the of the $p67^{phox}$ C-terminal SH3 domain. It is of note that this non-PxxP HTH secondary structure was also capable of binding to the $p67^{phox}$ C-terminal SH3 domain by itself with a binding affinity comparable to the PxxP peptide only (10 μM). However, this "independent" interaction was shown to be critically dependent of the conserved arginine residue of the Class II consensus motif of the $p47^{phox}$ (PAVPPR) contacting the negatively charged residues on the $p67^{phox}$ C-terminal SH3 specificity pocket coordinating the binding of the non-PxxP peptide (Kami et al. 2002), and therefore could perhaps be classified as another example of increasing the SH3 specificity and affinity by a secondary structure outside the PxxP motif.

5.6. Nef as an SH3-binding protein

Nef is a 27-34 kDa myristoylated protein of primate lentiviruses (HIV-1, -2, and SIVs), and important for development of high viremia and immunodeficiency in the infected host, in humans as well as in SIV infected macaques (reviewed in Cullen 1999; Renkema and Saksela 2000; Fackler and Baur 2002; Greenway et al. 2003). Despite the overall sequence variation among different Nef alleles, amino acids in the proline rich region that adopt the minimal consensus binding site for SH3 domains are very conserved between the various lentiviral Nef proteins (see Figure 6.a., Lee et al. 1996, Los Alamos Database; http://hiv-web.lanl.gov/content/index).

Several cellular functions of Nef, such as enhancement of HIV-1 replication, downregulation of cell-surface expression of human leukocyte antigen class I (HLA-I) and alteration of signal transduction pathways have been shown to require an intact PxxP motif (reviewed in Renkema and Saksela 2000; Geyer et al. 2001). In the mouse models of HIV-1 and SIV, both the HIV-1 and SIV Nef transgenic mice developed an AIDS-like disease (Hanna et al. 1998; Simard et al. 2002). Lack of the functional PxxP motif in the HIV-1 Nef transgene completely abolished the pathogenic effect of Nef suggesting that interaction with a cellular protein via the PxxP motif is required for the induction of the disease in these mice (Hanna et al. 2001). HIV-1 Nef has been reported to interact with several SH3 domain containing cellular proteins, including the Src family kinases Fyn, Hck, Lck, Lyn and Src and the guanine nucleotide exchange factors Vav1 and Vav2 (Arold et al. 1998; Renkema and Saksela 2000). Hanna et al. (2001) studied the effect of Hck by breeding the transgenic Nef mice with Hck-/- mice. It did not abolish the pathological phenotype but clearly delayed the disease development indicating that Hck is important but not essential for the Nef induced pathogenesis in this model.

a.



b.

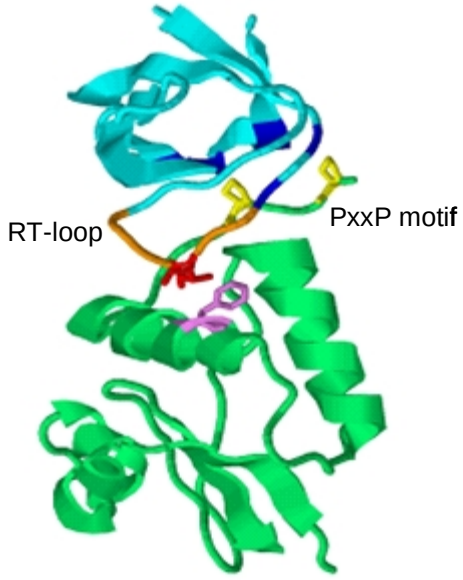


Figure 6.

a. Alignment of HIV-1 NL4-3 Nef R71 (Saksela et al. 1995) and SIV 239 mac Nef protein sequences showing the secondary structures (adapted from Arold and Baur, 2001). The numbering is according to Nef R71. The PPII helical region is indicated in a grey box. The residues involved in SH3 binding (PPII helical region together with the distant F90 residue) have been boxed with a solid black line. The mutated residues used in this study (**I**, **II**) have been boxed in grey.

b. HIV-1 Nef R71 (coloured in green) in complex with an Fyn I SH3 domain (coloured in cyan; Lee et al. 1996). The P72 and P75 residues of Nef contacting the SH3 have been coloured in yellow and the "tip" of the SH3 RT-loop in orange. The critical I96 residue of the Fyn I SH3 contacting the F90 residue in Nef (magenta) is indicated in red. Hydrophobic residues of the SH3 domain involved in ligand binding are shown in dark blue. The structure (1EFN) was obtained from Protein Data Bank (http://beta.rcsb.org/pdb/) and modified by RasMol.

Nef provides an example of the increment of SH3 binding specificity and affinity by regions that are distally located from the canonical PxxP region. Nef has been shown to have selective SH3-binding characteristics. The Src family kinases Hck and Lyn SH3 domains are recognised efficiently whereas the related Fyn and Lck SH3 domains poorly (Lee et al. 1995; Saksela et al. 1995). Affinity of the Hck SH3/Nef complex is unusually high (K_d value of 0.25 μM; Lee et al. 1995). A peptide spanning across the Nef PxxP region showed only a modest affinity for Hck SH3 domain (K_d value of 91 μM) suggesting additional contacts to be involved. In striking contrast to the Hck SH3/Nef binding, the related Fyn SH3 domain had almost a 100-fold lower affinity towards the full-length Nef (K_d value of >20 μM; Lee et al. 1995). Mutagenesis studies have shown that the key for the high affinity interaction is the Hck SH3 RT-loop isoleucine residue that corresponds to an arginine in Fyn SH3. By changing this arginine to isoleucine in Fyn SH3 RT-loop (Fyn SH3 R96I mutant) the high binding affinity towards Nef was achieved (K_d value of 0.38 μM; Lee et al. 1995). Structural studies of the Nef/SH3 complex have implied that the canonical Class II PxxP motif of HIV-1 Nef accommodates the conventional two binding pockets on the SH3 surface (Lee et al. 1996). The selectivity of binding is provided by the RT-loop that, as a flexible and highly variable structure (Lee et al. 1995; Lee et al. 1996; Arold et al. 1998), is capable of extending over the surface of Nef. The region in Nef that is in contact with the RT-loop is composed of multiple non-contiguous parts of the Nef protein located distally from the PPII helical region of Nef (see Figure 6.; Lee et al. 1996). The side chain of the critical isoleucine in the RT-loop is placed into a hydrophobic pocket between two α-helices in the Nef structure (Lee et al. 1996), shown in Figure 6 b. and schematically in Figure 7.

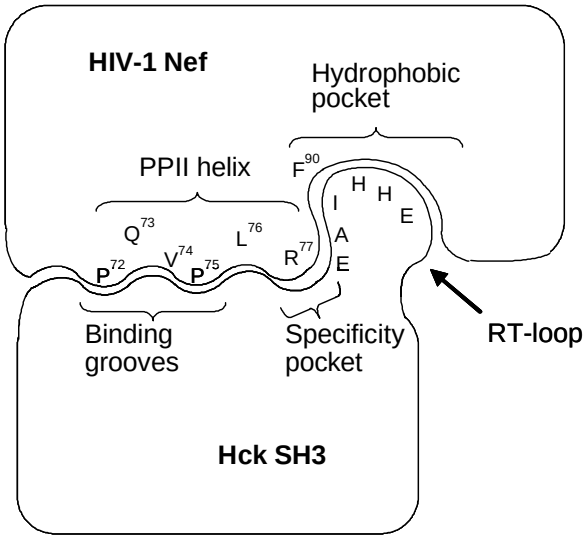


Figure 7.
Schematic drawing of the HIV-1 Nef in complex with Hck SH3 domain. Shown is the Class II consensus sequence ($P^{72}Q^{73}V^{74}P^{75}L^{76}R^{77}$) as well as the distant F90 residue contacting the Hck SH3 RT-loop isoleucine residue.

In HIV-2 and SIV Nef proteins this hydrophobic pocket is substituted with residues of different character than those of HIV-1 Nef resulting in differential binding to Src family kinase SH3 domains (Collette et al. 2000). On the other hand, a crystal structure of HIV-1 Nef in complex with the wild-type Fyn SH3 (Arold et al. 1997) showed that

the aliphatic moiety of the arginine side chain also docks the same hydrophobic pocket on the Nef surface as the isoleucine of the "Hck-like" Fyn SH3 R96I showed by Lee et al. (1996). It was shown that upon Nef binding the R77 residue in Nef makes a salt bridge to D100 residue in the Fyn SH3 RT-loop and simultaneously displaces the stabilising hydrogen bonds between the RT-loop and the main chain nitrogen atoms. This destabilisation of the RT-loop results in its accommodation on the Nef surface.

In structural studies the SH3 domains are usually complexed with peptide ligands whereas in the case of Nef almost a full-length protein is in complex with SH3 domains (Lee et al. 1996; Arold et al. 1997; Arold et al. 1998). It is therefore unclear whether the stabilising interactions between the loop structures of SH3 domains and the non-PxxP regions far outside the PxxP core in their binding partners are widespread or rather is the Nef/SH3 binding unusual in nature.

As discussed, HIV-1 Nef shares common features, like the conserved proline motif, with the simian immunodeficiency virus (SIV) and HIV-2 Nef proteins. However, the possible role of the SH3-mediated functions of SIV/HIV-2 Nef remains obscure. Downregulation of the major histocombatibility complex class I (MHC-I) cell surface expression by HIV-1 Nef is SH3-dependent but for the same function SIV Nef needs the unique C-terminal region that is not found in HIV-1 Nef (Swigut et al. 2000). Studies with macaques that were infected with SI viruses (SIV239mac strain) that contained either intact or PxxP-mutated Nef proteins have resulted in controversial data. The PxxP-mutant viruses were shown to relatively rapidly to revert back to wild-type viruses increasing progressively and development of simian AIDS (Khan et al. 1998). On the other hand, Kirchhoff and colleagues showed that the PxxP motif in SIV Nef is largely dispensable for replication and pathogenicity of the SIV239mac and that the reversions of the mutant PxxP motif occur often only at relatively late stages of the viral infection (Lang et al. 1997; Carl et al. 2000). SIV Nef has been shown to interact with the Src family kinases Lck, Hck (Greenway et al. 1999) and Src (Du et al. 1995; Lang et al. 1997) but apparently independently of the proline motif and SH3-binding. SIV Nef recognised the Hck SH3 domain poorly but instead was able to bind Fyn and Src SH3 domains to some extent (Greenway et al. 1999; Collette et al. 2000). Thus, so far no high-affinity ligands have been identified for SIV Nef. It could be argued that the lack of high-affinity binding partners is because the analyses have not been undertaken with the natural SIV host-derived cellular target proteins. However, it was shown recently that the simian Hck is 100% identical to the SH3 domain of the human Hck but was not any better in binding to SIV Nef compared to the human counterpart. Similar results were also obtained by using simian Lck and Fyn (Picard et al. 2002). It was shown earlier (Collette et al. 2000) by homology modelling and biochemical studies that the binding of an SH3 domain to SIV Nef is affected by the amino acid changes in the SIV Nef region that corresponds to the hydrophobic binding pocket in HIV-1 Nef accommodating the SH3 RT-loop. Site-directed mutagenesis of these residues in the SIV Nef was able to restore the Hck SH3-binding ability. This reflects the differential molecular binding surface in SIV Nef compared to the HIV-1 Nef and a possibility that the RT-loop region of the SH3 domains does not occupy the binding surface efficiently enough to achieve high-affinity binding.

5.7. Specificity versus promiscuity

The whole process of signal transduction leading to a cellular response is dependent on protein-protein interactions. A key question is, whether these SH3/ligand interaction pathways should be considered as pairwise, specific interactions or does a dynamic equilibrium between low-specifity and low-affinity interactions generally explain the relevant biological outputs (Ladbury and Arold 2000; Mayer 2001). As discussed, both the SH3 domains and proline rich ligands are abundant in cells resulting in hundreds of possible interaction pairs. Moreover, the affinities of most of the SH3-mediated interactions are low (K_d values ranging from high nanomolar to low micromolar). However, as discussed, such measurements have typically involved only short PxxP peptides instead of the corresponding native proteins. It was suggested (Ladbury and Arold 2000) that most of the SH3 domains do not have enough of intrinsic specificity to account for specific biological output but have differential mechanisms to circumvent the problem. Strategies for enhancing the specificity of binding involve the cellular co-localisation of proteins, for example on cellular membranes to yield high local protein concentrations. Another solution is the assembly of multiprotein complexes that would only completely assemble in a presence of a specific protein. For example, the intramolecular interaction of the Src SH3 domain with the kinase-linker region is clearly enhanced by the other domains of the Src kinase (Xu et al. 1997). A combined effect of more than one domain may in some cases add to the specificity. The Nck adaptor protein containing multiple SH3 domains may bind multiple sites in the ligand protein simultaneously to enhance binding specificity (Wunderlich et al. 1999). A recent study reported that PLCγ did not show detectable binding to SLP-76 alone but required the SPL-76 to be complexed with GADS for the interaction to occur. It was shown previously that SLP-75 undergoes changes in the secondary structure upon ligand binding (Liu et al. 2003) suggesting that cooperative interactions enhance the specificity of the PLCγ binding to the SLP-76/GADS complex (Houtman et al. 2004).

Absolute selectivity may not be necessary in most physiological situations. As a result, in this case a parallel processing of signals would occur where the protein with an SH3 domain would interact with multiple downstream effectors resulting in triggering more than one signalling pathway ultimately leading to a junction protein. The junction protein would be activated, for example, only if two incoming signals arrive simultaneously. This would imply a constantly changing network of interactions where the relevant output is achieved by changes from the dynamic equilibrium of all the interacting partners. Thus the specificity *in vivo* is not carried in an individual SH3 domain but is rather gained through the context where the binding partners are presented (Ladbury and Arold 2000; Mayer 2001). If this is generally true, many SH3 domains should be redundant and be able to functionally replace the native SH3 domain. A recent study showed that the yeast *S. cerevisiae* SH3 domain from the ShoI protein was functionally replaced with a set of non-yeast SH3 domains and that these other non-yeast SH3 domains were able to bind to the biologically relevant binding partner of ShoI SH3, the protein Pbs2 *in vitro* and *in vivo* supporting the idea of the SH3 domains having only a limited information content about the specificity (Zarrinpar et al. 2003b). On the other hand, in the same study a striking difference was observed when the yeast SH3 domains were tested for rescuing the ShoI function: there were no other cross-reacting SH3 domains found from the yeast, neither did the other yeast SH3 domains recognise the ShoI SH3 binding partner Pbs2 suggesting that this SH3/ligand pair contains enough information to encode interaction specificity among the yeast SH3

domains. Accordingly, it was proposed that the specificity in the yeast SH3 domains has evolved both through the positive selection of the Pbs2 peptide/ShoI SH3 ligand pair and also through negative selection against binding to competing with other SH3 domains in yeast. The generality of the negative selection to other yeast SH3 domains as well as the other proteomes is to be established.

Structures of two SH3 ligand complexes, Csk SH3 with PEP (Ghose et al. 2001) and $p67^{phox}$ C-terminal SH3 with the $p47^{phox}$ C-terminus (Kami et al. 2002) have revealed regions outside the PxxP motif to contribute to the specificity of the ligand binding (discussed in Section 5.5). NMR studies showed that in the Csk SH3/PEP interaction several residues in the PPII helix binding surface of the SH3 exhibited motion in slow millisecond timescale indicating of possible large changes in local structure. On the other hand, none of the residues comprising to the specificity surface exhibited those slow motions. It was suggested that regions exhibiting this kind of backbone dynamics would be more tolerant to the nature of the associating ligand as long as the overall features required for the interaction are present. This interpretation would fit to the promiscuous nature of SH3 mediated interactions (Ghose et al. 2001). In a more recent study the backbone dynamics concerning the specificity of binding was studied in more detail also by NMR (Dutta et al. 2004). $p67^{phox}$ C-terminal SH3 free state and in complex with the $p47^{phox}$ tail peptide revealed fast motions and large changes in backbone dynamics of residues in the specificity site. Again, the slow motions were seen only in the PxxP binding surfaces (like in the case of the Csk SH3/PEP). The authors proposed that the slow motions arise from conformational reorganisation of the SH3 and the PxxP defining residues. They assumed that several conformers having a slightly different structure at any time point. These variable structures would be also able to bind ligands of slightly different structures resulting in non-specific binding. However, these two studies concentrated on the backbone dynamics of these interactions. Since most of the SH3/ligand interactions occur via side chains, analysis of the side chain dynamics would be more informative in this kind of a specificity question.

5.8. Predicting binding partners

In many cases the cellular signalling complexes rely on protein-interaction domains that are able to recognise short target sequences. In a simplistic view, the inherent specificity of an SH3 domain determines the preference for target peptides that obey some common structural theme. Therefore attempts to develop computational tools to predict SH3-ligand pairs from sequence databases have been made. Two computational tools were recently published for prediction of the proline rich ligands for SH3 domains. Yaffe and co-workers (Yaffe et al. 2001; Obenauer et al. 2003) reported the search program Scansite (http://scansite.mit.edu/) and (Brannetti et al. 2000) introduced the iSPOT (http://cbm.bio.uniroma2.it/ispot/). Data from the SH3 sequence alignments and from SH3/ligand structures was combined with the experimental data from peptide panning experiments to develop an algorithm (SH3-SPOT) that was used to predict a preferred decapeptide sequence for any SH3 domain (Brannetti et al. 2000). Src SH3 domain as a test case resulted in correct prediction of ligand peptides even when the experimental data concerning the Src SH3 domain was omitted. On the other hand, when no binding data of the entire Src family SH3 ligands was added, correlation between the experimental and predicted data was very low. When the SH3-SPOT was used to search for binding partners of the Abl, Hck and Grb2 SH3 domains, known

interactors were ranked very high in the score, like the 3BP2 for Abl SH3, HIV-1 Nef and Cbl for Hck SH3 and Sos for Grb2 SH3, showing the validity of this approach.

Recently, a combined computational and experimental approach for predicting SH3 interaction networks in yeast was reported (Tong et al. 2002). Twenty different SH3 domains form *S cerevisiae* were used to screen random peptide libraries to define consensus sequences for each ligand. Based on these results, the yeast proteome was searched for potential natural SH3 ligands to derive a protein-protein interaction network. Each interaction inferred from the proteome search for a specific peptide ligand was then studied by yeast two-hybrid system after which the intersection of the experimental and predicted networks was assessed. From the total of 394 interactions the two-hybrid experiments generated a network of 233 interactions of which 59 were found by both approaches. *In vivo* relevance of the obtained interactions was studied with one SH3 domain-network and was found to correlate with the predicted interactions.

To eliminate some source of error obvious from the previous study (Tong et al. 2002), a different strategy to solve SH3 interaction networks was used (Landgraf et al. 2004). By this proteome peptide scanning approach an SH3 domain was challenged by a repertoire of peptides that it is likely to face inside the cell. A less strict, "relaxed consensus" binding sequence was determined for eight *S cerevisiae* SH3 domains whose "strict" binding preferences had been determined previously (Tong et al. 2002), and the relaxed sequences were used to search for ligands from the yeast proteome. About 1500 peptides were selected for each SH3 domain to be synthesised on membranes (SPOT synthesis technology). The membranes were probed by the corresponding SH3 domains and the interaction partners were detected. Many of the SH3 domains exhibited promiscuous binding to large numbers of common target peptides. One SH3 tested was capable of interacting with the predicted ligands *in vivo* but no correlation between the amount of the SH3 coprecipitated and the SPOT intensity or the corresponding dissociation constant was found (Landgraf et al. 2004).

Taken together, these approaches indicate that it is possible to predict a preferred target sequence based on the amino acid sequence of an SH3 domain that may help to help in the process of identifying the natural binding partners of the SH3 domain in question. In that sense these semi-computational tools may guide the research in to the right directions when studying SH3-mediated protein-protein interactions and elucidating cellular networks relying on them. But after all, these approaches will still be dependent on traditional experimental methods to verify the relevant interactions, as was also seen for example, in Tong et al. (2002). However, in the near future the high throughput experiments will be probably used more frequently.

A computational approach at best is only a model of an interaction and does not take into account for example unique molecular contacts in the SH3/target interaction. These are, like already discussed, the molecular contacts outside the ΦPxΦP core in the ligand and the variable and non-conserved RT- and n-Src loop structures on the SH3 binding surface. In addition, molecular context, localisation and the local concentrations of interacting proteins that also have an effect on the interaction are limitations for these prediction programmes.

5.9. Regulation of SH3 binding

The ability of a proline residue to exist stably as a *cis* isomer in a peptide bond could provide a means for regulation of the SH3/ligand binding recognition but thus far there is no evidence of such an event.

According to current views proteins with PDZ (postsynaptic density protein, disks large, zonula occludens) domains have been shown to play role in protein-protein interactions. However, PDZ domains were recently shown to bind also to a lipid ligand phosphatidylinositol 4,5-bisphosphate (PIP_2) to provide a novel strategy of regulation of interactions between a PDZ-containing protein and its (peptide or non-peptide) ligand (Zimmermann et al. 2002). A similar binding of a non-peptide ligand could be also surmised for SH3 domains but currently there is no evidence of such a phenomenon.

SH3 activity can be regulated by phosphorylation but the biological significance is still somewhat obscure. A major autophosphorylation site in the Bruton's tyrosine kinase (Btk) is the Y223 residue located in the SH3 domain (the first tyrosine in the conserved ALYDY-motif of the RT-loop ligand binding site, see Section 5.4.5.). Mutating this residue to phenylalanine potentiated the Btk transforming activity but did not affect its kinase activity suggesting that the SH3 domain regulates the Btk activity by modulating the binding to cellular targets (Park et al. 1996). *In vitro* binding studies have shown that when autophosphorylated, the Btk SH3 domain has altered specificity to target proteins as compared to the non-phosphorylated form (Morrogh et al. 1999). Similarly, a platelet derived growth factor (PDGF) induced the phosphorylation of c-Src Y138 residue located in the SH3 domain (the conserved Y residue in 3_{10} helix that is involved in ligand binding, see Section 5.4.5) in NIH3T3 cells and *in vitro*. Mutating this residue to phenylalanine did not affect the Src activation indicating that tyrosine phosphorylation on this site is not required for the activation induced by PDGF. On the other hand, phosphorylation of this residue reduced the ability of the SH3 domain to bind peptide ligands *in vitro* (Broome and Hunter 1997).

Phosphorylation of the flanking residues of the proline rich SH3 ligand may affect the SH3 binding. Mitogen activated protein (MAP) kinase is able to phosphorylate several serine residues in the C-terminal tail of Sos containing proline rich regions indicated in the Grb2 SH3 binding. This serine phosphorylation did not affect the Sos/Grb2 SH3 binding but instead had an effect on the ternary complex formation by decreasing the binding of this complex to epidermal growth factor receptor and Sch suggesting the mechanism of modulation of the Ras activation by forcing the Sos to other signalling pathways than to its ultimate target Ras (Rozakis-Adcock et al. 1995). Mutating the serine residues to alanines resulted in increase the binding affinity of Sos to Grb2 SH3 thus similarly providing a strategy of controlling the downstream target Ras activation (Corbalan-Garcia et al. 1996). In a more recent study the binding of the second SH3 domain of the adaptor protein Nck to PAK1 was shown to be dependent on the PAK autophosphorylation of the serine residue next to the NckSH3-2 binding motif. In the case of PAK1, serine phosphorylation abolished the NckSH3-2 binding *in vitro* and *in vivo* suggesting a mechanism to regulate PAK interactions with its SH3 containing downstream targets. Moreover, the PAK1 autophosphorylation on a serine residue also reduced the PIX SH3 binding to PAK1 on a differential site than the Nck SH3-2 binding site (Zhao et al. 2000).

Sam68 protein contains proline rich regions and was shown to interact with Fyn and PLCγ SH3 domains. These proline motifs are flanked with RG repeats that are, in turn, sites for arginine methylation by arginine N-methyltransferase. Arginine methylation was shown to reduce the binding of the SH3 domains to their target sequences whereas binding by WW (another proline recognising motif) domain containing protein was not affected suggesting that arginine methylation of proteins can affect certain SH3-mediated protein-protein interactions (Bedford et al. 2000).

Regulation of NADPH oxidase $p47^{phox}$ involves unusual intramolecular interactions between tandem SH3 domains of a $p47^{phox}$ dimer (Groemping et al. 2003; Yuzawa et al. 2004) and furthermore, represents an example how SH3 domains can be regulated via phosphorylation. The crystal structure revealed that the canonical ligand binding sites of both of the SH3 domains involved are brought side by side in such a way that they create a new binding groove that is capable of docking the N-terminal portion of the polybasic HTH region where the stretch GAPPR adopts a PPII helical conformation that contacts both of the SH3 domains. On one face the binding resembles the Class II binding whereas on the other side it is reminiscent of Class I binding. Serine phosphorylation of the polybasic region causes a conformational change that releases the binding and in the active conformation the substrate $p22^{phox}$ binds the tandem SH3 domain groove via its proline rich region (Groemping et al. 2003).

As there may be several SH3 domains in the same protein, in many cases there are also several (often overlapping) PxxP target sites. Depending on the physiological state of the cell these sites are used differentially to create a functional complex of interaction network. For example, Sos was recently reported to be involved in two distinct cellular pathways leading to Rac activation (Scita et al. 1999). In addition to the conventional view the Grb2 SH3/Sos-mediated Ras activation Sos was shown to form a tri-protein complex with Eps8 and E3b1/Abi-1. The formation of the complex was mediated via the SH3 domain of Eps8 binding to E3b1/Abi-1 and Abi-1 is used to bring together Eps8 and Sos. It was shown that both Grb2 SH3 and E3b1/Abi-1 SH3 domains utilise the same binding site in the Sos protein suggesting that Sos exists in two different molecular complexes in cells either bound to Grb2 SH3 or to the triprotein complex Sos/E3b1/Eps8. This leads to divergent signalling pathways depending on the SH3 ligand that occupies the Sos binding site and may allow a coordinated activation of Ras and Rac (Innocenti et al. 2002).

5.10. Inhibition and engineering of SH3-mediated interactions

Signal transduction pathways are targets for drug development. The importance of SH3-mediated interactions in critical signalling processes that result in malignancies and other pathological states has resulted in attempts to generate drugs that would interfere with the SH3 mediated processes. However, the promiscuous nature of the SH3-ligand binding is a challenge for these approaches. Ideally the reagents should be of low toxicity, very high selectivity and affinity to be used as targets for the SH3-mediated processes. Research concerning the development of SH3 inhibitors has so far at least provided new insights to drug design as well as offered reagents for a chemical biology approach to the SH3-mediated interactions.

Prompted by the studies on the natural SH3 specificity, for example, specific inhibitory peptides binding to an SH3 domain have been reported. Based on a known tight

interaction of the peptide from the C3G protein to the first SH3 domains of Crk and CrkL (Knudsen et al. 1995) a rational design together with series of mutagenesis rounds were performed to obtain the peptides for most optimal binding to Crk and CrkL SH3 domain (Posern et al. 1998). Affinities of these peptides increased up to 20-fold compared to the wild-type ones. Testing these peptides against a panel of other SH3 domains showed that they were highly selective to Crk and CrkL SH3 domains only; showing also only very limited binding to cellular proteins. Moreover, these peptides were capable of competing with the complexes between the Crk SH3 and its natural ligands in a concentration dependent manner. In a subsequent study these peptides were attached to an antennapedia-derived internalisation signal to create cell-penetrating peptides (Kardinal et al. 2000). These peptides were able to disrupt the Bcr-Abl/CrkL complexes and inhibited cell proliferation in the primary chronic myeloid leukaemia (CML) blast cell cultures as well as in cell lines established from Bcr-Abl positive patients. This suggested that the Bcr-Abl/CrkL complexes largely depend on the first SH3 domain of CrkL and that disrupting these SH3-mediated interactions could affect the proliferative signals in these cells.

Combinatorial approaches have been widely used in the drug design. A modified version of phage display (a mirror-image phage display) was developed to overcome the degradation of natural L-amino acids in cells and therefore provides a starting point for designing novel drugs. In this approach an SH3 domain composed of D-amino acids was used to select for peptide ligands made of L-amino acids. Consequently, the D-amino acid version of the peptide should interact with the natural SH3 composed of L-amino acids (Schumacher et al. 1995). Screening of this library by Src SH3 domain as a D-enantiomer resulted in ligand peptides with no sequence similarity to a proline rich peptide. A (cyclic) D-peptide of the ligand was created and it was shown to bind to L-Src SH3 domain. NMR studies showed that this peptide partly occupied the typical ligand binding site of the SH3 domain.

Since many SH3 domains occur in conjunction with SH2 domains, bivalent consolidated ligands were generated to recognise these domains simultaneously in order to enhance affinity and specificity of ligand binding (Xu et al. 1999). Among the tested, two types of ligands were shown to bind with high affinity (improvement of an order of a magnitude) and specificity to SH2-SH3 dual domains.

To study additional binding sites on the SH3 domain a bivalent ligand for the C-terminal SH3 domain of Sem5 (a *D. melanogaster* Grb2 homologue) was engineered by connecting the natural ligand peptide derived from Sos by a tetraglycine linker to a disulfide-closed random hexapeptide and the binding optimised by phage display (Ferguson et al. 2004). The selection revealed a single peptide whose affinity to its target SH3 domain was very high (K_d value of 0.048 μM), over three orders of magnitude higher than for the Sos-derived peptide alone. A variant containing a triglycine linker was even more potent in binding (K_d value of 0.025 μM). NMR structure revealed that the bivalent peptide was co-bound to the conventional SH3 domain ligand binding surface but made additional contacts with other regions of the SH3 domain. However, the additional region was different from the reported non-PxxP binding site of the Grb2 SH3 domain (Kami et al. 2002) suggesting that these ligands could be useful in studying more of the Grb2 SH3 domain-mediated interactions. Specificity of the bivalent peptide was not assessed in this study.

Several protein folds have been studied as alternative scaffold to be used as antibody mimetics. For example, a large library the tenth type III domain of human fibronectin (FN3) fold containing mutations in two of the loop structures was recently used to target SH3 domain of Src (Karatan et al. 2004) resulting in clones that bound to specifically to Src SH3 domain and were able to compete with a Src SH3 specific ligand peptide binding. Two of the Src SH3 specific monobodies were further examined by NMR spectroscopy revealing the other monobody that lacked a proline-rich sequence in its randomised loops to bind in a slightly different position on the Src SH3 surface as the Src SH3-specific peptide. The strongest binding monobody was also shown to recognise the recombinant Src SH3 domain as well as endogenously expressed Src from cell lysates further confirming the utility of these monobodies as affinity reagents both in research and diagnostics.

A split-pool combinatorial chemistry approach was used to engineer ligands that contain non-peptide binding elements for the Src SH3 domain. The library consisted of peptides containing an invariant Class I Src SH3 binding peptide and an array of organic monomers N-terminal to the proline-rich peptide (Combs et al. 1996; Feng and Schreiber 1997). The tightest binding ligand isolated showed an affinity of K_d value of 3.4 μM towards the Src SH3 domain. NMR structures of two of the isolated peptides in complex with Src SH3 showed that the non-peptide part occupied the specificity pocket of the Src SH3 domain by different means, mainly by hydrophobic contacts with the RT-loop, compared to peptide ligands (Feng et al. 1996).

As discussed (Section 5.5.), N-substitution of the PxxP defining proline residues is the main feature recognised by the SH3 domains (Nguyen et al. 1998). Accordingly, in order to achieve more specificity (and affinity) for the SH3 binding, consensus proline residues from a Sos-derived peptide were replaced with various nonnatural N-substituted glycine residues (dubbed peptoids) and were tested for binding to Sem5, Grb2, Crk and Src SH3 domains. It was shown that many of these peptoids were capable of replacing the proline in the consensus sequence and showed binding to the SH3 domains. The affinities of the peptoids were comparable to or better than the natural peptide. One of the peptoids was capable of binding to Grb2 SH3 domain with a very high affinity (K_d value of 0.040 μM) and specificity showing no (Sem5 and Crk) or negligible (Src) binding to other SH3 domains tested. Other specific domain/ligand pairs showed 5-25 fold improvement in the binding affinity. More recently the same authors replaced multiple prolines in the Sos-derived peptide with chemically more diverse N-substituted peptoids than in the previous study and studied their use us proline mimetics to enhance ligand selectivity (Nguyen et al. 2000). Single substitutions on the proline residues in the core resulted in peptoids showing significantly increased affinity and specificity. Substituting both of the prolines in the ΦPxΦP core resulted in peptoids that either bind as well or better than the wild-type ligands. In most compounds tested the double substitution caused an intermediate binding affinity when compared to the singly substituted peptoids. Substitution of (multiple) prolines outside the core region resulted in binding affinities close to the wild-type peptides. In addition, N-substituted peptoids were also introduced in peptides that contained flanking regions optimised for high-affinity binding to Src and Crk SH3. One of these compounds resulted in very high-affinity binding to Crk SH3 domain (K_d value of 0.008 μM) but it was still able to interact also with Src and Grb2 SH3 domains with a reasonable micromolar affinity. To engineer a ligand having an orthogonal selectivity for an SH3 domain, other peptoids were introduced into the Crk-peptide background. This manipulation reduced their

binding affinity to Crk SH3 (K_d values of 0.5 μM and 2 μM) but more importantly, showed no detectable binding to Src and Grb2 SH3 domains.

A completely non-peptide small-molecule ligand UCS15A was originally introduced as a Src inhibitor capable of interfering with the SH3-mediated protein-protein interactions of Src as well as other proteins. However, *in vitro* studies suggested that the inhibitor and its analogs do not bind to the SH3 domain but instead to the proline-rich regions of the target proteins (Oneyama et al. 2002; Oneyama et al. 2003). Recently, another group reported of designing of non-peptide small molecule ligands that recognise the SH3 domain (Inglis et al. 2004). Structure-based design was used to target the Tec SH3 domain with small heterocyclic compounds (for example, 2-aminoquinoline). NMR studies showed that the compound bound to the conserved residues in the SH3 specificity pocket. The affinity of the SH3/2-aminoquinoline was measured to be of K_d value 125 μM. The compound was capable of competing with the SH3 interaction with a proline-rich peptide. The interaction involved the conserved tryptophan (normally contacting one of the core proline residues) well as an aspartate side chain located in the RT loop. Derivatives of the 2-aminoquinoline were synthesised resulting in ligands that showed up to six-fold improvement in affinity compared to the original compound. They were able to compete with the proline-rich peptide binding for the SH3 domain. Specificity of, for example, the 2-aminoquinoline was tested in a competition assays for a set of different SH3 domains. It was found that the compound bound to Hck SH3 domain very poorly and not at all to Fyn SH3 domain. On the other hand, binding to the second SH3 domain of Nck was similar to the Tec SH3 domain. Hck and Fyn SH3 domains resembled more the Tec SH3 domain all having the aspartate in the same position in the RT-loop whereas Nck SH3-2 contained glutamate on this position showing the contextual discrimination of the aspartate and favouring of the glutamate in this position. However, some specificity could be achieved already with a simple compound providing a scaffold for further structure-affinity relationship studies.

6. AIMS OF THE STUDY

Nef has been shown to exhibit selective binding characteristics to a subset of Src family SH3 domains. Hck SH3 domain binds to Nef with an unusually high affinity and specificity. The structural basis of this high-affinity interaction lies on the "tip" of the Hck SH3 RT-loop that occupies a hydrophobic binding pocket on the Nef surface relatively distant from the canonical PxxP motif. The aim of the study was to characterise the SH3 binding function of Nef in correlation with association with a cellular protein as well as to study further the ligand recognition in the SH3-mediated protein interactions using the Nef/Hck SH3 interaction as a model.

Detailed aims of the present study were:

1. To confirm the effect of mutations of residues in Nef implicated in SH3 binding based on an X-ray structure of a Nef/SH3 complex on the capacity of Nef to bind Hck SH3.

2. To examine the role of SH3 binding and other Nef residues potentially involved in protein interactions in cellular association of Nef with PAK2.

3. To develop Hck-derived modified SH3 domains (RRT-SH3) that would bind to Nef with an unnaturally high affinity and study how these "tailor-made" SH3 domains could act as Nef inhibitors.

4. To examine if similar Hck-derived RRT-SH3 domains could be also targeted to ligands that normally bind poorly to Hck SH3.

7. MATERIALS AND METHODS

7.1. Plasmid constructs

7.1.1. Eukaryotic expression vectors and reporters

Eukaryotic expression vectors for HIV-1 Nef mutants were generated by PCR and cloned to pcDNA3 vector (Invitrogen). Expression vectors for activated cdc42 ($cdc42^{V12}$), Nck and myristoylated Nck SH3-2 (myr-Nck SH3-2), pEBB (Mizushima and Nagata 1990) and pEBB vector that contained a v-Src N-terminal myristoyl tail and a C-terminal HA-epitope tag (pEBB-myr-HA) were obtained from Dr. B. Mayer (University of Connecticut, Farmington). The NL4-3 Nef allele R71 (Saksela et al. 1995) and SIVmac239 Nef were cloned into pEBB vector provided with a myc epitope tag, the SH3 domains of Hck and the HIV-1 Nef or SIV Nef selected RRT-SH3 clones were cloned as GST fusions into pEBB-myr-HA vector. The vector for SV40-CD4 was from Dr. N. Landau (Salk Institute, La Jolla). The NFAT-luciferase construct containing three tandem short binding sites for NFAT promoter (-286 to -257 of human IL-2 enhancer) in front of minimal IL-2 promoter was provided by Dr. Y. Choi (Rockefeller University, New York). The pLacZ reporter plasmid was created by inserting β-galactosidase cDNA into pEBB.

7.1.2. Bacterial expression vectors

MBP-Hck SH3 was cloned by inserting a fragment encoding for a 56 amino acid Hck polypeptide (starting IVVA-, ending -VDSL) into pMAL-c2-S, a modification of a vector pMAL-c2 (New England Biolabs) containing three in-frame stop-codons after the polylinker. Bacterial expression vectors for HIV-1 Nef mutants are described in Saksela et al. 1995 or were generated by PCR and cloned to pGEX-4T-1 vector (Amersham Biosciences). GST-Hck SH3 vector has been described in Saksela et al. 1995. Bacterial GST-RRT-SH3 vectors were generated by PCR of the RRT-SH3 fragments from the corresponding phagemids, and cloned in pGEX-4T-1 (Amersham Biosciences).

7.1.3. Phagemids

To produce a monovalent phagemid for Hck SH3 expression, a fragment of human Hck (coding for amino acid residues 62-118) was cloned into the pCANTAB-5EP, a modified version of pCANTAB-5E (Amersham Biosciences) containing an extra PstI cloning site. To create the monovalent phage library of RRT-SH3 domains with randomised residues 69-74 of the human Hck, a degenerate sense primer extending over the RT-loop encoding region of Hck was used, and the resulting fragment was cloned into pCANTAB-5EP. The multivalent phagemid vector pG8H6 (Jacobsson and Frykberg 1996) was a gift from Dr. L. Frykberg (Uppsala, Sweden). The multivalent library of Hck SH3 derived RRT-SH3 domains was generated by PCR using a degenerate primer as for the monovalent library. In addition, a TAG-stop codon was inserted in front of the SH3 to reduce the fusion protein expression using a *supE E.coli* strain TG1.

7.2. Antibodies

The sheep polyclonal anti-HIV-1 Nef/anti-GST antiserum was provided by Dr. Mark Harris (Leeds University, UK) and the mouse monoclonal anti-HIV-1 Nef antibodies (2A3, 3A2, 6.2, 2H12, 3E6, 3F2.1, 3D12 and 2F2) were from Dr. Kai Krohn from our institute. The rabbit polyclonal anti-PAK antibody (N-20) and the anti-phosphotyrosine

antibody PY99 were purchased from Santa Cruz Biotechnology. Anti-HA antibody was from BabCO, anti-CD3 (HIT3a) and the phycoerythrin-conjugated anti-CD4 antibodies from Pharmingen, monoclonal anti-myc antibody 9E10 from Roche and anti-CD28 monoclonal antibody from CLB.

7.3. Cell culture and transfections

All the cell lines were obtained from American Type Culture Collection (ATCC). The human T-cell leukaemia cell line Jurkat (JE-6) and myeloid U937 and K562 cells were cultured in RPMI 1640 (Bio Whittaker) supplemented with 10% fetal calf serum (FCS, Bio Whittaker) and 2 mM glutamine (Invitrogen). Human embryonic kidney fibroblast cells (293T) were cultured in Dulbecco's modified Eagle's medium (DMEM, Bio Whittaker) supplemented with 10% FCS and 2 mM glutamine.

Transfections of 293T cells were done by the calcium phosphate precipitation method or by Lipofectamine transfection reagent (Invitrogen). Jurkat E-6 cells were transfected with the DMRIE-C (Invitrogen) reagent according to the manufacturer's instructions.

7.4. Immunoprecipitation and *In vitro* kinase assay (IVKA)

Cells were washed once with standard phosphate buffered saline (PBS) and lysed in *in vitro* kinase assay lysis buffer (50 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) pH 7.4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM ethylene glycol-O,O'-bis-[2-amino-ethyl]-N,N,N',N',-tetraacetic acid (EGTA), 1.5 mM $MgCl_2$, containing 1 mM phenylmethylsulfonyl fluoride (PMSF), 10 μg/ml aprotinin, 10 mM NaF and 1 mM sodium orthovanadate). Lysates were precipitated with anti-Nef and anti-PAK antibodies (**I**) or with anti-Nef antiserum (**III**) by using protein A- or protein G-Sepharose beads (Sigma). For IVKA, after washing with IVKA lysis buffer and kinase assay buffer (50 mM HEPES pH 7.4 and 5 mM $MgCl_2$) 32[P]γATP (2.5 μCi; Amersham Biosciences) was added to the immunoprecipitates followed by a 30-minute incubation at 37°C. The beads were washed with PBS, boiled in Laemmli sample buffer, followed by SDS-polyacrylamide gel electrophoresis (PAGE) and detection by autoradiography.

7.5. Western blotting

For Western blotting, immunocomplexes or total protein samples were separated by SDS-PAGE and transferred onto nitrocellulose membranes for subsequent immunoblotting with specific antibodies. Biotinylated secondary antibodies, and depending on the detection method, streptavidin-horseradish peroxidase or alkaline phosphatase (AP) conjugates were used followed by enhanced chemiluminescence (ECL) or Nitro blue tetrazolium chloride/5-Bromo-4-chloro-3-indolyl phosphate toluidine salt (NBT-BCIP) substrate, respectively. In the case of polyclonal anti-Nef (sheep) antibody (**II**) the detection was performed without a biotinylated secondary antibody by using an anti sheep IgG-AP GT-34 (Sigma) conjugate and NBT-BCIP.

7.6. Recombinant protein production in *E.coli*

Expression and purification of GST- and maltose binding-fusion proteins in *E.coli* were carried out according to manufacturers' instructions (Amersham Biosciences and New England Biolabs, respectively). After specific elution from their respective affinity beads the proteins were concentrated and buffers changed to desired ones by successive rounds of micro concentration with Centrex UF2 columns (Schleicher & Schuell). For

GST-Nef proteins (**I**) the buffer was changed to HBS (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA) containing 0.005 % (v/v) surfactant P20 and 1 mM dithiotreitol (DTT) in which they were subjected to thrombin cleavage (1 U thrombin/0.2 mg fusion protein) for 4-6 hours at 37°C. Maltose binding fusion proteins (**I**) were biotinylated using EZ-Link™ Sulfo-NHS-LC-Biotin reagent as suggested by the manufacturer (Pierce) and subjected to additional rounds of micro concentration and a buffer change to HBS to remove any free biotin. In other studies (**II-IV**) after eluting the recombinant fusion proteins the buffer was changed to PBS by micro concentration.

7.7. Surface plasmon resonance measurements (I)

Surface plasmon resonance (SPR) experiments were performed with the Biacore X apparatus (Biacore). An SA-sensor chip with preimmobilised streptavidin was coated either with biotinylated MBP (reference channel) or MBP-Hck SH3 (test channel) proteins (in a concentration of 100 ng/ml in HBS containing 1 mM DTT at the flow rate of 5 µl/minute). The attachment of the ligands was monitored by the changes of the refractive index and was set to ~2200 (MBP-Hck SH3) and 2000 (MBP) response units (RU) corresponding to the relative difference in their molecular weights. After immobilising the ligands the chip was subjected to three rounds of preregeneration cycles, which were also subsequently applied once after each Nef injection. A cycle consisted of one minute pulses of subsequent steps of glycine buffer pH 2.2, 0.05 % SDS and 4 M urea and 5 minutes of HBS running buffer containing 1 mM DTT. Some loss of the refractive index was observed during the first and second cycles of the preregeneration but no longer during the third cycle of this treatment. The injections of Nef were done in concentrations ranging from 4.0 to 0.0125 µM, with a flow rate of 5 µl/min in HBS containing 1 mM DTT. Each chip was used for approximately for 50 Nef injections during which no loss of the immobilised ligand or the capacity of the chip to bind to a standard Nef solution was observed. The sensorgrams, in which the refractive index values from the reference channel were subtracted to give corrected relative resonance units (cRU) were analysed by the BIAevaluation (v3.0) software (Biacore). The Scatchard plots and line fitting were based on values from the sensorgrams at 20-minutes post injection time points.

7.8. Production of the phage libraries

When creating the monovalent phage library (**II**) the phagemids containing the SH3 domains with randomised RT-loops (RRT-SH3) were electroporated into *E.coli* TG1 cells. A total of $137 \cdot 10^6$ individual colonies were pooled and grown overnight at 30°C in 2x YT medium containing 100 µg/ml ampicillin and 2% (w/v) glucose (2xYT-AG). Cultures were diluted one in ten in 2xYT-AG, and infected with $5 \cdot 10^8$ pfu/ml of M13KO7 helper phage (Amersham Biosciences) at 37°C. After two hours cells were pelleted and the medium changed to 2x YT medium containing 100 µg/ml ampicillin and 50 µg/ml kanamycin (2xYT-AK). The double resistant cultures were grown overnight at 37°C, the supernatant containing the infectious recombinant phages was collected, passed through a 0.45 µm filter and stored in aliquots at 4°C. The multivalent phage library (**IV**) of RRT-SH3 domains was constructed in the pG8H6 vector. A total of $420 \cdot 10^6$ individual colonies were obtained by electroporation into TG1 cells, and used to produce recombinant phages via infection with the M13KO7 helper phage. In addition, sublibraries of approximately 500 000 recombinant clones consisting of PCR amplified RRT-SH3 inserts initially selected using the multivalent system were inserted into monovalent pCANTAB-5EP vector as described above.

7.9. Phage selection

Six-well plates were coated with 10 µg/ml GST-Nef, GST-NefR90, GST-Nef-PA1 (**II**), GST-SIV Nef (**IV**) or plain GST in 50 mM sodium carbonate pH 9.6 at 4°C overnight. The wells were blocked with 5% milk in PBS/0.05% Tween 20, and washed briefly with PBS before 10^7-10^{11} pfu of recombinant phages per well were added (higher titers used in the early rounds of selection), followed by incubation for 2 hours at RT. In some experiments involving NefR90-coated wells, the phage solution was supplemented with 10 µg/ml of soluble wild-type Nef (**II**). After incubation with the phages the wells were washed six times (5 min) with PBS/0.05% Tween 20 and three times with PBS. In most experiments the bound phages were eluted with a small volume of PBS containing an excess of (150 µg/ml) of the corresponding Nef protein (**II**). Alternatively, the TG1 cells to be infected were added directly to the washed wells (**II**, **IV**). In both cases, the bacteria were first grown in 2xYT to log phase from an overnight culture, infected with the affinity-selected phages for 2 hours at 37°C, after which a sample of 1% was removed for determination of the infectious titer of the selected phages by plating on ampicillin plates. These plates also served as indicators for the enrichment of specific clones when compared to plates infected with phages from GST-coated wells processed in parallel. The remaining of the infected bacteria were supplemented with 100 µg/ml ampicillin and 2% glucose, and subjected to a subsequent infection with $5 \cdot 10^8$ pfu/ml of M13KO7 helper phages for 2 hours at 37°C, after which they were pelleted and resuspended into an equal volume of 2xYT/AK. After an overnight incubation the amplified recombinant phage supernatants were collected as described above, and used for the subsequent round of selection/infection. Usually after 8 cycles of selection ≥ 12 colonies were picked for preparation of phagemid DNA, and the RRT-SH3 inserts were sequenced using ABI Prism 310 (Applied Biosystems).

7.10. Competitive Nef/SH3 binding assay (II)

MaxiSorp F8 strips (Nunc) were coated with the different GST-Nef proteins (200 ng in 100 µl per well) overnight at 4°C, followed by a 30 min incubation with 1.5% BSA in washing buffer (WB; PBS/0.05% Tween 20) at RT, and washed twice with WB. The unlabelled SH3 proteins used as competitors were diluted into WB that contained 1.5% BSA and a large molar excess of plain GST (4 µM). 50 µl of this solution was mixed with an equal volume of the probe (biotinylated SH3 in WB) and added to the wells resulting in a final probe concentration of 66 nM (Hck SH3/Nef and RRT.A1/NefR90 assays). After a one-hour incubation at RT the wells were washed three times with WB, and 100 µl of 1:2000 dilution (in WB) of streptavidin-biotinylated horseradish peroxidase complex (Amersham Biosciences) was added per well. The plates were incubated 45 minutes at RT and washed again three times, after which their peroxidase activity was measured using 1,2-phenyldiamine-dihydrochloride (OPD; 0.6 mg/ml; Fluka AB) as a substrate. The enzymatic reactions were stopped after 10 minutes by adding 50 µl of 2 M sulphuric acid, followed by optical density measurement at 492 nm using a Victor 1420 Multilabel Counter (Wallac).

7.11. Luciferase reporter assays (III, IV)

Twenty hours after transfection using the pEBB-lacZ as an internal transfection efficiency control, an NFAT luciferase reporter plasmid and expression vectors, cells were either left untreated or stimulated with 100 ng/ml of PMA or 50 ng/ml of anti-CD3 antibody for 4-6 hours. After harvesting cells were washed with PBS, lysed to cell culture lysis reagent (Promega) and the luciferase activity measured with Promega

reagents and Luminova 1254 luminometer (ThermoLabsystems). In order to normalise the luciferase values for the transfection efficiency the β-galactosidase activities in the lysates were determined by using o-nitrophenyl ß-D-galactopyranoside (ONPG; Sigma) as a substrate and by measuring absorbances at 420 nm.

7.12. CD4 downregulation assay (III)

For the CD4 downregulation assay 293T cells were co-transfected with human CD4 and GFP alone or together with Nef R71 and RRT-SH3 clone B6. Forty-eight hours after transfection cells were harvested, washed with PBS containing 0.1 % BSA, and stained with PE-conjugated anti-human CD4 (dilution 1:10) for 30 minutes. After two washes with PBS/0.1 % BSA the stained cells were suspended in PBS containing 0.5 mM EDTA. The CD4 surface expression among the GFP positive cells was examined using FACScan (Becton Dickinson).

7.13. Glutathione-S-transferase pull-down assays (III, IV)

293T cells were harvested forty-eight hours after transfection of Nef and GST-RRT-SH3s expressing plasmids (**III**, **IV**), washed with PBS and lysed in *in vitro* kinase assay lysis buffer. Lysates (1 mg of total cellular protein) were incubated with glutathione Sepharose 4B beads for 4 hours at 4°C and washed with IVKA lysis buffer. The associated proteins were separated by SDS PAGE followed by immunoblotting.

To compare the ability of the RRT-SH3s to bind different phosphorylated substrates from cells (**III**), K562 and U937 cells were incubated with 1µM pervanadate at 37°C for 15 minutes. Pervanadate was prepared by mixing 20 µl of 1M Na-orthovanadate with 11 µl of 30% H_2O_2 and 69 µl of H_2O and incubated at room temperature for 15 minutes before use. After washing with PBS, $2 \cdot 10^7$ cells/ml were lysed in 1 ml of Nonidet P-40 lysis buffer (1% Nonidet P-40, 150 mM NaCl, 20 mM Tris-Cl pH 7.4, 0.5% sodium deoxycholate, 50 mM NaF, 1 mM PMSF and 10 µg/ml aprotinin). The lysates were preincubated with recombinant GST protein (10 µg/1 mg of total protein) at 4°C for two hours, followed by capture on glutathione Sepharose 4B beads. These pre-cleared lysates were split into aliquots corresponding to 500 µg of total cellular protein and incubated for 2 hours at 4°C with beads coated with 8 µg of GST or GST-SH3. The beads were washed three times with and once with PBS, boiled in Laemmli sample buffer, followed by SDS PAGE and anti-phosphotyrosine immunoblotting.

Epitope-tagged SIVmac239 and HIV-1 Nef proteins containing lysates were produced by transient transfection of 293T cells (**IV**). Approximately $20 \cdot 10^6$ transfected cells were collected 48 hours after transfection and lysed into Nonidet P-40 lysis buffer. The lysates were serially diluted (four-fold dilutions) into a similarly prepared lysate of untransfected 293T cells. After removal of small aliquots for monitoring their Nef content, 200 µl of these dilutions were incubated for 2 hours at 4°C with glutathione Sepharose 4B beads coated with 8 µg of plain GST of GST-SH3 proteins. After the incubation the beads were washed three times with PBS, and subjected to SDS-PAGE and immunoblotting.

8. RESULTS AND DISCUSSION

8.1. Role of SH3-binding in Nef/PAK2 complex

8.1.1. Structure-based mutagenesis of residues in the core domain of Nef (**I**)

HIV-1 Nef was reported to coprecipitate with a serine/threonine kinase NAK, for Nef associated kinase (Sawai et al. 1994) that was recently identified as PAK2 (Renkema et al. 1999). PAK proteins do not contain SH3 domains but former studies had suggested that the NAK association with Nef might involve SH3 binding (Wiskerchen and Cheng-Mayer 1996; Lang et al. 1997). The role of the SH3-binding capacity as well as a panel of other surface exposed residues of Nef in association with PAK2 was examined by point mutations introduced into Nef R71 background (summarised in Table I).

Table I
Summary of HIV-1 Nef mutants used in the study **I** and selectively in the studies **II-IV**.

Group	Variant
A	Parental NL4-3 Nef T71, and its wild-type variant R71
B	Mutations in the SH3-contacting residues in the PPII helix region[1]: P72A, P75A, P72A;P75A (PA-1), V74D, R77E, V74D;R77E
C	Mutations in the non-SH3-contacting residues in the PPII helix region: P69A, Q73P;L76A
D	Mutations in the SH3-contacting residues outside the PPII helix region: F90R (called as R90 in **II**)
E	Mutations in the non-SH3-contacting residues outside the PPII helix region: R106A, L112R, F121R

[1] Residues 71-77 in the NL4-3 Nef define the PPII helix region.

NL4-3 R71 (called as wild-type or just Nef from now on) is a derivative of the original laboratory adapted HIV-1 strain NL4-3 Nef (dubbed T71 in our studies), and has an arginine residue in the position 71 like majority of the naturally occurring Nef sequences have. The novel point mutations in the Nef sequence were based primarily on the Nef/SH3 X-ray structure (Lee et al. 1996) and the nuclear magnetic resonance (NMR) structure of unliganded Nef (Grzesiek et al. 1997). The minus-oriented SH3 ligand region in Nef involves the residues $P^{72}Q^{73}V^{74}P^{75}L^{76}R^{77}$ which are highly conserved in all Nef alleles (http://www.hiv.lanl.gov/content/index; Shugars et al. 1993). Of the SH3 contacting residues P72 and P75 are the PxxP defining prolines and the V74 and R77 are critical Class II (minus orientation) consensus residues. Due to the triangular structure of the PPII helix, Q73 and L76 are on the opposite side of the SH3-contacting residues and are not directly involved in the SH3 binding. The F90 residue in Nef is fairly distantly located from the PxxP/SH3 interface but it has been shown to contribute to the tertiary contacts between the SH3 RT-loop and Nef (Lee et al. 1996).

8.1.2. Mutations affecting the SH3 binding capacity of Nef (**I**)

The effects of the desired point mutations in Nef on its ability to bind SH3 proteins were studied by surface plasmon resonance using the Nef/Hck SH3 interaction (Lee et al.

1995) as a model. As reported (Lee et al. 1995), the wild-type Nef associated with Hck SH3 with a high affinity (K_d of 0.25 μM). A significant but not critical role of the R71 residue in the wild-type Nef allele was indicated here by the threefold lower affinity measured for the T71 allele. Disrupting the SH3-contacting residues in the PPII helix (Group B mutations in the Table I) resulted either in non-detectable SH3-binding affinity or in a large reduction in the affinity without completely abolishing the binding to Hck SH3 confirming that a functional PxxP motif is required for Hck SH3 binding. The residual binding of the singly mutated prolines in the PxxP motif probably arises from the other stabilising non-PxxP interactions between the Nef and Hck SH3 and that the residues replacing the prolines (alanines) still preserved the hydrophobic character of the region. As expected, the mutation of the critical F90 to an arginine (Group D in Table I) in the Nef hydrophobic pocket contacting the SH3 RT-loop reduced the binding affinity about 10-fold compared to the wild-type Nef (K_d of 1.99 μM). Changing residues not contacting the SH3 domain, but adjacent to the PxxP motif (Group C mutations in Table I) did not decrease but rather somewhat increased the Hck SH3-binding affinity of Nef. The slightly improved SH3-binding capacity of the Q73P;L76A mutant was probably due to its stabilising effect on the PPII helix (see Section 5.4.4; Yu et al. 1994). Mutating residues outside the PPII helix region (Group E mutations in Table I) did not affect the SH3-binding capacity of these Nef variants, and the affinities were comparable to the wild-type Nef binding to Hck SH3.

8.1.3. Mutations affecting the capacity of Nef to associate with NAK (PAK2) (**I**)

To study if NAK association of Nef correlated with SH3-binding capacity of the different mutants, these Nef variants together with the active p21 GTPase $cdc42^{V12}$ (to induce the autokinase activity of PAK2) were transfected to 293T cells followed by immunoprecipitation with Nef antibodies and a subsequent *in vitro* kinase assay to detect Nef-associated PAK2 activity. Only the wild-type Nef alleles coprecipitated PAK2. Without exceptions, all the mutants that failed to bind to Hck SH3 were also negative in the PAK2 association indicating that the SH3 binding of Nef is required. On the other hand, Nef variants with mutations in residues that were not involved in the SH3 binding function (Group C and E in Table I) were also unable to associate with PAK2.

For example, the double mutant Q73P;L76A that changes the non-SH3-contacting residues within the Nef PxxP motif was negative in association with PAK2 but was still capable of binding to Hck SH3 domain. Examination of the available structural data from this region of Nef revealed that Q73 and L76 are on the surface of Nef even when the SH3 is occupying the PxxP region. Together with L112 and F121 these residues form an exposed hydrophobic surface that could serve as a binding site for PAK2 or some other critical component in the PAK2/Nef/SH3 complex (a dashed area in Figure 8.). Additional Nef variants L112R and F121R (group E in Table I) were therefore generated and indeed found to be capable of SH3-binding but unable to associate with PAK2 providing more evidence for a possible role for this binding surface in additional non-SH3 mediated contacts with PAK2. However, residues of this hydrophobic surface have been also suggested to take part in the oligomerisation of Nef (Arold et al. 2000; Liu et al. 2000; Ye et al. 2004) and therefore the mutations affecting these residues would disrupt the Nef oligomers and the effects of these changes on PAK2 activation would be more indirect.

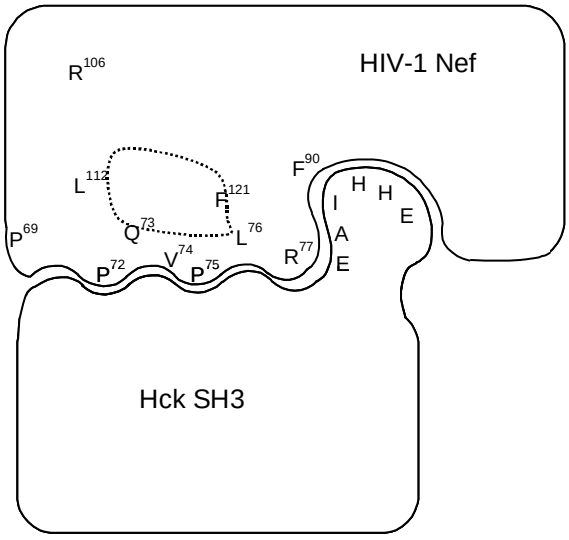


Figure 8.
Capacity of Nef to associate with PAK2 is dependent on the SH3 binding interface as well as P69 and R106 residues and the hydrophobic binding surface (dashed line) lined by the residues Q73, L76, L112 and F121.

The R106A and P69A mutants were negative in PAK2 association as had been earlier reported (Sawai et al. 1995; Wiskerchen and Cheng-Mayer 1996). The position P69 residue is the most amino terminal residue of the Nef core domain and the first proline of the tetraproline repeat in Nef, but did not contribute to the SH3 binding. Instead, it could be envisioned to play a role in the correct positioning of the N-terminal flexible arm of Nef. When this residue is mutated to an alanine PAK2 association could be abolished either by the interference of the N-terminal residues with the aforementioned putative hydrophobic PAK2 binding site or alternatively by disrupting regions N-terminal of the P69 that could be involved in additional contacts with PAK2. Sawai et al. (1995) showed that the conserved N-terminal region and membrane association of Nef are required for the NAK association that makes the latter possibility more probable.

In conclusion, the mutations affecting the residues that are part of the SH3 binding interface abolished binding of Nef to Hck SH3, as expected while the mutants with amino acid changes in residues not implicated in SH3 binding in the X-ray structure by Lee et al. (1996) behaved as the wild-type Nef. Only the wild-type Nef alleles were able to associate with PAK2. Variants with mutations affecting the SH3 binding function of Nef did not coprecipitate PAK2 indicating that the SH3 binding function of Nef is required for the PAK2 association. Since PAK2 itself does not contain an SH3 domain it is plausible that the interaction with Nef is indirect and mediated via some other SH3-containing protein. Mutations in the selected residues not involved in SH3 binding were also deficient in associating with PAK2 suggesting an important function of these residues in PAK2/Nef interactions, possibly forming a putative binding site for PAK2 itself or another component of this interaction (or due to more indirect effects via preventing Nef dimerisation/oligomerisation).

8.2. Phage-display of the libraries of modified SH3 domains (II, IV)

In contrast to the high affinity binding to Hck SH3, Nef has almost a 100-fold lower affinity to the SH3 domain of Fyn. Engineering of the Fyn SH3 domain to be more "Hck-like" resulted in high-affinity interaction (Lee et al. 1995). Prompted by this finding phage display approach was taken to study if it would be possible to generate a large panel of SH3 domains with desired ligand binding properties by manipulating the RT-loop.

A phage-display library of Hck-derived artificial SH3 domains, in which the six non-conserved, Hck-specific RT-loop residues were replaced by a random hexapeptide (termed RRT-SH3 for randomised RT-loop), was constructed. As discussed (see Section 5.4.3.), the RT-loop is among the most variable regions in the otherwise conserved SH3 domains and together with the n-Src loop aligns the binding surface for the PPII helical ligand (see Figure 4.). As the relatively well-conserved acidic residues of the RT-loop are involved in accommodating the basic residues of Class I and Class II ligands the central non-conserved and highly variable residues can engage in interactions that could be unique to individual SH3/ligand pairs and thereby offer more to the specificity of the SH3 binding.

The SH3 domains expressed on the surface of the bacteriophages were functional and specific for a cognate ligand, because binding of the wild-type Hck SH3 phage by wild-type Nef resulted in enrichment of infectious phage particles that was orders of magnitude higher than using PA-1 mutant (**II,** see Table I for the Nef mutant nomenclature). The monovalent library of RRT-SH3 domains was essentially used in (**I**) and the multivalent library in (**IV**).

8.3. RRT-SH3 proteins targeted to wild-type Nef (II)

8.3.1. Identification of the RRT-SH3 proteins with increased binding to wild-type Nef (**II**)

Phage selection of the monovalent RRT-SH3 library with wild-type Nef resulted in novel SH3 ligands that bound Nef in a strictly PxxP-dependent manner since the proline deficient PA-1 mutant of Nef was not able to bind to these clones. The randomised RT-loop region of the selected clones did not resemble wild-type Hck SH3 RT-loop or any other naturally occurring SH3 domain found in the GenBank database. Summary of the RT-loop sequences of the most prevalent RRT-SH3 clones after the 7th selection round is presented in Table II.

The RRT-SH3 clones were categorised in four sequence families (A, B, C and D) depending on three or more identical residues within each group. A most notable feature of all of the RRT-SH3 clones is a serine residue in the position 2/6 of the randomised sequence. Additionally, these wild-type Nef selected RT-loops were particularly rich in aromatic and proline residues. A possible role of the serine residue in the position 2/6 in the RT-loop was not studied in the case of wild-type Hck SH3.

Table II
Phage-selected RT-loop sequences, and the estimated binding affinities of the corresponding SH3 domains to wild-type or R90 variant of Nef. The measured affinities for the Hck SH3/Nef and Hck SH3/R90 complexes are from (**I**). The RRT-SH3 affinities for wild-type Nef are estimated from the data obtained from (**II**) according to the Cheng and Prusoff equation (see Section 8.3.2. for details). The affinity of RRT.A1 for R90 Nef was estimated to be 40 times higher than Hck SH3 (measured in **I**) and was therefore set to 50 nM. The affinities of the remaining RRT-SH3 domains for R90 variant were proportioned to this value according to the data obtained from (**II**).

SH3 domain	RT-loop	Estimated affinity (K_d) to wild-type Nef (M)	Estimated affinity (K_d) to R90 Nef (M)
Hck SH3	EAIHHE	$2.5 \cdot 10^{-7}$	$2.0 \cdot 10^{-6}$
Wild-type Nef-selected clones:			
RRT.A1	VSWSPD	$7.1 \cdot 10^{-9}$	$5.0 \cdot 10^{-8}$
RRT.A2	FSWSDT	$1.3 \cdot 10^{-8}$	$5.0 \cdot 10^{-8}$
RRT.A3	DSWSTS	-	-
RRT.A4	YSWSDM	$2.5 \cdot 10^{-8}$	$6.3 \cdot 10^{-8}$
RRT.B1	WSPFPS	$9.6 \cdot 10^{-9}$	$8.3 \cdot 10^{-7}$
RRT.B2	DSPFSF	$1.1 \cdot 10^{-8}$	$1.7 \cdot 10^{-7}$
RRT.B3	FSPFSF	-	-
RRT.B4	FSPFDW	$1.2 \cdot 10^{-8}$	$8.3 \cdot 10^{-8}$
RRT.B5	SSPFDW	-	-
RRT.B6	YSPFSW	$6.8 \cdot 10^{-9}$	$5.0 \cdot 10^{-8}$
RRT.C1	TSPFPW	$8.1 \cdot 10^{-9}$	$8.3 \cdot 10^{-7}$
RRT.C2	YSFFPW	$1.6 \cdot 10^{-8}$	$5.0 \cdot 10^{-7}$
RRT.C3	YSDFPW	$9.6 \cdot 10^{-9}$	$2.5 \cdot 10^{-7}$
RRT.C4	DSWFPW	$1.8 \cdot 10^{-8}$	$2.5 \cdot 10^{-7}$
RRT.D1	SSFYSS	$1.1 \cdot 10^{-8}$	$1.7 \cdot 10^{-7}$
Nef R90-selected clones:			
RRT.m1	QGFLDQ	$3.1 \cdot 10^{-7}$	$8.3 \cdot 10^{-9}$
RRT.m2	NAFLPS	$1.1 \cdot 10^{-7}$	$1.3 \cdot 10^{-8}$
RRT.m3	EAWSPL	$1.5 \cdot 10^{-8}$	$6.3 \cdot 10^{-9}$
RRT.m4	ESYSEW	$3.1 \cdot 10^{-7}$	$5.0 \cdot 10^{-8}$

8.3.2. RRT-SH3 domains can bind to wild-type Nef with high affinity (**II**)

We previously used SPR in measuring the dissociation constants for Nef/SH3 binding (**I**). According to the phage display results the novel RRT-SH3 clones were expected to bind to Nef at least an order of a magnitude better than Hck SH3 does. Limitations (aggregation and "stickiness") of using very low concentrations of recombinant Nef protein made SPR an unpractical method for measurement of the strongest of these

interactions. Therefore, to study the binding properties of the selected RRT-SH3 domains we developed a competitive assay, in which the RRT-SH3 proteins were tested for their ability to compete with a constant amount of biotinylated Hck SH3 for binding to immobilised Nef. However, since it is not formally possible to derive the exact K_d values based on these numbers, the strength of binding in (**II**) was expressed as such. Here the binding affinities of the selected RRT-SH3 domains used as competitors were roughly estimated by assuming that the ratio of RRT-SH3:Hck SH3 at 50% inhibition is proportional to the relative difference of their affinity to Nef as compared to the K_d value (0.25 μM) of the Hck SH3/Nef interaction according to the Cheng and Prusoff equation (Cheng and Prusoff 1973). All of the RRT-SH3 domains selected with wild-type Nef bound better than the parental Hck SH3. The estimated affinities from the data obtained from the competitive assay (**II**) are summarised in Table II.

8.3.3. RRT-SH3 domains recognise Nef by different strategies (**II**)

An isoleucine residue in Hck SH3 RT-loop packs against the F90 residue in a hydrophobic binding pocket of Nef and results in high-affinity binding (Lee et al. 1995; Lee et al. 1996). To study the dependence of the selected RRT-SH3 domains on a similar hydrophobic interaction involving F90, binding affinities of these RRT-SH3 clones against the Nef R90 mutant were determined. Since the wild-type Nef-selected RRT-SH3 clone A1 was in the initial experiments found to bind the R90 Nef mutant quite well, a competitive assay using this variant was developed. In this assay Hck SH3 was unable to reach the 50 % inhibition of the RRT.A1/Nef R90 interaction even with 16-fold molar excess which was the highest practical ratio for the assay involving 66 nM of labelled RRT-A1 (**II** and Figure 3. therein). Considering this result and the measured K_d value of 2.0 μM for the Hck SH3/Nef R90 interaction (**I**) together with the observation that binding of RRT-A1 to Nef R90 was clearly less strong than to wild-type Nef (K_d of 7.1 nM) the binding strength of RRT-A1/Nef R90 was estimated to be 50 nM (see also **II** and Figure 3. therein). The affinities of the rest of the RRT-SH3 domains to R90 were proportioned to this value. These estimated affinities from the data obtained from (**II**) are summarised in Table II.

The utility of the RRT.A1 in this assay readily showed that it did not critically depend on the intact F90 when binding to Nef. Among the other Nef selected RRT-SH3 clones the ability to bind to Nef R90 varied: some clones were efficient in competing with the RRT.A1 binding whereas others behaved more like Hck SH3 and were unable to compete with RRT.A1 for Nef R90 binding, and thus showing critical dependence on the F90 residue in Nef.

In conclusion, these results suggest that changing the composition of the Hck SH3 RT-loop creates novel SH3 domains that can bind Nef with high affinity (up to 40-fold better than Hck SH3). The binding of these SH3 domains to Nef is PxxP dependent. On the other hand, these RRT-SH3 domains recognised Nef with different molecular strategies. Depending of the RT-loop sequence the recognition of Nef was either “Hck-like” (depending on the F90) or did not discriminate between the wild-type or R90 mutant Nef. In the latter cases the binding affinity is presumably achieved by anchoring the engineered RT-loop residues to a surface of Nef that wild-type Hck SH3 does not use. However, binding to the R90 mutant was always weaker than to wild-type Nef.

8.4. RRT-SH3 proteins targeted to R90 variant of Nef (II)

8.4.1. Identification of RRT-SH3 proteins with increased binding to Nef R90 mutant

Our results suggested (see Section 8.3.3.) that depending on the RT-loop, the selected SH3 domains might recognise Nef by different strategies. Therefore, another selection using the Nef R90 mutant (see Table I) was performed to obtain clones that would preferentially select R90 instead of the wild-type Nef. RRT-SH3 clones showing an ability to bind efficiently to the Nef R90 variant could be rescued from the monovalent RRT-SH3 library. The most frequent Nef R90-selected RRT-SH3 domains (RRT.m1-m4) are presented in the bottom of Table II. The Nef R90-selected RRT-SH3 domains did not contain any of the sequence motifs characteristic to the wild-type Nef-selected RRT-SH3 domains. Also the serine 2/6 that was invariant in the Nef-selected clones was absent in all but one (RRT.m4) Nef R90-selected RRT-SH3 domains.

8.4.2. RRT-SH3 domains can bind to R90 Nef with high affinity (**II**)

A competitive binding assay for the RRT-SH3 clones that were selected with Nef R90 mutant revealed that all of these RRT-SH3 domains were efficiently able compete with binding of RRT.A1 to Nef R90 showing binding affinities comparable to the ones seen with the clones selected with wild-type Nef (see Table II). In this group the RRT.m3 seemed to bind to both Nef R90 and wild-type Nef whereas for the others the majority of the binding affinity was dictated by the mutant R90 residue in Nef. This shows that the engineered RT-loop provides them with a novel binding specificity that allows their targeting to a new ligand.

8.4.3. Engineering of binding interfaces may result in changes in binding characteristics

It was shown that by engineering the RT-loop of the Hck SH3 domain it is possible to increase the binding affinity for a known ligand Nef as well as find high-affinity SH3 ligands for a novel PxxP containing protein, in this case the R90 mutant of Nef. Structural basis of these interactions could help to elucidate how such protein-protein-interactions occur and would be of interest of future work. Structural studies have in many cases provided a framework for modifying binding interfaces. For example, engineering of the binding interface created a mutant growth hormone receptor (that was unable to bind wild-type hGH) but was able to bind to a variant hGH that was created by mutational randomisation of five hGH residues and selected by phage display. X-ray structures of the hGH liganded to the native growth hormone receptor and a variant hGH that was complexed to a mutant growth hormone receptor were compared to study the structural basis of human growth hormone (hGH) receptor recognition (Atwell et al. 1997). In another study a three-dimensional structure of VEGF (vascular endothelial growth factor) and its ligand binding site were used for rational design of mutants followed by phage display approach to select improved binding characteristics (Li et al. 2000).

Phage display has been widely used in engineering the binding properties of recombinant antibody fragments (reviewed in Azzazy and Highsmith 2002) as well as other structures. For example, and especially related to the current work, Panni and collaborators (2002) generated a repertoire of SH3 domains by engineering chemical characteristics at the proline-rich ligand (PxxP) binding surface of Abl SH3 domain and by phage display were able to recover specific variant SH3 domains that either bound to

a ligand peptide containing the Abl specific or to a ligand peptide containing a Src specific proline rich sequence (Panni et al. 2002).

8.5. The role of the RT-loop in SH3 ligand selection

As discussed in Section 5.5, studies with peptide-SH3 ligand complexes have revealed that relatively short flanking sequences as well as larger secondary structures in conjunction with the minimal PxxP-motif can improve the specificity and affinity of binding. The use of a separate surface formed by discontinuous amino acid residues in the ligand protein for SH3 recognition via the "tip" of the RT-loop is an arrangement so far described only for the Nef/Hck SH3 interaction. However, the important role of the RT-loop in SH3 ligand selection seen in the Nef/SH3 interactions may as well be more universal and not just a unique property of Nef. It is therefore of interest if other PxxP-containing SH3 ligand proteins could serve as targets for the RRT-SH3 library derived from the Hck SH3 domain. On the other hand it is likely that the invariant Hck SH3 encoded regions in an RRT-SH3 domain could pose some restrictions to efficient ligand binding using the RRT-SH3 strategy based on the Hck SH3 "backbone" and that only certain kind of PxxP motif holding ligands could be targeted.

We have used several other PxxP containing ligands (including, for example, Sos, p85 regulatory subunit of the PI3K, PAK1 and Btk as well as cytoplasmic tails of CD3ε, CD28 and ADAM 15) in the affinity selection experiments and have been able to select RRT-SH3 domains that can bind tightly to each of these PxxP proteins. Moreover, all of these RRT-SH3 domains looked very different when compared to each other as well as to the wild-type Hck SH3 RT-loop sequence EAIHHE (Manuscript in preparation). Examples of the RT-loop sequences selected with other PxxP proteins than with Nef are shown in Table III.

Table III
Examples of the RT-loops retrieved from the selection of the RRT-SH3 library with other PxxP proteins than Nef.

SH3 domain selected by	RT-loop
Sos	RALELG
p85	EAYAPW
PAK1	DPLYVG
Btk	NDRDVG
CD3ε cytoplasmic tail	VARATP
CD28 cytoplasmic tail	VGIYWD
ADAM15 cytoplasmic tail	TGEDRN

Notably, unlike Nef these proteins are not normally high-affinity binding partners of native Hck SH3. It is of interest that these proteins contain both classical consensus sequences as well as proline-rich motifs without a classical consensus sequence that suggest that the optimal PxxP recognition is not required in order to develop high-affinity binding RRT-SH3 domains thus further emphasising the role of the RT-loop in the ligand recognition. In addition, the invariant Hck SH3 encoded region did not interfere with the affinity selection and thus it is also possible to select Hck SH3-derived RRT-SH3 domains even for an atypical ligand. However, using this kind of an Hck

SH3-derived library in targeting atypical SH3 ligands might not be most successful approach to select for the highest specificity or affinity although many of the atypical ligands also may look like conventional ligand consensus sequences (see Section 5.4.6.).

Finding artificial Hck SH3 derived ligands for these proteins suggests that the modification in the RT-loop is sufficient in ligand targeting and that the central "tip" of the RT-loop can play a dominant role in the affinity and specificity of the ligand recognition. Our data argue that individual SH3-domains can have a capacity for mediating highly specific and strong interactions, and do not necessarily need for example, multiprotein complexes to achieve this as has been postulated for instance by Ladbury and Arold (2000). However, the frequency of this kind of SH3 ligand recognition as a general phenomenon (apart from Nef/Hck SH3) in nature remains unknown. SH3 domains are often used in transient multiprotein complexes in cellular signal transduction pathways and extremely high binding affinities there would be deleterious for the signal to be transmitted as they would not be easily reversed. Thus, it may not have been evolutionary beneficial in most instances to develop SH3 domains with very high binding affinities. On the other hand, most of the SH3 binding studies have been or still are performed with short peptide ligands that are not readily comparable to native proteins provided that if the RT-loop is required for the specific ligand recognition, these peptides may lack proper contacts with the RT-loop or other possible specificity determinants on the SH3 molecule.

8.6. Nef selected RRT-SH3 domains can act as Nef inhibitors

Whether or not it is common that natural SH3 domain interactions have high binding affinities, a tempting idea is that these kinds of engineered molecules that have been introduced here could be used as potent inhibitors of natural SH3 interactions. Targeting of cellular SH3 domains could provide a means for example, for regulating or inhibiting the signalling pathways mediated by them. These reagents could serve for example, as tools for studying signalling pathways as such or ultimately be developed for use in gene therapy.

8.6.1. RRT-SH3 domains associate efficiently with Nef in co-transfected cells (**III**)

Association of the Nef selected RRT-SH3 clones with Nef *in vivo* was studied by coexpressing the parental Hck SH3 domain and RRT-SH3 domains (A1, A2, B4, B6 and C1, see Table II) with wild-type or PA-1 variant (Table I) of Nef in 293T cells. Cells were lysed, associated protein complexes captured with glutathione Sepharose 4B beads and detected with anti HA (SH3) and anti-Nef antibodies. All the RRT-SH3 domains studied associated with Nef considerably more efficiently than Hck SH3. The Nef PA-1 mutant failed to associate with the wild-type Nef specific RRT-SH3 domains in cotransfected cells, as already suggested based on the *in vitro* phage display binding data (**II**).

8.6.2. RRT-SH3 domains bind poorly to many normal cellular ligands of Hck SH3 (**III**)

In order to use the targeted RRT-SH3 domains as intracellular inhibitors, they should ideally bind only to their specific target, but not to the natural substrates of their parental SH3 domains. This implies that if the RT-loop indeed is a critical specificity determinant in SH3 ligand selection, modifications in the RT-loop that provide

increased affinity towards the targeted ligand, should concomitantly decrease binding to the natural ligands of the parental SH3 domain (with an unmodified RT-loop).

Myeloid cell lines U937 and K562 were used as sources for endogenous Hck SH3-binding proteins because Hck is naturally expressed particularly in cells of the myeloid lineage. As many of the SH3 ligands are signalling proteins that become tyrosine phosphorylated upon activation, cells were treated with the tyrosine phosphatase inhibitor sodium orthovanadate to accumulate tyrosine phosphorylated proteins. Cell lysates were incubated with recombinant GST or GST-SH3 proteins immobilised on glutathione Sepharose 4B beads and the associated proteins were blotted with anti-phosphotyrosine antibodies to compare profiles of proteins co-precipitating from cell extracts with wild-type Hck SH3 or the RRT-SH3 domains.

Several phosphorylated proteins could be precipitated by the Hck SH3 domain. Some were common to both cell lines whereas others were preferentially expressed in either of them. As hypothesised above, most of these phosphoproteins were either preferentially or exclusively associated with wild-type Hck SH3 but not with the Nef-targeted RRT-SH3 domains. Therefore, most of the cellular target proteins of Hck SH3 seemed to require the natural Hck SH3 RT-loop residues for binding. This result strongly argues that the RT-loop is critically involved also in recognition of the natural ligands of Hck, and supports the idea that the RRT-SH3 approach might be useful for targeting specific SH3/ligand complexes in the cell without interfering with others.

The binding of the PPII helical region of the ligand protein to its cognate SH3 surface does not introduce major changes on the interface nor is it very different from one SH3 domain to another (Larson and Davidson 2000). Despite our encouraging results it is therefore clear that the specificity that can be introduced by the engineered RT-loop residues has its limitations. However, the affinity of the wild-type Hck SH3 to Nef is of an order of a magnitude greater than the affinity of a typical SH3/ligand complex (Lee et al. 1995) and the relative affinities of these RRT-SH3/Nef complexes even stronger (**II**) thus suggesting that the optimal levels of expression can be found to sufficiently inhibit the target molecule without harmful effect via any residual binding to other SH3 ligand proteins in cell.

8.6.3. RRT-SH3s are able to inhibit the association of Nef with PAK2 (**III**)

As shown (**I**), association of Nef with PAK2 is dependent on SH3 binding capacity of Nef. The ability of the Nef-selected RRT-SH3 domains in disrupting the PAK2 association with Nef was tested by transiently transfecting the 293T cells with Nef, PAK2 and RRT-SH3 expression vectors together with the $cdc42^{V12}$ followed by immunoprecipitation with anti-Nef antibodies and subsequent *in vitro* kinase assay. The absence of the SH3 domains resulted in a strong Nef-associated PAK2 signal. Co-expression of the Hck SH3 reduced the signal to some extent while as the selected RRT-SH3 domains were almost completely able to abolish the Nef/PAK2 association.

8.6.4. RRT-SH3s are able to inhibit Nef-induced activation of NFAT (**III**)

Manninen et al. (2000) showed recently that Nef can activate nuclear translocation of the transcription factor NFAT (the nuclear factor of activated T-cells) by synergising with the MAPK-signalling pathway in inducing gene expression via the ARRE2 element (antigen receptor responsive element of the IL-2 gene; Rao et al. 1997) that is an example a composite NFAT/AP-1 (activator protein-1) binding site. The activation of

NFAT-dependent transcription also correlates with the capacity of Nef for SH3 binding (Manninen et al. 2001).

The effect of the Nef-selected RRT-SH3 domains in the Nef-induced NFAT transcriptional activation was studied by cotransfecting Jurkat T-cells with Nef and an ARRE2-driven luciferase reporter plasmid together with RRT-SH3 expression vectors, Hck SH3 or the control vector. A beta-galactosidase expression vector was added as a control of transfection efficiency. Cells were stimulated with phorbol 12-myristate, 13-acetate (PMA) to activate MAPK signalling, lysed and examined for luciferase and beta-galactosidase activities.

Upon PMA stimulation the NFAT activity was increased approximately 30-fold in cells expressing Nef compared to PMA-stimulated cells not expressing Nef. All of the RRT-SH3 domains strongly reduced the NFAT-activation induced by Nef+PMA down to 10% of the control cells. Surprisingly, Hck SH3 was as powerful as the RRT-SH3 domains in inhibiting the Nef-induced NFAT activation although it was considerably less potent in association with Nef in cells as well as in the abolishing the Nef/PAK2 interaction. However, the Nef-independent NFAT activation via T-cell receptor (TCR) ligation by anti-CD3 and anti-CD28 antibodies clearly showed that upon TCR stimulation the RRT-B6 did not interfere with NFAT activation via its natural pathway whereas ectopically expressed Hck SH3 was also a potent inhibitor of this Nef-independent NFAT activation. Thus the NFAT inhibition by the B6 appeared to be specifically due to inhibition of Nef function whereas Hck SH3 inhibited also some other signal transducing factors contributing the ARRE2-mediated gene expression in activated Jurkat cells. This agrees with the results seen earlier (Section 8.6.2) that the RRT-SH3 domains show decreased binding to cellular targets that are strongly associating with the parental Hck SH3.

8.6.5. RRT-SH3s do not interfere with an SH3-independent function of Nef (**III**)

One of the best-characterised functions of Nef in cells is the downregulation of cell surface expression of the CD4 molecule and coupling of it to the endocytic machinery (Aiken et al. 1994). CD4 downregulation was shown to be independent on the SH3 binding capacity of Nef (Saksela et al. 1995). Nevertheless, we were interested to see if the SH3 binding could affect this function of Nef perhaps via a more general inhibition of Nef. It is possible that the binding of the RRT-SH3 domains can also affect the SH3 independent functions in their target proteins by some other means of action. The randomised hexapeptide in the RT-loop might bind to other sequences of the target protein and interfere for example, with the subsequent signalling events that are important for the protein function. It is also possible that the RRT-SH3 domains cause more indirect effects on cellular processes. For example, the presence of high-affinity SH3 domains could cause “overactivation” of some other, non-SH3-mediated interactions because of the saturation of the SH3 recognition sequences in the target molecule (in this case, Nef) and subsequent lack of competition of the binding sites. In addition, blocking the target protein function could be also caused by physical hindrance of the binding site, introduction of a change in the protein conformation, stability or subcellular localisation. These non-competitive means of blocking the effect of the target protein function could be also appended to the inhibitory SH3 domains to make them even more powerful, where applicable.

293T cells were cotransfected with human CD4, green fluorescent protein (GFP) alone or together with Nef, with or without RRT-B6 and studied with flow cytometry. As expected, Nef induced CD4 downregulation that could not be markedly affected by coexpressing the RRT-B6 indicating that the engagement of the SH3 on Nef does not interfere with this function.

8.7. Targeting SIV Nef with the RRT-SH3 approach

8.7.1. Selection of the multivalent phage-display library of RRT-SH3 domains by SIVmac239 Nef (**IV**)

Because no natural high-affinity SH3 ligands for SIV Nef have been found we were interested to use the RRT-SH3 approach to study if the SIVmac239 Nef (SIV Nef) is capable of such an interaction. The multivalent RRT-SH3 library and a monovalent sublibrary (see Materials and methods) were selected against the recombinant GST-SIV Nef using GST only as a negative control. Only two different RT-loop sequences: DGWWG and EGWWG (termed as D and E, respectively) were retrieved from these selections. These sequences did not resemble any natural SH3 RT-loop in the GenBank database or the previously selected sequences by HIV-1 Nef (**II**). Moreover, SIV Nef appeared to select specifically clones containing only five randomised residues in the RT-loop whereas the HIV-1 Nef selected clones that all contained six residues in the RRT region.

Finding the five-residue RRT-sequences from the library originally designed to have six random residues in the RT-loop suggests some kind of a selective advantage of these clones for binding to SIV Nef. The group of the shorter RT-loop clones that were subsequently selected by SIV Nef may have arisen during the original cloning procedure or they may have evolved by a deletion during the course of phage selection and amplification.

8.7.2. RRT-SH3 domains associate efficiently with SIV Nef *in vitro* (**IV**)

To study the association of the SIV Nef with its cognate RRT-SH3 domains also in other expression systems, D and E as well as the control Fyn and Hck SH3 domains (Collette et al. 2000) were expressed as recombinant GST fusion proteins and were coupled to glutathione Sepharose 4B affinity beads. SIV Nef, and for comparison, HIV-1 Nef were ectopically expressed in 293T cells and the lysates were used for co-precipitation studies. The SIV Nef selected clones D and E bound strongly to SIV Nef whereas both the Hck and Fyn SH3 were not able to pull down visible amounts of SIV Nef in this assay. This implies that substituting the DGWWG or EGWWG sequences in the Hck SH3 RT-loop results in a specific high-affinity binding to SIV Nef. The binding of D and E to SIV Nef was at least two orders of magnitude greater when compared to Fyn SH3. As expected, HIV-1 Nef was capable of binding to Hck SH3 but the binding was much weaker compared to the D or E binding to SIV Nef. This suggested the affinities of SIV Nef/D or SIV Nef/E binding to be on a low nanomolar range. These results show that the lack of intracellular SH3 partners that would associate strongly with SIV Nef is not resulting for example, from a non-functional PPII helix-region. In respect it can be noted that these RRT-SH3 domains were, however, not tested in binding to the PxxP mutant of SIV Nef and therefore it cannot be formally excluded that the binding would take place also without an intact PxxP motif in SIV Nef, but based on our studies with HIV-1 Nef we consider this to be unlikely.

8.7.3. RRT-SH3 domains associate efficiently with SIV Nef in cells (**IV**)

Association of the SIV Nef selected RRT-SH3 clones with SIV Nef *in vivo* was studied by coexpressing D, E and for a comparison, previously HIV-1 selected clones A1 and B6 (**II**) together either SIV Nef or HIV-1 Nef vectors in 293T cells. Cells were lysed, the associated protein complexes captured with glutathione Sepharose 4B beads and detected with anti HA (SH3) and anti-myc (Nef) antibodies. D and E were able to co-precipitate SIV Nef whereas the A1 or B6 were not, indicating their specificity to HIV-1 Nef. In addition to the strong binding of the A1 and B6 to HIV-1 Nef, D as well as E were also capable of binding to HIV-1 Nef to some extent showing that their engineered RT-loops also posed some determinants recognised by HIV-1 Nef.

8.7.4. SIV Nef selected RRT-SH3s are able to inhibit Nef-induced activation of NFAT (**IV**)

Induction of NFAT transcriptional activation (Manninen et al. 2000) was shown to be a conserved property among various Nef alleles, including SIV Nef (Manninen et al. 2001). To verify that the association between the D and E clones with SIV Nef (see previous chapter) takes place already in cells as well as to test the D and E clones for their ability to serve as cellular inhibitors, this conserved function of SIV Nef was utilised. For the transcriptional assay Jurkat T-cells were co-transfected as in (**III**), stimulated with PMA and measured for luciferase and β-galactosidase activities. Upon PMA stimulus both the HIV-1 Nef and SIV Nef were able to activate calcium signalling leading to NFAT dependent transcription, the induction being more potent for HIV-1 Nef. SIV Nef selected clones D and E were able to inhibit the SIV Nef-induced NFAT activity more than 75% of the value without inhibitors while as HIV-1 Nef selected clones B6 only slightly inhibited and A1 was virtually without an effect in this assay. In agreement with (**III**) A1 and B6 acted as potent inhibitors of the HIV-1 Nef induced NFAT transcriptional activity and as suggested by the slight association in pull-down experiments, the SIV Nef selected D and E were also capable of inhibiting the HIV-1 Nef induced transcriptional activation to some 50% of the value generated by Nef only.

8.7.5. The role of the SH3 mediated interactions in SIV Nef functions

This study shows that SIV Nef is capable of high affinity interaction with SH3 domains although a natural example of this is lacking. Association with SIV Nef-targeted RRT-SH3 domains took place in cells and resulted in inhibition of NFAT transcriptional activation, thus also suggesting that these engineered SH3 domains could be useful as intracellular inhibitors of SIV Nef. On the other hand, these domains were not strictly specific for SIV Nef but were also capable of inhibiting the HIV-1 Nef function to some extent. Recent studies suggest that SIV Nef is less dependent on SH3 binding in mediating its cellular functions than HIV-1 Nef (Lang et al. 1997; Carl et al. 2000; Swigut et al. 2000). This suggests that HIV-1 and SIV underwent convergent evolution where different molecular strategies were selected with similar outcome at least in the Nef/SH3 binding interface (Picard et al. 2002). However, the strict conservation of the PxxP region among various Nef alleles (including HIV-1, HIV-2 and SIV) suggests a role in controlling the host proteins and that there may be cellular SH3 proteins that are able to bind also SIV Nef. Based on this study the potential SIV Nef associating SH3 domains are likely to be different from those interacting with HIV-1 Nef.

Infection of the natural host with SIV is considered to be nonpathogenic (an exception reported; Ling et al. 2004) and in that sense the biological significance of a possible SIV

Nef/SH3 interaction in controlling the virus-host interactions is obscure. However, a recent report indicated that a cross-species transmission of a naturally occurring SIV between two African non-human primates has the ability to induce AIDS in the recipient species but the outcome of the infection may vary considerably (Apetrei et al. 2004). The possible mechanisms underlying these events were not studied in great detail so it could be envisioned that virus-host communication events could also involve Nef/SH3 dependent interactions.

8.8. Tools for interfering with cellular protein interactions

Intracellular antibodies (intrabodies) represent a new application of antibodies. This concept involves intracellular expression of antibody fragments that are directed to specific subcellular compartments to inactivate the target proteins (reviewed in Kontermann 2004; Stocks 2004). Intrabodies have been successively used as inhibitors of proteins that are for example, involved HIV infection (Rondon and Marasco 1997) and in cancer (reviewed in Hudson and Souriau 2003; Souriau and Hudson 2003), the ultimate goal being used as therapeutic agents. Common problems have been encountered by the use of intrabodies are the instability, incorrect folding and solubility (Cattaneo and Biocca 1999; Wörn et al. 2000; Wörn and Plückthün 2001). However, Lener et al. (2000) demonstrated recently that a panel of single-chain anti-p21 Ras antibodies that were selected from the phage display library by recombinant p21 Ras proteins were able to inactivate Ras when expressed intracellularly although they were not able to interfere with the Ras GTPase activity and the Rac binding activity *in vitro*. These antibodies were aggregating in cells suggesting that they may be in intermediate state of the folding process that allowed sequestration of Ras in intracellular aggregates, diverting it from the natural intracellular location.

Due to the problems encountered with intrabodies the use of inhibitory RRT-SH3 could provide a potential alternative since these domains provide a stable fold and solvent exposed residues at the site of the ligand recognition. Moreover, these residues can be randomised to create specific sequences for every target protein in question. However, it cannot be completely excluded that the aforementioned problems would not in some cases also interfere with the use of intracellular SH3 domains. Obviously, this strategy is useful only for proteins that normally serve as SH3 ligands.

Most probably the applications of inhibitory SH3 domains would first concentrate in basic research, like in studying cellular signalling pathways. For example, it would be of interest to study if expression of the artificial and specific SH3 domains could modulate the signalling pathways regulated by their target proteins like the Sos protein-mediated activation of the p21Ras signalling pathway (McCormick 1993). The signalling is initiated via a growth factor binding to its receptor causing the receptor phosphorylation that leads to subsequent signal transduction effects. The adaptor protein Grb2 is able to bind via its SH2 domain to the phosphorylated receptor and the SH3 of Grb2 in turn binds to a Sos protein and the complex is recruited on the plasma membrane bringing it close to Ras protein that is activated by Sos. The activated Ras mediates the incoming signal further to the nucleus. In aberrant Ras signalling, often seen in cancer, the Ras is in the active state at all times and for example, promotes cells for growth (reviewed in Coleman et al. 2004). Using Sos/SH3 interaction as a model, it can be hypothesised that by engineering artificial SH3 domains that will bind tightly to

Sos, the natural Grb2 function could be circumvented and the recruitment of the natural Grb2/Sos complex to plasma membrane could be blocked.

The use of the inhibitory SH3 domains as gene therapy reagents would require further development. Transportation of these domains inside the cells might be achieved by using various internalisation peptides like HIV-1 TAT or antennapedia that have been successfully used for delivering cargos inside the cells (reviewed in Temsamani and Vidal 2004). As a proof-of concept for gene therapy based on inhibitory SH3 domains as reagents blocking for example the Nef functions could be studied in transgenic mice. As discussed (Hanna et al. 1998), in the mouse model of HIV-1 Hck was found to be important but not essential for the disease progression. It would be of interest to see the effect of these inhibitors expressed as transgenes in the disease development in this model. Use of the inhibitory SH3 domains in inhibiting the Nef functions and as a cure for AIDS encounters the challenges and problems associated with gene therapy, thus, targeting the right gene in the right place at the right time and simultaneously overcome all the adverse effects like host defence related factors, low efficacy or circulating times. Although highly important, Nef is not the only factor causing the development of AIDS and in addition to the inhibitory SH3 domains, other remedies are also needed.

Considering the SH3/ligand interaction, the canonical PxxP motif presumably guides the interaction whereas the variable "tip" of the RT-loop anchors to a binding surface to provide more affinity and specificity for the binding. Elucidation of the structural basis of the RT-loop/ligand interaction could provide tools for molecular modelling of small drug-like molecules that would prevent this interaction and result in disrupting the whole SH3/ligand interaction provided that the RT-loop is the critical component of the high-affinity interaction. These molecular models of small drug-like molecules could further be used in drug development. For instance, it has been shown that intracellular peptide libraries can be produced in mammalian cells for phenotypic screens. Random peptides have been successfully inserted in the loop structures of green fluorescent protein (GFP) followed by a functional screen in yeast (Abedi et al. 1998) and mammalian cells (Peelle et al. 2001). It can be envisaged that the inhibitory molecules of the RT-loop/ligand interaction could be attached as a part of a GFP for demonstrating the individual phenotypes and the possible inhibitory effect.

8.9. Summary

Protein interaction modules mediate many dynamic and ordered protein-protein interactions that are critical for the assembly of multiprotein complexes and signalling pathways that control the cell behaviour. Disruption of these sensitive multiprotein networks may result in pathological effects. During evolution many kinds of different protein interaction modules have emerged. SH3 domains were among the first ones identified and have thereafter been found in large numbers in various types of proteins.

SH3 domains are small (approximately 60 amino acids) modular proteins that mediate inter- and intramolecular protein interactions by recognising short proline-rich stretches that adopt a left-handed polyproline type II (PPII) helical conformation. The PPII helix is also a relatively abundant structure on the surface of various proteins and due to the biochemical characteristics is well suited for mediating protein-protein interactions. The main focus of the work presented here was to characterise the ligand recognition in

SH3-mediated interactions, mainly by using the HIV-1 Nef/Hck SH3 interaction as a model.

The HIV-1 Nef protein has been shown to interfere with cellular SH3-mediated processes. We studied the SH3 binding capacity of Nef as well as the ability to associate with a cellular protein PAK2 by mutagenesis approach. Our results indicated that mutations affecting the residues of the SH3 binding interface abolished binding of Nef to Hck SH3, while the mutations not implicated in SH3 binding were like the wild-type Nef. The SH3-binding surface together with another site containing residues not involved in SH3-binding were required for the PAK2 association. The non-SH3 binding surface was presumed to form a binding site for PAK2 or some other component of this interaction or alternatively, to be involved in a more indirect Nef oligomerisation.

Based on the previous structural and functional data of the specificity of HIV-1 Nef/Hck SH3 interaction we were interested to see if manipulation of the "tip" of the Hck SH3 RT-loop could result in differential binding characteristics towards a variety of PxxP-containing ligands. Libraries of Hck SH3 domains containing engineered RT-loops (RRT-SH3) were expressed on the surface of bacteriophage particles followed by selection with wild-type and R90 mutants (defective in proper interaction with Hck SH3) of HIV-1 Nef. Because of no reported natural high-affinity SH3 domains binding to SIV Nef we also examined the capacity of SIV Nef for the high-affinity SH3-interaction by this approach. All the selections revealed RRT-SH3 clones that bound tightly to their cognate targets indicating that the RT-loop is a critical determinant of these interactions. Some of the RRT-SH3 domains were able to recognise the Nef protein via divergent molecular strategies compared the wild-type Hck SH3.

A panel of RRT-SH3 domains selected by HIV-1 Nef were studied for their ability to function as intracellular inhibitors of this protein. Our data showed that these RRT-SH3 domains could efficiently associate with Nef in cells and inhibit the cellular functions of Nef (disrupting the association of Nef with PAK2 and inhibiting the Nef-induced NFAT transcriptional activation) without interfering with the SH3 independent function of Nef (CD4 receptor downmodulation). Moreover, the RRT-SH3 domains recognised only poorly the expected cellular substrates of Hck SH3 domain thus suggesting a role for the engineered RT-loop in ligand recognition. The SIV Nef selected RRT-SH3 domains could also associate with SIV Nef in cells and inhibit the SIV Nef-induced NFAT transcriptional activation. This suggests a potential role of these RRT-SH3 domains to act as cellular inhibitors of these Nef proteins.

The utility of the RRT-SH3 domains could be assessed in many ways. The applications would most probably concentrate in studying for example, cellular signal transduction pathways or mechanisms of protein-protein interactions in general. Moreover, knowing the molecular basis underlying these interactions could help in developing small-molecular components that would interfere with the correct docking of the RT-loop on the ligand surface. These drug-like molecules could be further developed to functional drugs of the SH3-mediated interactions.

9. ACKNOWLEDGEMENTS

This study was carried out at the Laboratory of Molecular Medicine, Institute of Medical Technology, University of Tampere.

I wish to express my gratitude to my supervisor, Professor Kalle Saksela, for his constant support and encouragement during these years. His optimism and devotion to science have helped me throughout the study. I also wish to thank Professor Olli Silvennoinen for the excellent research facilities at IMT.

I am very grateful to Professor Kari Keinänen and Docent Jari Ylänne for critical reviewing of the manuscript and for their valuable suggestions for improvements. Jim Rowland is acknowledged for revising the language.

I wish to express my sincere thanks to all my co-authors, Professor Mauno Vihinen, Drs Bruce Mayer, Wange Lu, Aki Manninen and Herma Renkema as well as Päivi Huotari and Kari Poikonen, for their important contribution to this study.

My warmest thanks go to all past and present colleagues at the Saksela lab. Aki Manninen and Herma Renkema as well as Katarina Hattula in the early phase of the study are acknowledged for sharing their expertise and friendship. I want to thank Päivi Huotari for being good company both in the lab as well as during the shopping adventures etc. Kati Pulkkinen, Kristina Lehtinen, Tiina Tissari, Marika Vähä-Jaakkola, Tapio Kesti and all the others are thanked for help and numerous happy moments, pieces of good advice for normal life and for creating a nice and inspiring working atmosphere.

Kati Takaluoma and Saara Aittomäki are acknowledged for their help in the lab as well as friendship and encouragement. Special thanks go to Ilkka Junttila for interesting conversations and support. Paula Kosonen is greatly thanked for help during the years.

I am grateful to all the colleagues and other staff from IMT for help whenever needed. Kimmo Savinainen is warmly thanked for realistic view of life and friendship during the years. Heimo Koskinen, Kaarin Forsman and Eila Vanha-aho are also acknowledged for making my everyday life in the lab a lot easier.

My heartfelt thanks go to my parents, Aila and Jalo, to my sisters Ilona and Kirsi and brother Aki for their love, care and encouragement. Jyrki's family, all the friends and relatives are thanked for their interest in my work, support and for great time disrupting the normal working routines.

Finally, my deepest gratitude goes to my dear husband Jyrki for his love, patience and never-ending support and to our daughter Jenni, our pride and joy, for everything.

This study was supported by grants from the Medical Research Fund of the Tampere University Hospital, the Farmos Research and Science Foundation and the Science Fund of the City of Tampere.

Tampere, March 2005

Marita Hiipakka

10. REFERENCES

Abedi M, Caponigro G and Kamb A (1998). Green fluorescent protein as a scaffold for intracellular presentation of peptides. Nucleic Acids Res. 26(2): 623-630.

Adzhubei AA and Sternberg MJE (1993). Left-handed polyproline II helices commonly occur in globular proteins. J. Mol. Biol. 229(2): 472-493.

Aiken C, Konner J, Landau NR, Lenburg ME and Trono D (1994). Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. Cell 76(5): 853-864.

Alexandropoulos K and Baltimore D (1996). Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein, Sin. Genes Dev.10(11): 1341-1355.

Apetrei C, Gormus B, Pandrea I, Metzger M, ten Haaft P, Martin LN, Bohm R, Alvarez X, Koopman G, Murphey-Corb M, Veazey RS, Lackner AA, Baskin G, Heeney J and Marx PA (2004). Direct inoculation of simian immunodeficiency virus from sooty mangabeys in black mangabeys (Lophocebus aterrimus): first evidence of AIDS in a heterologous African species and different pathologic outcomes of experimental infection. J. Virol. 78(21): 11506-11518.

Arold S, Franken P, Strub MP, Hoh F, Benichou S, Benarous R and Dumas C (1997). 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 5(10): 1361-1372.

Arold S, O'Brien R, Franken P, Strub MP, Hoh F, Dumas C and Ladbury JE (1998). RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. Biochemistry 37(42): 14683-14691.

Arold S, Hoh F, Domergue S, Birck C, Delsuc MA, Jullien M and Dumas C (2000). Characterization and molecular basis of the oligomeric structure of HIV- 1 Nef protein. Protein Sci. 9(6): 1137-1148.

Arold, S and Baur AS (2001) Dynamic Nef and Nef dynamics: how structure could explain the complex activities of this small HIV protein. Trends Biochem. Sci. 26(6):356-363.

Atwell S, Ultsch M, De Vos AM and Wells JA (1997). Structural plasticity in a remodeled protein-protein interface. Science 278(5340): 1125-1128.

Azzazy HME and Highsmith J, W. Edward (2002). Phage display technology: clinical applications and recent innovations. Clin. Biochem. 35(6): 425-445.

Bakal CJ and Davies JE (2000). No longer an exclusive club: eukaryotic signalling domains in bacteria. Trends Cell Biol. 10(1): 32-38.

Ball LJ, Jarchau T, Oschkinat H and Walter U (2002). EVH1 domains: structure, function and interactions. FEBS Lett. 513(1): 45-52.

Barnett P, Bottger G, Klein AT, Tabak HF and Distel B (2000). The peroxisomal membrane protein Pex13p shows a novel mode of SH3 interaction. EMBO J. 19(23): 6382-6391.

Bedford MT, Frankel A, Yaffe MB, Clarke S, Leder P and Richard S (2000). Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. J. Biol. Chem. 275(21): 16030-16036.

Berry D, Nash P, Liu S, Pawson T and McGlade C (2002). A high-affinity Arg-X-X-Lys SH3 binding motif confers specificity for the interaction between Gads and SLP-76 in T cell signaling. Curr. Biol. 12(15): 1336-1341.

Borinstein SC, Hyatt MA, Sykes VW, Straub RE, Lipkowitz S, Boulter J and Bogler O (2000). SETA is a multifunctional adapter protein with three SH3 domains that binds Grb2, Cbl, and the novel SB1 proteins. Cell. Signal. 12(11-12): 769-779.

Bottger G, Barnett P, Klein AT, Kragt A, Tabak HF and Distel B (2000). Saccharomyces cerevisiae PTS1 receptor Pex5p interacts with the SH3 domain of the peroxisomal membrane protein Pex13p in an unconventional, non-PXXP-related manner. Mol. Biol. Cell 11(11): 3963-3976.

Brannetti B, Via A, Cestra G, Cesareni G and Citterich MH (2000). SH3-SPOT: An algorithm to predict preferred ligands to different members of the SH3 gene family. J. Mol. Biol. 298(2): 313-328.

Broome MA and Hunter T (1997). The PDGF receptor phosphorylates Tyr 138 in the c-Src SH3 domain in vivo reducing peptide ligand binding. Oncogene 14(1): 17-34.

Buday L, Wunderlich L and Tamas P (2002). The Nck family of adapter proteins: regulators of actin cytoskeleton. Cell. Signal. 14(9): 723-731.

Carl S, Iafrate AJ, Lang SM, Stolte N, Stahl-Hennig C, Matz-Rensing K, Fuchs D, Skowronski J and Kirchhoff F (2000). Simian immunodeficiency virus containing mutations in N-terminal tyrosine residues and in the PxxP motif in Nef replicates efficiently in rhesus macaques. J.Virol. 74(9): 4155-4164.

Cattaneo A and Biocca S (1999). The selection of intracellular antibodies. Trends Biotechnol. 17(3): 115-121.

Cestra G, Castagnoli L, Dente L, Minenkova O, Petrelli A, Migone N, Hoffmuller U, Schneider-Mergener J and Cesareni G (1999). The SH3 domains of endophilin and amphiphysin bind to the proline-rich region of synaptojanin 1 at distinct sites that display an unconventional binding specificity. J. Biol. Chem. 274(45): 32001-32007.

Cheadle C, Ivashchenko Y, South V, Searfoss GH, French S, Howk R, Ricca GA and Jaye M (1994). Identification of a Src SH3 domain binding motif by screening a random phage display library. J. Biol. Chem. 269(39): 24034-24039.

Chellgren BW and Creamer TP (2004). Short sequences of non-proline residues can adopt the polyproline II helical conformation. Biochemistry 43(19): 5864-5869.

Cheng Y and Prusoff WH (1973). Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochem. Pharmacol. 22(23): 3099-3108.

Cicchetti P, Mayer BJ, Thiel G and Baltimore D (1992). Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. Science 257(5071): 803-806.

Cohen GB, Ren R and Baltimore D (1995). Modular binding domains in signal transduction proteins. Cell 80: 237-248.

Coleman ML, Marshall CJ and Olson MF (2004). Ras and Rho GTPases in G1-phase cell-cycle regulation. Nat. Rev. Mol. Cell. Biol. 5(5): 355-366.

Collette Y, Arold S, Picard C, Janvier K, Benichou S, Benarous R, Olive D and Dumas C (2000). 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. 275(6): 4171-4176.

Combs AP, Kapoor TM, Feng S, Chen JK, Daude-Snow LF and Schreiber SL (1996). Protein structure-based combinatorial chemistry: discovery of non-peptide binding elements to Src SH3 domain. J. Am. Chem. Soc. 118: 287-288.

Corbalan-Garcia S, Yang SS, Degenhardt KR and Bar-Sagi D (1996). Identification of the mitogen-activated protein kinase phosphorylation sites on human Sos1 that regulate interaction with Grb2. Mol. Cell. Biol. 16(10): 5674-5682.

Cullen BR (1999). HIV-1 Nef protein: an invitation to a kill. Nat. Med. 5(9): 985-986.

Dong J, Misselwitz R, Welfle H and Westermann P (2000). Expression and purification of dynamin II domains and initial studies on structure and function. Protein Expr. Purif. 20(2): 314-323.

Du Z, Lang SM, Sasseville VG, Lackner AA, Ilyinskii PO, Daniel MD, Jung JU and Desrosiers RC (1995). Identification of a Nef allele that causes lymphocyte activation and acute disease in macaque monkeys. Cell 82(4): 665-674.

Dutta K, Shi H, Cruz-Chu ER, Kami K and Ghose R (2004). Dynamic influences on a high-affinity, high-specificity interaction involving the C-terminal SH3 domain of p67phox. Biochemistry 43(25): 8094-8106.

Fackler OT and Baur AS (2002). Live and let die: Nef functions beyond HIV replication. Immunity 16(4): 493-497.

Farazi TA, Waksman G and Gordon JI (2001). The biology and enzymology of protein N-myristoylation. J. Biol. Chem. 276(43): 39501-39504.

Feller SM, Wecklein H, Lewitzky M, Kibler E and Raabe T (2002). SH3 domain-mediated binding of the Drk protein to Dos is an important step in signaling of Drosophila receptor tyrosine kinases. Mech. Dev. 116(1-2): 129-139.

Feng S, Chen JK, Yu H, Simon JA and Schreiber SL (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. Science 266(5188): 1241-1247.

Feng S, Kasahara C, Rickles RJ and Schreiber SL (1995). Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. Proc. Natl. Acad. Sci. U.S.A. 92(26): 12408-12415.

Feng S, Kapoor TM, Shirai F, Combs AP and Schreiber SL (1996). Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis. Chem. Biol. 3(8): 661-670.

Feng S and Schreiber SL (1997). Enantiomeric binding elements interacting at the same site of an SH3 protein receptor. J. Am. Chem. Soc. 119: 10873-10874.

Ferguson MR, Fan X, Mukherjee M, Luo J, Khan R, Ferreon JC, Hilser VJ, Shope RE and Fox RO (2004). Directed discovery of bivalent peptide ligands to an SH3 domain. Protein Sci. 13(3): 626-632.

Fernandez-Ballester G, Blanes-Mira C and Serrano L (2004). The tryptophan switch: changing ligand-binding specificity from type I to type II in SH3 domains. J. Mol. Biol. 335(2): 619-629.

Finan P, Shimizu Y, Gout I, Hsuan J, Truong O, Butcher C, Bennett P, Waterfield M and Kellie S (1994). An SH3 domain and proline-rich sequence mediate an interaction between two components of the phagocyte NADPH oxidase complex. J. Biol. Chem. 269(19): 13752-13755.

Fischer EH (1999). Cell signaling by protein tyrosine phosphorylation. Adv. Enzyme Regul. 39: 359-369.

Freund C, Dotsch V, Nishizawa K, Reinherz EL and Wagner G (1999). The GYF domain is a novel structural fold that is involved in lymphoid signaling through proline-rich sequences. Nat. Struct. Biol. 6(7): 656-660.

Geyer M, Fackler OT and Peterlin BM (2001). Structure-function relationships in HIV-1 Nef. EMBO Rep. 2(7): 580-585.

Ghose R, Shekhtman A, Goger MJ, Ji H and Cowburn D (2001). A novel, specific interaction involving the Csk SH3 domain and its natural ligand. Nat. Struct. Biol. 8(11): 998-1004.

Gorina S and Pavletich NP (1996). Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. Science 274(5289): 1001-1005.

Gout I, Dhand R, Hiles ID, Fry MJ, Panayotou G, Das P, Truong O, Totty NF, Hsuan J, Booker GW and et al. (1993). The GTPase dynamin binds to and is activated by a subset of SH3 domains. Cell 75(1): 25-36.

Grabs D, Slepnev VI, Songyang Z, David C, Lynch M, Cantley LC and De Camilli P (1997). The SH3 domain of amphiphysin binds the proline-rich domain of dynamin at a single site that defines a new SH3 binding consensus sequence. J. Biol. Chem. 272(20): 13419-13425.

Greenway AL, Dutartre H, Allen K, McPhee DA, Olive D and Collette Y (1999). Simian immunodeficiency virus and human immunodeficiency virus type 1 Nef proteins show distinct patterns and mechanisms of Src kinase activation. J.Virol. 73(7): 6152-6158.

Greenway AL, Holloway G, McPhee DA, Ellis P, Cornall A and Lidman M (2003). HIV-1 Nef control of cell signalling molecules: multiple strategies to promote virus replication. J. Biosci. 28(3): 323-335.

Gregorieff A, Cloutier JF and Veillette A (1998). Sequence requirements for association of protein-tyrosine phosphatase PEP with the Src homology 3 domain of inhibitory tyrosine protein kinase p50(csk). J. Biol. Chem. 273(21): 13217-13222.

Groemping Y, Lapouge K, Smerdon SJ and Rittinger K (2003). Molecular basis of phosphorylation-induced activation of the NADPH oxidase. Cell 113(3): 343-355.

Grzesiek S, Bax A, Clore GM, Gronenborn AM, Hu JS, Kaufman J, Palmer I, Stahl SJ and Wingfield PT (1996). 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. 3(4): 340-345.

Grzesiek S, Bax A, Hu JS, Kaufman J, Palmer I, Stahl SJ, Tjandra N and Wingfield PT (1997). Refined solution structure and backbone dynamics of HIV-1 Nef. Protein Sci. 6(6): 1248-1263.

Guijarro JI, Sunde M, Jones JA, Campbell ID and Dobson CM (1998). Amyloid fibril formation by an SH3 domain. Proc. Natl. Acad. Sci. U.S.A. 95(8): 4224-4228.

Hanna Z, Kay DG, Rebai N, Guimond A, Jothy S and Jolicoeur P (1998). Nef harbors a major determinant of pathogenicity for an AIDS- like disease induced by HIV-1 in transgenic mice. Cell 95(2): 163-175.

Hanna Z, Weng X, Kay DG, Poudrier J, Lowell C and Jolicoeur P (2001). The pathogenicity of human immunodeficiency virus (HIV) type 1 Nef in CD4C/HIV transgenic mice is abolished by mutation of its SH3-binding domain, and disease development is delayed in the absence of Hck. J. Virol. 75(19): 9378-9392.

Harkiolaki M, Lewitzky M, Gilbert RJ, Jones EY, Bourette RP, Mouchiroud G, Sondermann H, Moarefi I and Feller SM (2003). Structural basis for SH3 domain-mediated high-affinity binding between Mona/Gads and SLP-76. EMBO J. 22(11): 2571-2582.

Hing H, Xiao J, Harden N, Lim L and Zipursky SL (1999). Pak functions downstream of Dock to regulate photoreceptor axon guidance in Drosophila. Cell 97(7): 853-863.

Holmberg CI, Tran SEF, Eriksson JE and Sistonen L (2002). Multisite phosphorylation provides sophisticated regulation of transcription factors. Trends Biochem. Sci. 27(12): 619-627.

Horita DA, Baldisseri DM, Zhang W, Altieri AS, Smithgall TE, Gmeiner WH and Byrd RA (1998). Solution structure of the human Hck SH3 domain and identification of its ligand binding site. J. Mol. Biol. 278(1): 253-265.

Houtman JC, Higashimoto Y, Dimasi N, Cho S, Yamaguchi H, Bowden B, Regan C, Malchiodi EL, Mariuzza R, Schuck P, Appella E and Samelson LE (2004). Binding specificity of multiprotein signaling complexes is determined by both cooperative interactions and affinity preferences. Biochemistry 43(14): 4170-4178.

Hudson PJ and Souriau C (2003). Engineered antibodies. Nat. Med. 9(1): 129-134.

Hunter T (1997). Oncoprotein networks. Cell 88(3): 333-346.

Inglis SR, Stojkoski C, Branson KM, Cawthray JF, Fritz D, Wiadrowski E, Pyke SM and Booker GW (2004). Identification and specificity studies of small-molecule ligands for SH3 protein domains. J. Med. Chem. 47(22): 5405-5417.

Innocenti M, Tenca P, Frittoli E, Faretta M, Tocchetti A, Di Fiore PP and Scita G (2002). Mechanisms through which Sos-1 coordinates the activation of Ras and Rac. J. Cell. Biol. 156(1): 125-136.

Jacobsson K and Frykberg L (1996). Phage display shot-gun cloning of ligand-binding domains of prokaryotic receptors approaches 100% correct clones. Biotechniques 20(6): 1070-1081.

Jardetzky TS, Brown JH, Gorga JC, Stern LJ, Urban RG, Strominger JL and Wiley DC (1996). Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. Proc. Natl. Acad. Sci. U.S.A. 93(2): 734-738.

Kami K, Takeya R, Sumimoto H and Kohda D (2002). Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. EMBO J. 21(16): 4268-4276.

Kaneko T, Kumasaka T, Ganbe T, Sato T, Miyazawa K, Kitamura N and Tanaka N (2003). Structural insight into modest binding of a non-PXXP ligand to the signal transducing adaptor molecule-2 Src homology 3 domain. J. Biol. Chem. 278(48): 48162-48168.

Kang H, Freund C, Duke-Cohan JS, Musacchio A, Wagner G and Rudd CE (2000). SH3 domain recognition of a proline-independent tyrosine-based RKxxYxxY motif in immune cell adaptor SKAP55. EMBO J. 19(12): 2889-2899.

Kaplan KB, Bibbins KB, Swedlow JR, Arnaud M, Morgan DO and Varmus HE (1994). Association of the amino-terminal half of c-Src with focal adhesions alters their properties and is regulated by phosphorylation of tyrosine 527. EMBO J. 13(20): 4745-4756.

Kaplan KB, Swedlow JR, Morgan DO and Varmus HE (1995). c-Src enhances the spreading of src-/- fibroblasts on fibronectin by a kinase-independent mechanism. Genes Dev.9(12): 1505-1517.

Karatan E, Merguerian M, Han Z, Scholle MD, Koide S and Kay BK (2004). Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain. Chem. Biol. 11(6): 835-844.

Kardinal C, Konkol B, Schulz A, Posern G, Lin H, Adermann K, Eulitz M, Estrov Z, Talpaz M, Arlinghaus RB and Feller SM (2000). Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients. FASEB J. 14(11): 1529-1538.

Kato JY, Takeya T, Grandori C, Iba H, Levy JB and Hanafusa H (1986). Amino acid substitutions sufficient to convert the nontransforming p60c-src protein to a transforming protein. Mol. Cell. Biol. 6(12): 4155-4160.

Kato M, Miyazawa K and Kitamura N (2000). A deubiquitinating enzyme UBPY interacts with the Src homology 3 domain of Hrs-binding protein via a novel binding motif PX(V/I)(D/N)RXXKP. J. Biol. Chem. 275(48): 37481-37487.

Kay BK, Williamson MP and Sudol M (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J. 14(2): 231-241.

Kelly MA, Chellgren BW, Rucker AL, Troutman JM, Fried MG, Miller AF and Creamer TP (2001). Host-guest study of left-handed polyproline II helix formation. Biochemistry 40(48): 14376-14383.

Khan IH, Sawai ET, Antonio E, Weber CJ, Mandell CP, Montbriand P and Luciw PA (1998). Role of the SH3-ligand domain of simian immunodeficiency virus Nef in interaction with Nef-associated kinase and simian AIDS in rhesus macaques. J. Virol. 72(7): 5820-5830.

Kishan KV, Scita G, Wong WT, Di Fiore PP and Newcomer ME (1997). The SH3 domain of Eps8 exists as a novel intertwined dimer. Nat. Struct. Biol. 4(9): 739-743.

Kishan KV, Newcomer ME, Rhodes TH and Guilliot SD (2001). Effect of pH and salt bridges on structural assembly: molecular structures of the monomer and intertwined dimer of the Eps8 SH3 domain. Protein Sci. 10(5): 1046-1055.

Knudsen BS, Zheng J, Feller SM, Mayer JP, Burrell SK, Cowburn D and Hanafusa H (1995). Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins. EMBO J. 14(10): 2191-2198.

Kontermann RE (2004). Intrabodies as therapeutic agents. Methods 34(2): 163-170.

Kurakin AV, Wu S and Bredesen DE (2003). Atypical recognition consensus of CIN85/SETA/Ruk SH3 domains revealed by target-assisted iterative screening. J. Biol. Chem. 278(36): 34102-34109.

Ladbury JE and Arold S (2000). Searching for specificity in SH domains. Chem. Biol. 7(1): R3-R8.

Landgraf C, Panni S, Montecchi-Palazzi L, Castagnoli L, Schneider-Mergener J, Volkmer-Engert R and Cesareni G (2004). Protein interaction networks by proteome peptide scanning. PLoS Biol. 2(1): E14.

Lang SM, Iafrate AJ, Stahl-Hennig C, Kuhn EM, Nisslein T, Kaup FJ, Haupt M, Hunsmann G, Skowronski J and Kirchhoff F (1997). Association of simian immunodeficiency virus Nef with cellular serine/threonine kinases is dispensable for the development of AIDS in rhesus macaques. Nat. Med. 3(8): 860-865.

Larson SM and Davidson AR (2000). The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. Protein Sci. 9(11): 2170-2180.

Lee CH, Leung B, Lemmon MA, Zheng J, Cowburn D, Kuriyan J and Saksela K (1995). 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. 14(20): 5006-5015.

Lee CH, Saksela K, Mirza UA, Chait BT and Kuriyan J (1996). Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell 85(6): 931-942.

Lehto VP, Wasenius VM, Salven P and Saraste M (1988). Transforming and membrane proteins. Nature 334(6181): 388.

Lener M, Horn IR, Cardinale A, Messina S, Nielsen UB, Rybak SM, Hoogenboom HR, Cattaneo A and Biocca S (2000). Diverting a protein from its cellular location by intracellular antibodies. The case of p21Ras. Eur. J. Biochem. 267(4): 1196-1205.

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP and Bork P (2004). SMART 4.0: towards genomic data integration. Nucleic Acids Res. 32 Database issue: D142-D144.

Lewitzky M, Kardinal C, Gehring NH, Schmidt EK, Konkol B, Eulitz M, Birchmeier W, Schaeper U and Feller SM (2001). The C-terminal SH3 domain of the adapter protein Grb2 binds with high affinity to sequences in Gab1 and SLP-76 which lack the SH3-typical P-x-x-P core motif. Oncogene 20(9): 1052-1062.

Lewitzky M, Harkiolaki M, Domart M, Jones EY and Feller SM (2004). 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. 279(27): 28724-28732.

Li X, Multon MC, Henin Y, Schweighoffer F, Venot C, Josef J, Zhou C, LaVecchio J, Stuckert P, Raab M, Mhashilkar A, Tocque B and Marasco WA (2000). Grb3-3 is up-regulated in HIV-1-infected T-cells and can potentiate cell activation through NFATc. J. Biol. Chem. 275(40): 30925-30933.

Lim WA, Richards FM and Fox RO (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. Nature 372(6504): 375-379.

Ling B, Apetrei C, Pandrea I, Veazey RS, Lackner AA, Gormus B and Marx PA (2004). Classic AIDS in a sooty mangabey after an 18-year natural infection. J. Virol. 78(16): 8902-8908.

Liu LX, Heveker N, Fackler OT, Arold S, Le Gall S, Janvier K, Peterlin BM, Dumas C, Schwartz O, Benichou S and Benarous R (2000). Mutation of a conserved residue (D123) required for oligomerization of human immunodeficiency virus type 1 Nef protein abolishes interaction with human thioesterase and results in impairment of Nef biological functions. J. Virol. 74(11): 5310-5319.

Liu Q, Berry D, Nash P, Pawson T, McGlade CJ and Li SS (2003). 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 11(2): 471-481.

Mahoney NM, Rozwarski DA, Fedorov E, Fedorov AA and Almo SC (1999). Profilin binds proline-rich ligands in two distinct amide backbone orientations. Nat. Struct. Biol. 6(7): 666-671.

Manninen A, Renkema GH and Saksela K (2000). Synergistic activation of NFAT by HIV-1 Nef and the Ras/MAPK pathway. J. Biol. Chem. 275(22): 16513-16517.

Manninen A, Huotari P, Hiipakka M, Renkema GH and Saksela K (2001). Activation of NFAT dependent gene expression by Nef: conservation among different Nef alleles, dependence on SH3-binding and membrane association, and cooperation with protein kinase C-theta. J. Virol. 75(6): 3034-3037.

Manning G, Whyte DB, Martinez R, Hunter T and Sudarsanam S (2002). The protein kinase complement of the human genome. Science 298(5600): 1912-1934.

Manser E, Loo TH, Koh CG, Zhao ZS, Chen XQ, Tan L, Tan I, Leung T and Lim L (1998). PAK kinases are directly coupled to the PIX family of nucleotide exchange factors. Mol. Cell 1(2): 183-192.

Mayer BJ, Hamaguchi M and Hanafusa H (1988). A novel viral oncogene with structural similarity to phospholipase C. Nature 332(6161): 272-275.

Mayer BJ (2001). SH3 domains: complexity in moderation. J. Cell. Sci. 114(Pt 7): 1253-1263.

McCormick F (1993). Signal transduction. How receptors turn Ras on. Nature 363(6424): 15-16.

McGee AW, Dakoji SR, Olsen O, Bredt DS, Lim WA and Prehoda KE (2001). Structure of the SH3-guanylate kinase module from PSD-95 suggests a mechanism for regulated assembly of MAGUK scaffolding proteins. Mol. Cell 8(6): 1291-301.

McPherson PS (1999). Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis. Cell. Signal. 11(4): 229-238.

Mizushima S and Nagata S (1990). pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. 18(17): 5322.

Moarefi I, LaFevre-Bernt M, Sicheri F, Huse M, Lee CH, Kuriyan J and Miller WT (1997). Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement. Nature 385(6617): 650-653.

Mongiovi AM, Romano PR, Panni S, Mendoza M, Wong WT, Musacchio A, Cesareni G and Di Fiore PP (1999). A novel peptide-SH3 interaction. EMBO J. 18(19): 5300-5309.

Morrogh LM, Hinshelwood S, Costello P, Cory GO and Kinnon C (1999). The SH3 domain of Bruton's tyrosine kinase displays altered ligand binding properties when auto-phosphorylated in vitro. Eur. J. Immunol. 29(7): 2269-2279.

Musacchio A, Noble M, Pauptit R, Wierenga R and Saraste M (1992). Crystal structure of a Src-homology 3 (SH3) domain. Nature 359(6398): 851-855.

Musacchio A, Saraste M and Wilmanns M (1994). High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. Nat. Struct. Biol. 1(8): 546-551.

Musacchio A (2002). How SH3 domains recognize proline. Adv. Protein Chem. 61: 211-268.

Nguyen JT, Turck CW, Cohen FE, Zuckermann RN and Lim WA (1998). Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. Science 282(5396): 2088-2092.

Nguyen JT, Porter M, Amoui M, Miller WT, Zuckermann RN and Lim WA (2000). Improving SH3 domain ligand selectivity using a non-natural scaffold. Chem. Biol. 7(7): 463-473.

Nishida M, Nagata K, Hachimori Y, Horiuchi M, Ogura K, Mandiyan V, Schlessinger J and Inagaki F (2001). Novel recognition mode between Vav and Grb2 SH3 domains. EMBO J. 20(12): 2995-3007.

Nishizawa K, Freund C, Li J, Wagner G and Reinherz EL (1998). Identification of a proline-binding motif regulating CD2-triggered T lymphocyte activation. Proc. Natl. Acad. Sci. U.S.A. 95(25): 14897-14902.

Noble ME, Musacchio A, Saraste M, Courtneidge SA and Wierenga RK (1993). 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. 12(7): 2617-2624.

Obenauer JC, Cantley LC and Yaffe MB (2003). Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res. 31(13): 3635-3641.

Ogura K, Nagata K, Horiuchi M, Ebisui E, Hasuda T, Yuzawa S, Nishida M, Hatanaka H and Inagaki F (2002). Solution structure of N-terminal SH3 domain of Vav and the recognition site for Grb2 C-terminal SH3 domain. J. Biomol. NMR 22(1): 37-46.

Oneyama C, Nakano H and Sharma SV (2002). UCS15A, a novel small molecule, SH3 domain-mediated protein-protein interaction blocking drug. Oncogene 21(13): 2037-2050.

Oneyama C, Agatsuma T, Kanda Y, Nakano H, Sharma SV, Nakano S, Narazaki F and Tatsuta K (2003). Synthetic inhibitors of proline-rich ligand-mediated protein-protein interaction: potent analogs of UCS15A. Chem. Biol. 10(5): 443-451.

Otsu M, Hiles I, Gout I, Fry MJ, Ruiz-Larrea F, Panayotou G, Thompson A, Dhand R, Hsuan J and Totty N (1991). Characterization of two 85 kd proteins that associate with receptor tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase. Cell 65(1): 91-104.

Panni S, Dente L and Cesareni G (2002). In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. J. Biol. Chem. 277(24): 21666-21674.

Pappu RV and Rose GD (2002). A simple model for polyproline II structure in unfolded states of alanine-based peptides. Protein Sci. 11(10): 2437-2455.

Park H, Wahl MI, Afar DE, Turck CW, Rawlings DJ, Tam C, Scharenberg AM, Kinet JP and Witte ON (1996). Regulation of Btk function by a major autophosphorylation site within the SH3 domain. Immunity 4(5): 515-525.

Pawson T and Nash P (2000). Protein-protein interactions define specificity in signal transduction. Genes Dev.14(9): 1027-1047.

Pawson T, Raina M and Nash P (2002). Interaction domains: from simple binding events to complex cellular behavior. FEBS Lett. 513: 2-10.

Pawson T and Nash P (2003). Assembly of cell regulatory systems through protein interaction domains. Science 300(5618): 445-452.

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

Peelle B, Lorens J, Li W, Bogenberger J, Payan DG and Anderson DC (2001). Intracellular protein scaffold-mediated display of random peptide libraries for phenotypic screens in mammalian cells. Chem. Biol. 8(5): 521-534.

Pellicena P and Miller WT (2001). Processive phosphorylation of p130Cas by Src depends on SH3-polyproline interactions. J. Biol. Chem. 276(30): 28190-28196.

Picard C, Greenway A, Holloway G, Olive D and Collette Y (2002). Interaction with simian Hck tyrosine kinase reveals convergent evolution of the Nef protein from simian and human immunodeficiency viruses despite differential molecular surface usage. Virology 295(2): 320-327.

Pires JR, Hong X, Brockmann C, Volkmer-Engert R, Schneider-Mergener J, Oschkinat H and Erdmann R (2003). The ScPex13p SH3 domain exposes two distinct binding sites for Pex5p and Pex14p. J. Mol. Biol. 326(5): 1427-1435.

Pisabarro MT and Serrano L (1996). Rational design of specific high-affinity peptide ligands for the Abl-SH3 domain. Biochemistry 35(33): 10634-10640.

Pisabarro MT, Serrano L and Wilmanns M (1998). Crystal structure of the Abl-SH3 domain complexed with a designed high-affinity peptide ligand: implications for SH3-ligand interactions. J. Mol. Biol. 281(3): 513-521.

Pleiman CM, Hertz WM and Cambier JC (1994). Activation of phosphatidylinositol-3' kinase by Src-family kinase SH3 binding to the p85 subunit. Science 263(5153): 1609-1612.

Ponting CP and Russell RR (2002). The natural history of protein domains. Annu. Rev. Biophys. Biomol. Struct. 31: 45-71.

Pornillos O, Alam SL, Davis DR and Sundquist WI (2002). Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. Nat. Struct. Biol. 9(11): 812-817.

Posern G, Zheng J, Knudsen BS, Kardinal C, Muller KB, Voss J, Shishido T, Cowburn D, Cheng G, Wang B, Kruh GD, Burrell SK, Jacobson CA, Lenz DM, Zamborelli TJ, Adermann K, Hanafusa H and Feller SM (1998). Development of highly selective SH3 binding peptides for

Crk and CRKL which disrupt Crk-complexes with DOCK180, SoS and C3G. Oncogene 16(15): 1903-1912.

Rao A, Luo C and Hogan PG (1997). Transcription factors of the NFAT family: regulation and function. Annu. Rev. Immunol. 15: 707-747.

Ren R, Mayer BJ, Cicchetti P and Baltimore D (1993). Identification of a ten-amino acid proline-rich SH3 binding site. Science 259(5098): 1157-1161.

Ren R, Ye ZS and Baltimore D (1994). Abl protein-tyrosine kinase selects the Crk adapter as a substrate using SH3-binding sites. Genes Dev. 8(7): 783-795.

Renkema GH, Manninen A, Mann DA, Harris M and Saksela K (1999). Identification of the Nef-associated kinase as p21-activated kinase 2. Curr. Biol. 9(23): 1407-1410.

Renkema GH and Saksela K (2000). Interactions of HIV-1 Nef with cellular signal transducing proteins. Front. Biosci. 5: D268-D283.

Rickles RJ, Botfield MC, Weng Z, Taylor JA, Green OM, Brugge JS and Zoller MJ (1994). Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. EMBO J. 13(23): 5598-5604.

Rickles RJ, Botfield MC, Zhou XM, Henry PA, Brugge JS and Zoller MJ (1995). Phage display selection of ligand residues important for Src homology 3 domain binding specificity. Proc. Natl. Acad. Sci. U.S.A. 92(24): 10909-10913.

Rondon IJ and Marasco WA (1997). Intracellular antibodies (intrabodies) for gene therapy of infectious diseases. Annu. Rev. Microbiol. 51: 257-283.

Rozakis-Adcock M, van der Geer P, Mbamalu G and Pawson T (1995). MAP kinase phosphorylation of mSos1 promotes dissociation of mSos1-Shc and mSos1-EGF receptor complexes. Oncogene 11(7): 1417-1426.

Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR, Hariharan IK, Fortini ME, Li PW, Apweiler R, Fleischmann W, Cherry JM, Henikoff S, Skupski MP, Misra S, Ashburner M, Birney E, Boguski MS, Brody T, Brokstein P, Celniker SE, Chervitz SA, Coates D, Cravchik A, Gabrielian A, Galle RF, Gelbart WM, George RA, Goldstein LS, Gong F, Guan P, Harris NL, Hay BA, Hoskins RA, Li J, Li Z, Hynes RO, Jones SJ, Kuehl PM, Lemaitre B, Littleton JT, Morrison DK, Mungall C, O'Farrell PH, Pickeral OK, Shue C, Vosshall LB, Zhang J, Zhao Q, Zheng XH, Zhong F, Zhong W, Gibbs R, Venter JC, Adams MD and Lewis S (2000). Comparative genomics of the eukaryotes. Science 287(5461): 2204-2215.

Saksela K, Cheng G and Baltimore D (1995). Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4. EMBO J. 14(3): 484-491.

Sauro HM and Kholodenko BN (2004). Quantitative analysis of signaling networks. Prog. Biophys. Mol. Biol. 86(1): 5-43.

Sawai ET, Baur A, Struble H, Peterlin BM, Levy JA and Cheng-Mayer C (1994). Human immunodeficiency virus type 1 Nef associates with a cellular serine kinase in T lymphocytes. Proc. Natl. Acad. Sci. U.S.A. 91(4): 1539-1543.

Sawai ET, Baur AS, Peterlin BM, Levy JA and Cheng-Mayer C (1995). A conserved domain and membrane targeting of Nef from HIV and SIV are required for association with a cellular serine kinase activity. J. Biol. Chem. 270(25): 15307-15314.

Schumacher C, Knudsen BS, Ohuchi T, Di Fiore PP, Glassman RH and Hanafusa H (1995). The SH3 domain of Crk binds specifically to a conserved proline-rich motif in Eps15 and Eps15R. J. Biol. Chem. 270(25): 15341-15347.

Scita G, Nordstrom J, Carbone R, Tenca P, Giardina G, Gutkind S, Bjarnegard M, Betsholtz C and Di Fiore PP (1999). EPS8 and E3B1 transduce signals from Ras to Rac. Nature 401(6750): 290-293.

Shugars DC, Smith MS, Glueck DH, Nantermet PV, Seillier-Moiseiwitsch F and Swanstrom R (1993). Analysis of human immunodeficiency virus type 1 nef gene sequences present in vivo. J. Virol. 67(8): 4639-4650.

Simard MC, Chrobak P, Kay DG, Hanna Z, Jothy S and Jolicoeur P (2002). Expression of simian immunodeficiency virus Nef in immune cells of transgenic mice leads to a severe AIDS-like disease. J. Virol. 76(8): 3981-3995.

Smotrys JE and Linder ME (2004). Palmitoylation of intracellular signaling proteins: regulation and function. Annu. Rev. Biochem. 73(1): 559-587.

Souriau C and Hudson PJ (2003). Recombinant antibodies for cancer diagnosis and therapy. Expert Opin. Biol. Ther. 3(2): 305-318.

Sparks AB, Quilliam LA, Thorn JM, Der CJ and Kay BK (1994). Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. J. Biol. Chem. 269(39): 23853-23856.

Stahl ML, Ferenz CR, Kelleher KL, Kriz RW and Knopf JL (1988). Sequence similarity of phospholipase C with the non-catalytic region of src. Nature 332(6161): 269-272.

Stapley BJ and Creamer TP (1999). A survey of left-handed polyproline II helices. Protein Sci. 8(3): 587-595.

Stocks MR (2004). Intrabodies: production and promise. Drug Discover. Today 9(22): 960-966.

Sudol, M (1998). From Src homology domains to other signaling modules: proposal of the 'protein recognition code'. Oncogene 17(11 Reviews): 1469-1474.

Sudol M, Sliwa K and Russo T (2001). Functions of WW domains in the nucleus. FEBS Lett. 490(3): 190-195.

Swigut T, Iafrate AJ, Muench J, Kirchhoff F and Skowronski J (2000). Simian and human immunodeficiency virus Nef proteins use different surfaces to downregulate class I major histocompatibility complex antigen expression. J. Virol. 74(12): 5691-5701.

Tavares GA, Panepucci EH and Brunger AT (2001). Structural characterization of the intramolecular interaction between the SH3 and guanylate kinase domains of PSD-95. Mol. Cell 8(6): 1313-1325.

Temsamani J and Vidal P (2004). The use of cell-penetrating peptides for drug delivery. Drug Discover. Today 9(23): 1012-1019.

Tong AH, Drees B, Nardelli G, Bader GD, Brannetti B, Castagnoli L, Evangelista M, Ferracuti S, Nelson B, Paoluzi S, Quondam M, Zucconi A, Hogue CW, Fields S, Boone C and Cesareni G (2002). A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science 295(5553): 321-324.

Touw IP, De Koning JP, Ward AC and Hermans MHA (2000). Signaling mechanisms of cytokine receptors and their perturbances in disease. Mol. Cell. Endocrinol. 160(1-2): 1-9.

Weng Z, Rickles RJ, Feng S, Richard S, Shaw AS, Schreiber SL and Brugge JS (1995). Structure-function analysis of SH3 domains: SH3 binding specificity altered by single amino acid substitutions. Mol. Cell. Biol. 15(10): 5627-5634.

Whisstock JC and Lesk AM (1999). SH3 domains in prokaryotes. Trends. Biochem. Sci. 24(4): 132-133.

Viguera AR, Arrondo JL, Musacchio A, Saraste M and Serrano L (1994). Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. Biochemistry 33(36): 10925-10933.

Williamson MP (1994). The structure and function of proline-rich regions in proteins. Biochem. J. 297(Pt 2): 249-260.

Wiskerchen M and Cheng-Mayer C (1996). HIV-1 Nef association with cellular serine kinase correlates with enhanced virion infectivity and efficient proviral DNA synthesis. Virology 224(1): 292-301.

Wörn A, Auf der Maur A, Escher D, Honegger A, Barberis A and Plückthün A (2000). Correlation between in vitro stability and in vivo performance of anti-GCN4 intrabodies as cytoplasmic inhibitors. J. Biol. Chem. 275(4): 2795-2803.

Wörn A and Plückthün A (2001). Stability engineering of antibody single-chain Fv fragments. J. Mol. Biol. 305(5): 989-1010.

Wu X, Knudsen B, Feller SM, Zheng J, Sali A, Cowburn D, Hanafusa H and Kuriyan J (1995). Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. Structure 3(2): 215-226.

Wunderlich L, Goher A, Farago A, Downward J and Buday L (1999). Requirement of multiple SH3 domains of Nck for ligand binding. Cell. Signal. 11(4): 253-262.

Xu Q, Zheng J, Xu R, Barany G and Cowburn D (1999). Flexibility of interdomain contacts revealed by topological isomers of bivalent consolidated ligands to the dual Src homology domain SH(32) of abelson. Biochemistry 38(12): 3491-3497.

Xu W, Harrison SC and Eck MJ (1997). Three-dimensional structure of the tyrosine kinase c-Src. Nature 385(6617): 595-602.

Yaffe MB, Leparc GG, Lai J, Obata T, Volinia S and Cantley LC (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. Nat. Biotechnol. 19(4): 348-353.

Ye H, Choi HJ, Poe J and Smithgall TE (2004). Oligomerization is required for HIV-1 Nef-induced activation of the Src family protein-tyrosine kinase, Hck. Biochemistry 43(50): 15775-15784.

Yu H, Rosen MK, Shin TB, Seidel-Dugan C, Brugge JS and Schreiber SL (1992). Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science 258(5088): 1665-1668.

Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW and Schreiber SL (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76(5): 933-945.

Yuzawa S, Ogura K, Horiuchi M, Suzuki NN, Fujioka Y, Kataoka M, Sumimoto H and Inagaki F (2004). Solution Structure of the Tandem Src Homology 3 Domains of p47phox in an Autoinhibited Form. J. Biol. Chem. 279(28): 29752-29760.

Zarrinpar A and Lim WA (2000). Converging on proline: the mechanism of WW domain peptide recognition. Nat. Struct. Biol. 7(8): 611-613.

Zarrinpar A, Bhattacharyya RP and Lim WA (2003a). The structure and function of proline recognition domains. Sci. STKE 2003(179): RE8.

Zarrinpar A, Park SH and Lim WA (2003b). Optimization of specificity in a cellular protein interaction network by negative selection. Nature 426(6967): 676-680.

Zhao ZS, Manser E and Lim L (2000). Interaction between PAK and Nck: a template for Nck targets and role of PAK autophosphorylation. Mol.Cell Biol. 20(11): 3906-3917.

Zimmermann P, Meerschaert K, Reekmans G, Leenaerts I, Small JV, Vandekerckhove J, David G and Gettemans J (2002). PIP_2-PDZ domain binding controls the association of syntenin with the plasma membrane. Mol. Cell 9: 1215–1225.

11. ORIGINAL COMMUNICATIONS

The original communications I-III are reprinted with permission from Elsevier and the original communication IV with permission from the Society for General Microbiology.