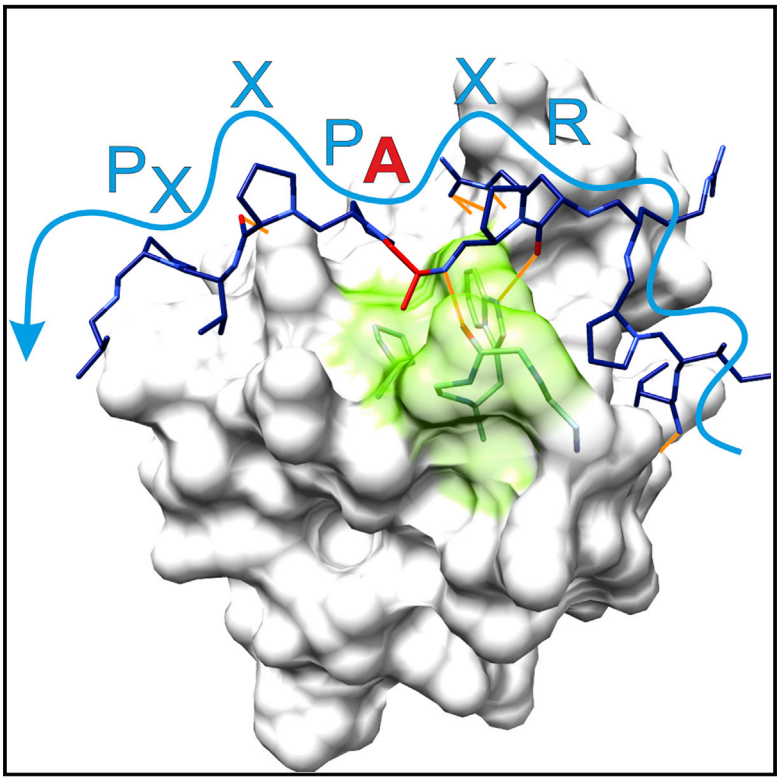
Article

Structure

Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs

Graphical abstract



Highlights

- SNX9 SH3 targets the RxAPxxP core motif found in EEEV nsP3 and cellular targets
- Binding induces structural and motional changes beyond the protein-ligand interface
- With a K_d of 0.3 μM EEEV nsP3 interacts with SNX9 SH3 28 times stronger than dynamin
- SNX9 SH3 interaction with the same motif in HTLV-1 Gag is important for infectivity

Authors

Helena Tossavainen, Hasan Uğurlu, Mikael Karjalainen, ..., Riku Fagerlund, Kalle Saksela, Perttu Permi

Correspondence

perttu.permi@jyu.fi

In brief

Tossavainen et al. reveal SNX9 SH3 domain target specificity by solving its structure in complex with an RxAPxxP motif-containing peptide found in EEEV nsP3. They show that the viral peptide binds the domain tighter than cellular ligands, and that the same motif found in HTLV-1 is required for efficient infection.

Tossavainen et al., 2022, Structure *30*, 1–12
June 2, 2022
https://doi.org/10.1016/j.str.2022.03.006

CellPress

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

Structure

CellPress
OPEN ACCESS

Article

Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs

Helena Tossavainen,[1,4] Hasan Uğurlu,[2,4] Mikael Karjalainen,[3] Maarit Hellman,[3] Lina Antenucci,[1,3] Riku Fagerlund,[2] Kalle Saksela,[2] and Perttu Permi[1,3,5,*]

[1]Department of Biological and Environmental Science, University of Jyvaskyla, Jyvaskyla FI-40014, Finland
[2]Department of Virology, University of Helsinki and Helsinki University Hospital, Helsinki FI-00014 Finland
[3]Department of Chemistry, Nanoscience Center, University of Jyvaskyla, Jyvaskyla FI-40014, Finland
[4]These authors contributed equally
[5]Lead contact
*Correspondence: perttu.permi@jyu.fi
https://doi.org/10.1016/j.str.2022.03.006

SUMMARY

Class I SH3 domain-binding motifs generally comply with the consensus sequence [R/K]xØPxxP, the hydrophobic residue Ø being proline or leucine. We have studied the unusual Ø = Ala-specificity of SNX9 SH3 by determining its complex structure with a peptide present in eastern equine encephalitis virus (EEEV) nsP3. The structure revealed the length and composition of the n-Src loop as important factors determining specificity. We also compared the affinities of EEEV nsP3 peptide, its mutants, and cellular ligands to SNX9 SH3. These data suggest that nsP3 has evolved to minimize reduction of conformational entropy upon binding, hence acquiring stronger affinity, enabling takeover of SNX9. The RxAPxxP motif was also found in human T cell leukemia virus-1 (HTLV-1) Gag polyprotein. We found that this motif was required for efficient HTLV-1 infection, and that the specificity of SNX9 SH3 for the RxAPxxP core binding motif was importantly involved in this process.

INTRODUCTION

Signaling networks that control cellular behavior depend on regulated proximity of proteins via interactions that are guided by modular protein units dedicated to this task. Several different families of such protein interaction domains (PIDs) are encoded by eukaryotic genomes, an archetype and the most copious family (~300 members) of which in the human proteome is the Src homology 3 (SH3) domain.

SH3 domains are small (~60 residues) protein modules that recognize their protein targets via proline-rich binding motifs that most typically contain consensus sequences [R/K]xØPxxP (class I sites) or PxØPx[R/K] (class II sites), where x is any and Ø is a hydrophobic residue (Feng et al., 1994; Lim et al., 1994). However, several other types of minimal consensus target sequences have been identified for different SH3 domains, and additional residues around or in some cases distant from these sites in the target protein contribute to define the specificity and affinity of SH3 binding (Saksela and Permi, 2012). The proline-rich stretches recognized by SH3 domains are representatives of peptide modules known as short linear motifs (SLiMs), abundant in cell signaling, and typically localized within natively disordered regions in target proteins.

Viruses and other pathogens have evolved to also encode proteins containing SH3-binding SLiMs in order to manipulate host cell signaling for promoting their replication and to interfere with host immune defenses. Although nanomolar binding affinities are an exception among SH3-mediated complexes between cellular proteins, and most of these interactions are relatively weak and transient (Kazlauskas et al., 2016), SH3 binding by pathogen proteins tend to be much stronger, as exemplified by recruitment of the Hck tyrosine kinase by the HIV-1 pathogenicity factor Nef (Saksela et al., 1995). Since this discovery in 1995, numerous examples of viral hijacking of SH3-mediated host cell processes have been reported (e.g., Heikkinen et al., 2008; Neuvonen et al., 2011; Schmotz et al., 2019).

For example, the Old World alphaviruses, including the arthralgia/arthritis- and rash-causing Chikungunya (CHIKV), Sindbis (SINV), and Semliki forest (SFV) viruses promote their replication by targeting the SH3 domain of host cell amphiphysins via a proline-rich motif in the non-structural protein 3 (nsP3) (Kim et al., 2016; Neuvonen et al., 2011; Tossavainen et al., 2016). On the other hand, the nsP3 of the eastern equine encephalitis virus (EEEV), the deadliest of the New World alphaviruses, associated with a fatality rate of 30%–70% in humans and severe neurologic impairment in survivors, does not target amphiphysin but instead can bind to the SH3 domain of the host cell sorting nexins SNX9 and SNX33 (Frolov et al., 2017).

Sorting nexins are a large and diverse group of proteins involved in cellular membrane trafficking and protein sorting,

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

and are grouped together because they all contain another canonical PID, namely a phospholipid-binding module known as the PX domain (Worby and Dixon, 2002). Sorting nexins 9, 18, and 33 form an SNX9 subfamily of sorting nexins that have related but not identical roles in endocytosis and membrane remodeling, and share a similar overall architecture, including an SH3 domain (Lundmark and Carlsson, 2009).

The SNX9-family SH3 domains are very similar in sequence, and bind to target motifs that conform to a class I consensus sequence RxØPxxP. However, these SH3 domains show an interesting preference for target motifs where the Ø position is occupied by an alanine (Howard et al., 1999; Kleino et al., 2009; Teyra et al., 2017). By contrast, alanine is rarely found in this position in the binding motifs for other SH3 domains, which prefer leucine or proline, and also accept isoleucine or valine as the Ø residue (Teyra et al., 2017). To better understand viral strategies for robust SH3 domain engagement, and to elucidate the molecular basis of the curious specificity of SNX9 SH3 for RxAPxxP-containing target motifs, we have solved the structure of the SNX9 SH3 domain in complex with a peptide derived from EEEV nsP3, introduced targeted mutations into this peptide, and compared their effects together with cellular ligands on SNX9 SH3 binding affinity.

These studies led us to identify another viral ligand for the SNX9 SH3 domain, namely the matrix (MA) subunit of the Gag polyprotein of human T cell leukemia virus-1 (HTLV-1), a deltaretrovirus causing an aggressive form of cancer known as adult T cell leukemia. We found that the SNX9 SH3 binding motif in HTLV-1 MA was required for efficient HTLV-1 infection, and that the specificity of SNX9 SH3 for the RxAPxxP core binding motif was importantly involved in this process.

RESULTS

Structure of the SNX9 SH3-EEEV nsP3 peptide complex

We solved the NMR solution structure of the complex using a sample that contained a ^{15}N, ^{13}C-labeled SH3 domain and unlabeled peptide AERLIPRRPA^{-1}P^{0}P^{1}VPVPARIPSPR of the hypervariable domain of EEEV nsP3, corresponding to residues 1,663–1,685 of the viral polyprotein. Ligand interface residues are numbered as in Yu et al. (1994), with the first proline in RxAPxxP designated as position 0. Chemical shift assignments were obtained using standard triple-resonance experiments for the SH3 domain and ^{15}N, ^{13}C-filtered experiments for the peptide (Permi and Annila, 2004; Sattler et al., 1999). Similarly, intradomain, intrapeptide, and intermolecular NOE restraints for structure calculation were derived from ^{15}N- and ^{13}C-edited and -filtered NOE spectra. Additional TALOS-N dihedral angle restraints based on assigned chemical shifts (Shen and Bax, 2013) were included for the SH3 domain. The final energy-minimized ensemble of 20 structures is shown in Figure 1 and the structural statistics in Table 1.

As is typical for SH3 domains, the structure of the SNX9 SH3 domain in the complex form is very similar to that in the free form, with an average heavy atom root-mean-square deviation (RMSD) of 1.1 Å for residues in the interaction interface (PDB: 2ENM; mouse SNX9 SH3, 89% sequence identity).

The complex structure revealed a dual-interaction interface, composed of the common combined hydrophobic core-charged specificity pocket interaction between an SH3 domain and a class I consensus motif, and of an additional hydrophobic binding surface, which accommodates peptide residues N terminal to the consensus motif (Figure 1). The core interaction involves SH3 Y9, F11, W39, P53, and Y56, which form two shallow hydrophobic xP sites on the surface of the domain and accommodate ligand A^{-1}P^{0} and V^{2}P^{3} moieties. E14, N17, E19 in the RT loop form the specificity pocket, within which E19 makes a salt bridge with R^{-3}. H-bonds between G37 O (n-Src loop)–A^{-1} H, W39 Hε1 (β3)–R^{-3} O, and Y56 Hη (3_{10} helix)–P^{1} O and hydrophobic contacts between sidechains of Y9 and V^{4} further stabilize the core interaction. The peptide adopts the typical extended polyproline II helix (PPII) conformation from R^{-3} ψ to P^{3} φ. In the extended N-terminal interaction, peptide L^{-7} and P^{-5} are key residues. Both pack against hydrophobic surfaces formed by SH3 N18, V35, E41, G50, L51 (L^{-7}) and V35, G36, and W39 (P^{-5}). Apart from V35, all interface residues are strictly conserved in vertebrate SNX9 SH3 amino acid sequences (Figure S1). V35 is conservatively replaced by an isoleucine in ∼15% of the sequences.

A similar hydrophobic extension to the canonical binding site has previously been described in the complex structure of the SH3 domain of the proto-oncogene tyrosine-protein kinase Src (c-Src) and VSL12 peptide, a class I peptide with a leucine at the −1 position from a biased phage display library (Bacarizo and Camara-Artigas, 2013; Feng et al., 1995). Although both SH3 domains bind their respective target peptides strongly, with a ∼0.5 μM affinity (Table 2; Feng et al., 1995), the question remains: how are these ligands specifically recognized by different SH3 domains? An overlay of the complex structures shows interesting differences, namely how only a few SH3 surface residues can confer specificity (Figure 2). In position −1, the backbones adapt to fit an alanine in EEEV nsP3 or a leucine in VSL12. In the latter, the bulky methyl groups are buried in a hydrophobic cleft bordered by the C-terminal residue of the n-Src loop, D117, and N135 in the 3_{10} helix. In SNX9 SH3, the surface is shallower with fewer available hydrophobic contacts for A^{-1} because a glycine (G38) occupies the position of D117. However, owing to the small size of glycine and the fact that the n-Src loop in SNX9 SH3 is one residue longer, the loop is able to approach the peptide backbone within H-bonding distance, thus gaining a favorable contribution to binding from a polar interaction instead. However, as discussed below, in addition to the direct local interaction between the binding partners, factors not decipherable from a static complex structure also contribute to specificity.

From position −1 toward the N termini, the peptide backbones diverge, although remaining between the n-Src and RT loops. With their arginine at position −3, both peptides canonically interact with the specificity pocket negatively charged residue, glutamate (E19) in SNX9 SH3 and aspartate (D99) in c-Src SH3. However, the arginine-glutamate interaction takes place above the conserved tryptophan, whereas the arginine-aspartate interaction is coplanar to the tryptophan aromatic ring. The former arrangement allows R^{-3}, E14, N17, and E19 to form an H-bond network while the latter optimally positions R^{-3} and W118 sidechains for a cation-π interaction. In both complexes, tryptophan ε1 is H-bonded to peptide backbone, to R^{-3} in the SNX9 and to P^{-2} in the c-Src complex. In the

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

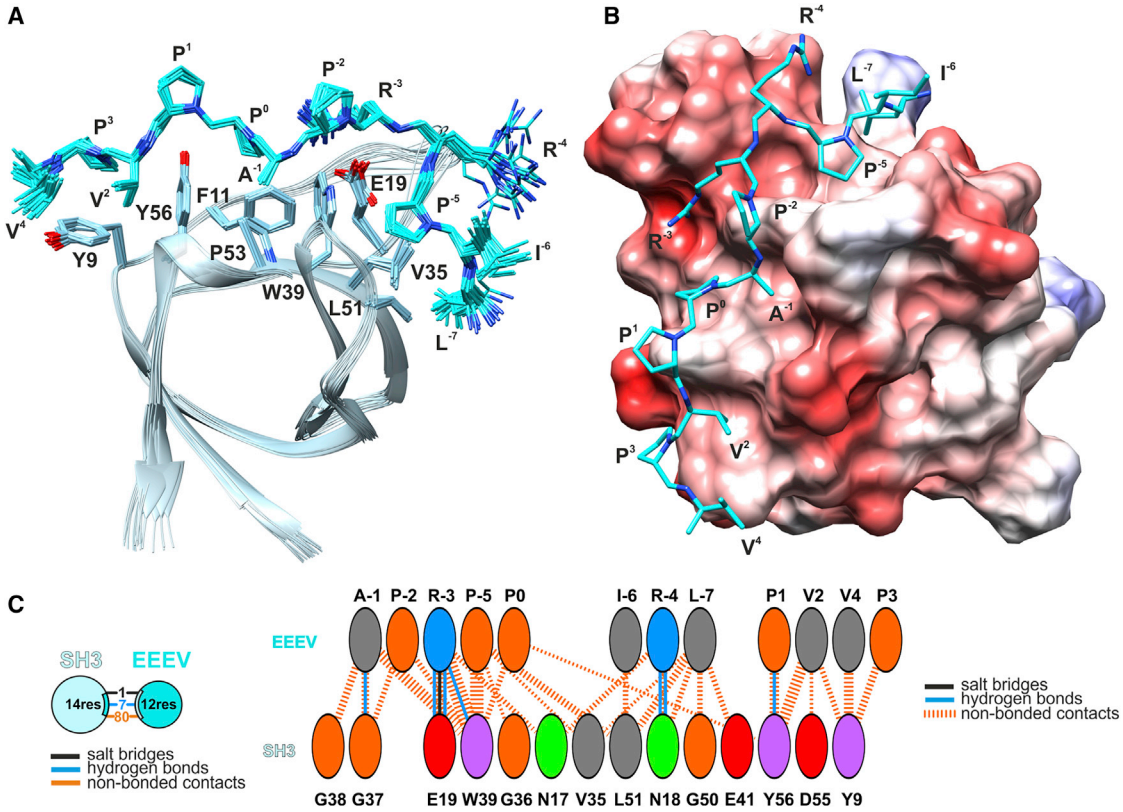


Figure 1. Solution NMR structure of the SNX9 SH3-EEEV nsP3 peptide complex

(A) The ensemble of 20 SNX9 SH3 structures. For the peptide (cyan) all side chains and for the SH3 domain side chains of residues forming the interaction interface are drawn and annotated with residue type and number. The disordered N- and C-terminal regions in both the peptide and the domain, and H atoms, are omitted for clarity.

(B) Surface representation of the interaction. Structure figures were created with UCSF Chimera (Pettersen et al., 2004).

(C) Interaction interface as represented by PDBsum (Laskowski et al., 2018) for the first conformer of the ensemble of structures. See also Figure S1.

c-Src complex, a buried water molecule bound to R^{-3} and D117 backbone atoms adequately fills the remaining space between the RT and n-Src loops (Bacarizo and Camara-Artigas, 2013). The presence of interfacial water molecules cannot be inferred from the solution NMR SNX9 SH3-EEEV nsP3 complex structure, but an identical arrangement is not feasible as G38 (the SNX9 equivalent of D117) and R^{-3} are too distant. The bulky interaction between P^{-5} and V35 in the SNX9 SH3 complex notably shifts the peptide backbone toward the RT loop as opposed to the corresponding interaction between A^{-5} and E115 side chain in the c-Src SH3 complex. The pseudocyclic interaction between L^{-6} O and R^{-3} guanidino group of VSL12 draws the L^{-6} sidechain in the vicinity of Y131. With a less curved backbone conformation, EEEV nsP3 creates a longer interaction interface bringing L^{-7} next to L51, which corresponds to Y131. Both leucines reside in hydrophobic clefts formed by N18, L51, and corresponding E98, Y131.

A^{-1} remarkably amplifies EEEV nsP3 affinity to SNX9 SH3

We further studied the significance of A^{-1} to the binding affinity of the viral peptide by replacing it by a leucine, the common residue at this position in typical SH3 ligands. We found that the leucine mutant has a drastically lower affinity to SNX9 SH3: in an NMR ligand titration study, SNX9 SH3 cross peaks moved linearly upon addition of EEEV $A^{-1}L$, indicating a fast-dissociating complex with micromolar to millimolar affinity (Figure 3). Curve fitting to changes in peak position in function of concentration ratios resulted in a K_d of 29 μM, which nicely agrees with the value obtained with ITC (28 μM; Table 2). This is a more than 90 times lower affinity than for the wild-type (WT) peptide, 0.3 μM.

Interestingly, in the complex between *Saccharomyces cerevisiae* Nbp2p SH3 domain and Ste20 peptide (PDB: 2LCS), replacement of an alanine in Ste20 position −1 with a threonine leads to a 29 times reduction in affinity (Gorelik and Davidson, 2012). The structure near Ste20 A^{-1} is remarkably similar to that in the EEEV nsP3 complex (Figure S2). Most importantly, both have a glycine as the C-terminal residue in the n-Src loop, enabling proximity of peptide backbone and the loop.

To confirm the importance of A-to-L swapping in position −1 for the recognition of two different class I-type ligands between SNX9 and c-Src SH3 domains, we also performed an NMR titration assay with c-Src SH3 and the $A^{-1}L$ mutant of EEEV nsP3, expecting this peptide to bind strongly because of the L^{-1} preference of c-Src SH3. Indeed, this is what we observed (Figure 3). The K_d is significantly lower than 10 μM, as shown from core

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

Table 1. NMR restraints and structural statistics for the ensemble of 20 SNX9 SH3-EEEV nsP3 peptide complex conformers of least restraint violations

Completeness of resonance assignments (%)[a]	
Backbone	97.0
Side chain, aliphatic/aromatic	97.1/100.0
Experimental restraints	
Distance restraints	
Total	2078
Short range (i-j $\leq$ 1)	1028
Medium range (1 < i-j < 5)	246
Long range (i-j $\geq$ 5, intramolecular)	632
Intermolecular	172
Dihedral angle restraints	95
No. of restraints per restrained residue	23.9
No. of long-range restraints per restrained residue	9.2
Residual restraints violations[b]	
Average no. of distance violations per structure	
0.1–0.2 Å	33.4
0.2–0.5 Å	0.7 (max. 0.3)
>0.5 Å	0
Average no. of dihedral angle violations per structure	
1–5°	0
>5°	0
Model quality[b]	
RMSD backbone/heavy atoms (Å)	0.3/0.6
RMSD bond lengths (Å)/bond angles (°)	0.013/2.1
Molprobity Ramachandran statistics (%)[b]	
Most favored regions	97.1
Allowed/disallowed regions	2.9/0.0
Global quality scores (raw/Z score)[b]	
Verify3D	0.23/-3.69
Prosall	0.41/-0.99
PROCHECK(φ-ψ)	−0.44/-1.42
PROCHECK (all)	−0.46/-2.72
Molprobity clash score	1.03/1.35
Model contents	
No. of ordered residues[b]/total no. of residues	72/90
BMRB accession number	34628
PDB code	7OJ9

[a]Backbone includes Cα, Cβ, N, and H, except the N-terminal amide. For side chains, excluded are the exchanging groups (K, amino; R, guanidino; S/T/Y hydroxyl; H δ1/ε2), as well as all unprotonated Cs and Ns. Only ^{1}H resonances were assigned for the EEEV nsP3 peptide. Not included in the tabulation: the unstructured N-terminal three-residue cloning artefact in SNX9 SH3.

[b]Ordered residues: 1–14, 17–61 (SNX9 SH3), 1,667–1,678 (EEEV nsP3). Computed using PSVS (Bhattacharya et al., 2007).

binding peak transitions from free to bound form without peak broadening. This further consolidates the difference is preferences in position −1 residue. Many peaks of c-Src SH3 residues interacting with the N-terminal portion of the peptide are broadened or have disappeared at 1:0.5 or 1:1 protein to peptide concentration ratios, which likely indicates that binding of the flanking sequence is not as tight. Lower affinity probably originates from the differences in amino acid composition between the two SH3 domains concentrated in the portions of the n-Src and RT loop that form the binding pocket for the flanking sequence (Figure 3).

Next, we studied the SNX9 SH3-EEEV nsP3 $A^{-1}L$ complex by performing atomistic molecular dynamics (MD) simulations of this complex, and, for comparison, of the EEEV nsP3 WT complex (Figure S3). The mutation does not significantly alter the conformational flexibility within the complex, as deduced from the very similar RMSF values over the trajectories. The per-residue solvent-accessible surface areas (SASAs) show minute differences, the largest of them, 15–24 Å^2, located in the first half of the RT loop (D10, A12, P15). What instead appear to be clearly different between the two complexes are the stabilities of certain H-bonds. In the WT complex, the interfacial W39 Hε1–R^{-3} O and G37 O–A^{-1} H-bond occupancies were ~73% and 31%, whereas in the $A^{-1}L$ complex these H-bonds were not present. On the other hand, an interfacial H-bond not present in the WT complex appeared between A12 backbone and R^{-3} sidechain in the mutant complex. Other mutation-related differences in H-bonding were observed within the SH3 RT loop. Namely, in the mutant complex occupancies in the middle of the loop (F11-V22) decrease, whereas they increase in residues toward the tip of the loop (F11-A13, E14-N17, P15-N18). All other H-bonds in the complexes, in total 36 with >30% occupancy, remain alike.

The MD trajectories reveal four relatively stable interfacial solvent molecules in both complexes (I^{-6}-wat-N17 bb, R^{-4} bb-wat-E19 sc, R^{-3} sc-wat-A12, P^{0}-wat-D55 bb). Nine SH3-bound solvent molecules are similarly present in both complexes. Two of these in the mutant complex show significantly higher occupancies (F11 H-wat-V22 O, E19-wat-T21), perhaps being able to compensate for the loss of inter-strand H-bonds (F11–V22) preset in the WT complex.

Distance measurements from the trajectories show that, as the larger Leu^{-1} does not fit to the position carved for an Ala^{-1}, the longer Leu^{-1} side chain pushes the peptide backbone farther from the SH3 surface. The Ala/Leu^{-1} methyl group is kept at a position similar with respect to the W39 indole ring, but the Trp sidechain in the mutant complex moves slightly outward; that is, toward the n-Src loop. Concomitantly, R^{-3} sidechain shifts toward the tip of the RT loop. These rearrangements are likely to be coupled to the changes in H-bonds and backbone φ/ψ angles in the RT loop of the mutant complex. In all, MD data suggest that the mutation causes structural differences in the vicinity of the mutation site but also in the RT loop.

This conclusion is supported by the chemical shift differences in the ^{1}H, ^{15}N HSQC spectra of WT and $A^{-1}L$ complexes (Figure 3). The two spectra appear noticeably different. The largest differences, Δδ = 0.15–0.45 ppm, are observed for A13, N17-L20 in the RT loop, V35, G37 in the n-Src loop, W39 ε1, and L40, V52, T54 around the tryptophan. The slightly different position of the Trp aromatic ring could affect peak positions. Additionally, as H^{N} chemical shifts are sensitive to H-bond lengths (Wagner et al., 1983), differences in ^{1}H, ^{15}N peak positions could arise from rearrangement of the H-bond network. That relatively

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

Table 2. Binding affinities and thermodynamic parameters for SNX9 SH3 interaction with EEEV nsP3, mutated EEEV nsP3 peptides, and cellular ligands

Peptide	Peptide sequence	K_d (μM)	ΔH	ΔG	−TΔS
EEEV nsP3	AERLIPRRPAP^{0}PVPVPARIPSPR	0.3 ± 0.03	−41.3 ± 0.5	−37.4 ± 0.2	3.9 ± 0.8
A^{-1}L	AERLIPRRPLP^{0}PVPVPARIPSPR	27.7 ± 7.45	−21.3 ± 0.4	−25.5 ± 0.0	−4.8 ± 0.3
R^{-8}A	AEALIPRRPAP^{0}PVPVPARIPSPR	0.63 ± 0.12	−42.5 ± 0.4	−35.5 ± 0.5	7.0 ± 0.8
L^{-7}S	AERSIPRRPAP^{0}PVPVPARIPSPR	1.0 ± 0.13	−44.6 ± 0.5	−34.2 ± 0.3	10.4 ± 0.7
L^{-7}Q	AERQIPRRPAP^{0}PVPVPARIPSPR	0.9 ± 0.09	−47.7 ± 0.3	−34.6 ± 0.2	13.1 ± 0.1
P^{-5}A	AERLIARRPAP^{0}PVPVPARIPSPR	1.0 ± 0.10	−42.7 ± 0.4	−34.2 ± 0.2	8.4 ± 0.6
P^{-2}A	AERLIPRRAAP^{0}PVPVPARIPSPR	0.6 ± 0.04	−42.6 ± 0.6	−35.7 ± 0.2	6.9 ± 0.4
P^{1}A	AERLIPRRPAP^{0}AVPVPARIPSPR	0.4 ± 0.01	−45.6 ± 1.0	−36.9 ± 0.1	8.7 ± 1.1
V^{4}S, P^{5}S	AERLIPRRPAP^{0}PVPSSARIPSPR	0.2 ± 0.00	−47.2 ± 1.0	−38.0 ± 0.0	9.2 ± 1.0
hADAM9	QGNLIPARPAP^{0}APPLYSSLT	7.0 ± 0.40	−41.3 ± 0.3	−29.5 ± 0.2	11.8 ± 0.1
hDYN1	TSSPTPQRRAP^{0}AVPPARPGS	8.5 ± 0.31	−56.4 ± 1.4	−29.0 ± 0.1	27.5 ± 1.3

P^0 is indicated and the site of mutation highlighted by underlining. Residues at the binding interface are aligned in EEEV nsP3 peptide sequence. ΔH, ΔG, and −TΔS are expressed in kJ/mol. All values are represented as mean ± SD. The stoichiometry, N, was 1 in all interactions. See also Figure S7 for the binding isotherms and fitted curves.

large differences ($\Delta\delta > 0.05$ ppm) in peak positions are also observed for M8, V22, N32, and N44, although they are distant from the mutation site, suggests that binding-induced global effects (see below) are somewhat different in the two complexes. Accordingly, differences in methyl group chemical shifts of residues both near and distant from the binding site are observed between ^{1}H, ^{13}C HSQC spectra of WT and A^{-1}L EEEV nsP3 complexes (Figure S4).

Ligand binding induces structural and motional changes beyond the protein-ligand interface

The binding interface between SNX9 SH3 and EEEV nsP3 peptide is mainly hydrophobic, as is the case for many other SH3 domains and their ligands. However, thermodynamic studies on SH3 interactions have shown that target binding is invariably driven by a favorable enthalpic contribution. This apparent inconsistency is the demonstration of the underlying complexity of the recognition of ligands by SH3 domains. Binding energetics are influenced by redistribution of conformations of both the domain and the ligand, by reorganization of the H-bond network, by global changes in SH3 domain dynamics, and by interfacial water molecules (e.g., Cordier et al., 2000; Ferreon and Hilser, 2003; Palencia et al., 2004; Wang et al., 2001). We studied the global effects of EEEV nsP3 peptide binding to the SNX9 SH3 domain by comparing chemical shifts, H to deuterium (H/D) exchange rates, and ^{15}N relaxation data of the free and bound states of the domain.

The majority of chemical shift perturbations upon ligand binding are expectedly observed for residues in the binding surface (Figure S4); however, several amide and methyl groups distant from the site experience notable chemical shift perturbations. Seven out of the eight amide peaks not in direct contact with the peptide but experiencing significant $\Delta\delta$s are among those that become more protected from solvent exchange upon ligand binding (Figure S5), suggesting that H-bonding might contribute to the observed chemical shift perturbations. Shifting of amide proton peaks from higher to lower field, as is the case for M8, L20, V22, L40, and V57, has been attributed to shorter H-bonds (Wagner et al., 1983). Methyl group shifts of A13, V22, and V52, which are within ~5 Å from F13, on the other hand, might be affected by repositioning of the aromatic ring. However, structural rearrangement deeper within the domain is likely to take place, as the farthest perturbed methyl groups (I31) are located >12 Å from the peptide.

In the free form, only seven backbone amide signals persisted to the first acquired ^{1}H, ^{15}N HSQC spectrum, at approximately 15 min from dissolution of the protein sample into D_2O (Figure S5). Five of them are strand residues of the SH3 β-barrel, whereas two reside in the RT loop. All are H-bonded in the ensemble of structures of free mouse SNX9 SH3 (PDB: 2ENM). All vanished during the acquisition of the second spectrum so that, at ~25 min, all protons had exchanged to deuterons. A substantially different extent of protection from exchange was observed for the complex form. In the first acquired spectrum, 34 out of the 61 amide peaks of the domain were present, and, after ~22 h, a set of 20 amide signals still persisted. These are located in the β-barrel strands, the RT loop, and the 3_{10} helix, and include all those that were momentarily protected in the free form. Upon complex formation, protection from exchange is increased at locations both near and remote from the binding site, the most distant protected amides being at ~15 Å distance from the peptide. This indicates that increased protection includes factors other than simple blocking of solvent access by the peptide. Protection depends on H-bond strength but also on all other energy terms associated with the closed and open forms of that bond, its surroundings, as well as the solvent (Wang et al., 2001).

We measured ^{15}N relaxation data from free and bound SNX9 SH3 and interpreted the data using the model-free approach (Lipari and Szabo, 1982) to extract residue-specific generalized order parameters S^2, internal correlation times τ_e, and exchange rates R_{ex}, parameters reporting on the spatial restriction of the amide bond vector motion, the rate of this motion, and conformational exchange, respectively (Figure S6). The backbones (residues K4-L60) in both forms are relatively rigid, with $S^2 < 0.8$ only for residues in the n-Src loops. Overall, ligand binding slightly reduces flexibility. The largest ligand-induced increase in restriction of motion is

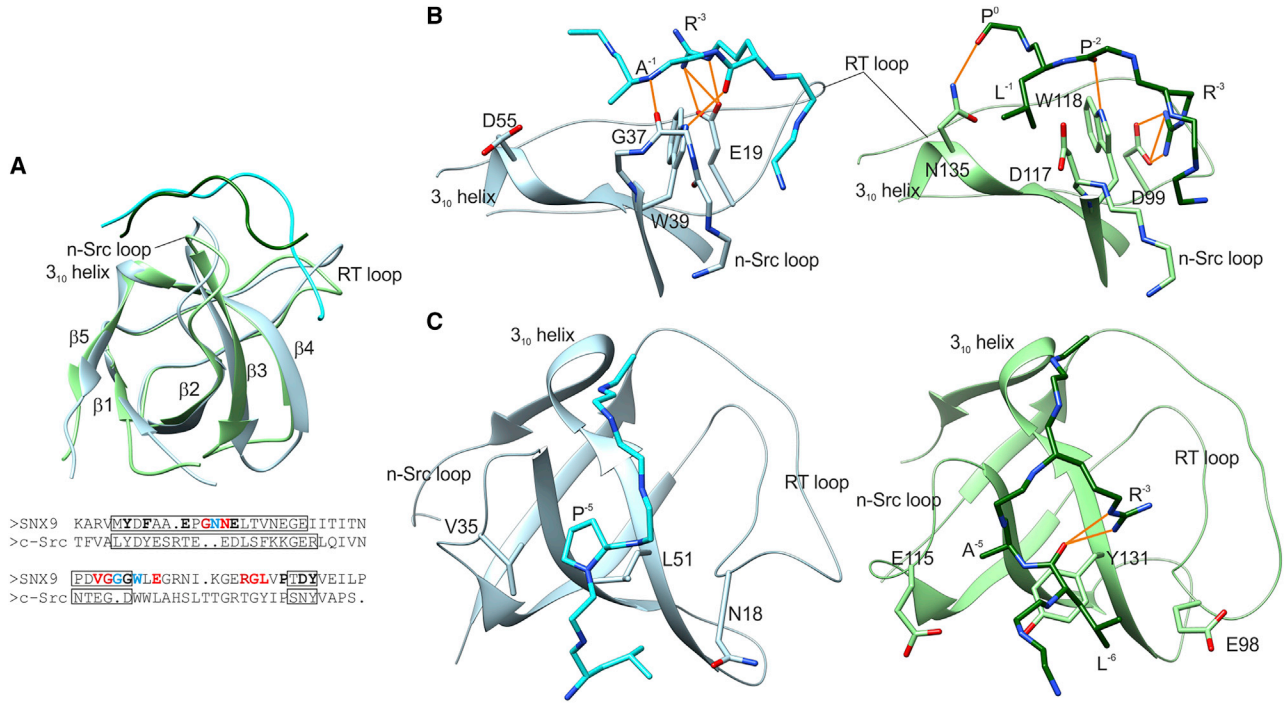


Figure 2. Comparison of SNX9 SH3-EEEV nsP3 and c-Src-VSL12 complex structures

(A) Overlay of structures, showing that, overall, the structures superimpose well, but n-Src and RT loop conformations are clearly distinct. SNX9 SH3-EEEV nsP3 is colored in light blue/cyan. The amino acid sequence alignment of the SH3 domains is shown below the structures. The RT, n-Src, and 3_{10} regions are boxed. Bold characters in the SNX9 SH3 sequence indicate residues that interact with the peptide, in black characters are indicated those that interact with the core RxAPxxP motif, in red those that interact with the flanking sequence, and in cyan those that interact with both. Interaction is defined as residues being less than 4 Å apart. The sequence alignment was created from structure superposition. The PDB ID of the c-Src-VSL12 complex structure is PDB: 4hvw.

(B) Structures highlight the Ala^{-1}/Leu^{-1} interactions in both complexes. H-bonds are shown with orange lines.

(C) Structures show the interactions between the N termini of the EEEV nsP3/VSL12 peptides and the respective SH3 domains. See also Figure S3.

observed for the last residue in the β1 strand (V7), residues in the tip of the RT loop (E14-E19), and the n-Src loop (G37). Changes in S^2 upon binding can be converted into changes in $-T\Delta S_{conf}$ of individual amide bonds (Yang and Kay, 1996). Their sum yields an overall unfavorable conformational entropy contribution of 45 kJ/mol at 25°C, which indicates that ΔS_{conf} derived from SH3 backbone fast-timescale motions makes a substantial contribution to overall binding entropy. However, EEEV nsP3 rigidifies upon binding to SNX9 SH3, a typical process for the proline-rich motifs targeting SH3 domains that contributes negatively to ΔS_{conf} (Aitio et al., 2012; Tossavainen et al., 2016), which implies that the overall reduction of conformational entropy must be highly compensated by other factors, $\Delta S_{solvent}$ in particular, contributing to entropy, as total $-T\Delta S$ is 4 kJ/mol. Contribution of peptide binding to slow-timescale motions can be recognized by monitoring changing R_{ex} values. E14 R_{ex} decreases, perhaps because its amide forms an H-bond with the side chain of E19, which in turn is stabilized by a salt bridge with R^{-3}. With an occupancy of 64%, the MD studies suggest a relatively stable H-bond for E14 H. G38 whose R_{ex} also decreases, is not H-bonded, nor does it appear to become more buried, but perhaps benefits from an overall conformational stabilization of the n-Src loop as the peptide binds in its vicinity. R_{ex} may also reflect changes in exchange between solvent and solute molecules. This might be the case for N18, D55, and V57, for which R_{ex} increases. They have in common, according to the MD studies, an interaction with a bridging solvent molecule (N18–wat–I^{-7}, 55% occ. and V57–wat–T54, 38% occ., D55–wat–P11, occ. 30%). Of note is that changes in the dynamics between the free and bound forms are not limited to the binding interface residues. Amides of residues as distant as I45 (14 Å from the peptide) were affected by binding. Overall, changes in chemical shift and amide protection as well as dynamics point to binding inducing global changes, consistent with observations for other SH3 domains.

Targeted mutations introduced in the EEEV nsP3 peptide and comparison with cellular ligands reveal conformational entropy-loss-optimized binding of viral peptides

We performed ITC experiments with nine peptides of different amino acid composition to characterize the thermodynamics of EEEV nsP3 binding to SNX9 SH3, and to compare its affinity with those of known cellular ligands of SNX9 SH3, namely ADAM9 and dynamin 1 (Howard et al., 1999; Shin et al., 2007). Both cellular ligands encompass two proline-rich regions, and, of those, peptides most similar to EEEV nsP3 were chosen; that is, the second region of ADAM9 (residues 800–819) and

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

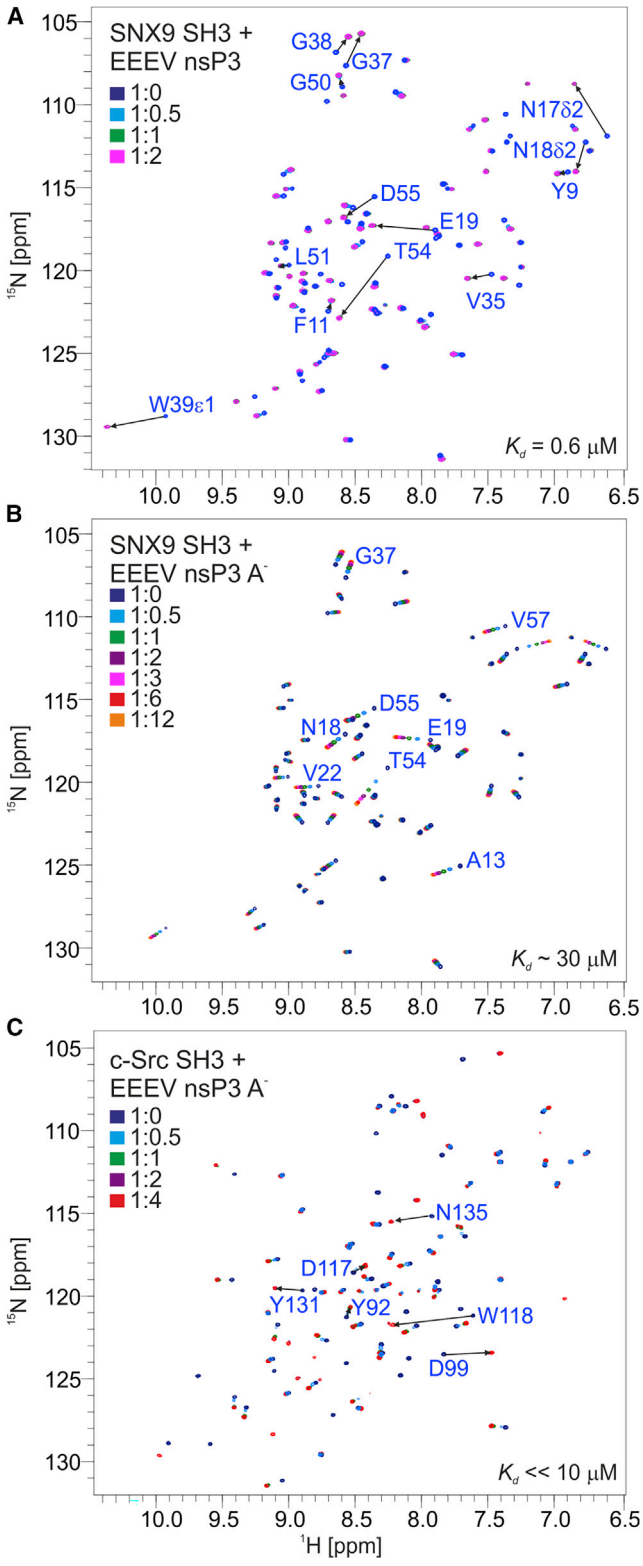


Figure 3. NMR titration data demonstrating the impact of alanine-to-leucine mutation to affinity

Shown are overlays of ^{1}H, ^{15}N HSQC spectra of SH3 domains acquired at increasing peptide concentrations. Color code for the SH3 to peptide ratios is shown in each spectrum.

(A) SNX9 SH3 titrated with EEEV nsP3 peptide. Arrows show cross-peak movement between free and bound forms for residues in contact with the peptide.

(B) SNX9 SH3 titrated with EEEV nsP3 A^{-1}L peptide.

(C) c-Src SH3 titrated with EEEV nsP3 A^{-1}L peptide. Arrows show presumed movement between free and bound form cross peaks. c-Src N, H assignments from Biological Magnetic Resonance DataBank (BMRB) entry 4888, but the sequence numbering corresponds to that in the 4hvw structure. See also Figure S2.

the first region of dynamin 1 (residues 776–795). ITC results are presented in Table 2 and binding isotherms in Figure S7.

In agreement with other SH3-proline-rich ligand interactions, binding of EEEV nsP3 peptide to SNX9 SH3 is characterized by large favorable enthalpy and unfavorable entropy. As mentioned above, this is inconsistent with a mainly hydrophobic interaction interface, and other factors than a simple local interaction between protein and ligand contribute to the thermodynamics. Indeed, as deduced from Δδ, H/D, and dynamics data, ligand binding to SNX9 SH3 transmits as structural and dynamical rearrangements throughout the domain. Also, changes in solute-solvent interactions as well as structural rearrangement of the ligand upon binding can affect the overall energetics. In the following, we ponder on the local effects of mutations on binding thermodynamics.

Apart from A^{-1}L, EEEV nsP3 peptide mutations have small effects on binding affinity, and the dissociation constants are within 0.7 μM from that of the WT peptide. However, the thermodynamic parameters of the mutants show a general tendency for more favorable enthalpies, counterbalanced by more unfavorable entropies compared with the EEEV nsP3 peptide.

For L^{-7}S/Q and P^{-5}A mutations, the increased entropic penalty might partly arise from polar residues or alanine not fitting as well in the hydrophobic sites as a leucine or proline, leading to unsatisfactory water dehydration. Additionally, a polar residue in the place of a leucine might affect the stability of the interfacial water molecule between I^{-6} and N17 backbones (MD study, Figure S3) and thus affect both entropy and enthalpy.

The solvent-exposed proline to alanine (P^{-2}A, P^{1}A) mutations, as well as the double mutation V^{4}S-P^{5}S also show an increase in ΔH and −TΔS of similar magnitudes. Reorganization of the network of water molecules (Breiten et al., 2013; Darby et al., 2019) and loss of conformational restrictions and polyproline II helical propensity provided by proline in the free form of the peptide (Brown and Zondlo, 2012) can contribute to the compensating changes in the enthalpy and entropy of binding, respectively.

These opposing effects are also active in the case of the cellular ligands ADAM9 and dynamin 1 binding to SNX9 SH3. These peptides have clearly lower affinities to SNX9 SH3 than the viral peptide or its mutants, apart from EEEV nsP3 A^{-1}L. Binding of ADAM9 results in a clear increase in −TΔS, while ΔH remains unchanged. ADAM9 entropic term entails unfavorable contributions from P^{1} to A substitution, as was the case for EEEV nsP3 P^{1}A, and, conversely, favorable contribution from a proline instead of a valine in position 2. In terms of PPII propensity, V^{4}L-P^{5}Y substitutions are mutually compensating as proline and leucine are high-PPII-propensity residues,

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

whereas valine and tyrosine are among residues that disfavor PPII the most (Brown and Zondlo, 2012). Replacement of a solvent-exposed arginine with an alanine in position −4 in ADAM9 is likely to unfavorably affect conformational entropy (Trbovic et al., 2009).

Dynamin 1 binding results in a large increase in $-T\Delta S$. As for ADAM9, in terms of conformational restriction and PPII propensity of the free form, the P^{-2} to R, P^{1} and P^{5} to A substitutions are likely to significantly contribute to the entropic penalty, compensated by the V^{4} to P substitution. The unfavorable entropy is largely counterbalanced by favorable enthalpy, but, as the peptide sequences are rather different, especially in the N-terminal binding motif, the origins of $\Delta\Delta H$ compared with EEEV nsP3 remain unsolved by the current data. Of note is that, similar to EEEV nsP3, both cellular ligands use their N-terminal flanking residues in binding (Figure S4); that is, the weaker affinity does not originate from a smaller interaction area.

SNX9 SH3 binding site in HTLV-1 Gag

Inspired by the data on the SNX9 SH3–EEEV nsP3 complex, we searched sequence databases for potential new viral ligands of SNX9. We noted that the matrix subunit of the Gag protein of human T cell leukemia virus-1 (HTLV-1) contains a peptide that conforms to the RxAPxxP consensus (Figure 4A), and interestingly shows a nine-residue stretch (PSRPAPPPP) of perfect sequence identity with human ADAM15 and *Escherichia coli* EspF proteins, both of which have been previously identified as SNX9 SH3 binding partners (Alto et al., 2007; Kleino et al., 2009).

We therefore investigated whether HTLV-1 Gag binds to SNX9 in human cells that co-express these proteins, and whether the HTLV-1 RxAPxxP motif would play a role in this. In addition to SNX9, we also tested the two other members of the homologous SH3 domain-containing SNX protein subfamily, namely SNX18 and SNX33. Expression vectors for these SNX proteins were co-transfected into HEK293 cells together with WT HTLV-1 Gag or variants of it dubbed Gag-AxxA and Gag-A>L, carrying mutations in the predicted SH3 binding motif. Gag-AxxA mutant contains an alanine substitution of both of the key proline residues defining a class I motif (RxxPxxP) plus a charge-reversing substitution of the third critical residue in the −3 position of this motif (i.e., RPAPPPP to EPAAPPA), which together can be expected to completely abrogate all SH3 binding capacity. By contrast, Gag-A>L contains only a subtle alanine-to-leucine change of the hydrophobic x residue of the first xP dipeptide of the RxxPxxP motif. Based on our data on the EEEV nsP3/SNX9 SH3 interaction presented above, this mutation would be expected to strongly and specifically reduce affinity for SNX9 SH3 without affecting binding of typical class I-targeting SH3 domains.

As shown by the co-precipitation data shown in Figure 4B, SNX9, -18, and -33 could all interact with HTLV-1 Gag in HEK293 cells, and, in each case, the capacity for SNX-binding of Gag-AxxA as well as Gag-A>L was insufficient for mediating Gag-SNX co-precipitation, and only traces of mutant Gag proteins could be detected in SNX pull-downs (reproducibly only with SNX18).

The capacity of HTLV-1 Gag to interact with SNX protein in human cells led us to examine the relevance of this SH3-mediated interaction for HTLV-1 replication and spread in tissue culture. To this end, we utilized a firefly luciferase-encoding HTLV-1 pseudovirus single-round replication system developed by Mazurov and colleagues (Mazurov et al., 2010; Shunaeva et al., 2015). This elegant system has been designed to express luciferase only after the HTLV-1 genomic RNA has been reverse transcribed and integrated in the infected cells. We co-transfected HEK293 cells with this engineered genomic HTLV-1 plasmid together with a vector expressing WT HTLV-1 Gag, Gag-AxxA, or Gag-A>L and followed the subsequent infectious spread of HTLV-1 in these cultures based on firefly luciferase activity. In order to monitor uniform transfection efficiency in these experiments, a vector for Renilla luciferase was also included in the transfection reactions, which, in direct contrast to HTLV-1-encoded firefly luciferase, is specifically expressed only in the transfected but not in the infected cells. Only experiments in which the Renilla luciferase values differed less than 25% from the average between the transfections with the three Gag variants were considered for further analysis.

Figure 5 shows results from three independent experiments comparing the infectious potential of HTLV-1 assembled from Gag-WT, Gag-AxxA, or Gag-A>L. Complete disruption of the class I SH3 binding site in Gag (AxxA mutation) resulted in a profound decrease in HTLV-1 infectivity to an average of 17% of that of the WT virus. The SNX selective A-to-L mutation in the −1 position of the HTLV-1 Gag class I SH3 binding motif also strongly reduced HTLV-1 infectivity (down to 31% of WT), in agreement with the large but not complete loss in the SNX9 SH3 binding affinity documented for this amino acid change by ITC measurements (Table 2).

Based on these data, we conclude that the interaction of HTLV-1 Gag with host cell SNX9 (and possibly SNX18/33) is important for the assembly of infectious HTLV-1 and its spread from one cell to another, and that the unique specificity of SNX SH3 domains for class I ligands with an alanine in −1 position guides this interaction, which HTLV-1 uses to hijack the host cell protein sorting machinery to facilitate its replication.

DISCUSSION

The sophisticated cellular machinery in eukaryotes is based on tight spatial and temporal regulation of key molecular interactions. Moreover, signal transduction is mediated by relatively low affinity but highly specific protein-protein interactions via SLiM recognition elements. Dissection of binding affinity from binding specificity is governed by a relatively high penalty in ΔS_{conf} of SLiM interactions; i.e., the recognition amino acid sequence, which bears a significant degree of disorder as a free ligand, undergoes folding upon binding. This contributes negatively to ΔS_{conf}, resulting in highly specific but relatively low-affinity interactions.

If a pathogen is to competitively bind a cellular motif, it does so with a motif that is fine-tuned for higher affinity. Displacement by an improved motif has been described as the mechanism of action for many SH3 domain-targeting viruses. For instance, extending the core PxxP binding motif with residues in the folded parts of HIV-1 Nef increases its affinity to Hck SH3 300-fold (Lee et al., 1995, 1996). SFV, SINV, and CHIKV avidly interact with amphiphysin 1 and 2 SH3 domains (Neuvonen et al., 2011). CHIKV's exceptional nanomolar binding affinity arises from the virus exploiting successive basic residues in its motif, which interact with the large negatively charged patch on the

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

CellPress
OPEN ACCESS

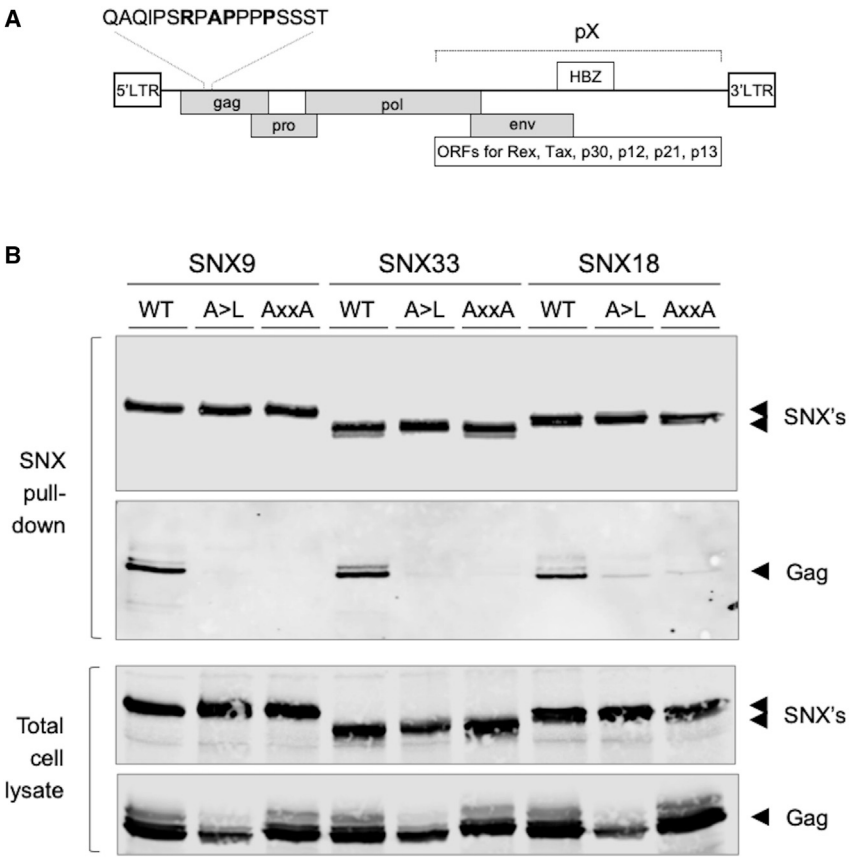


Figure 4. HTLV-1 genomic organization and interaction of SNX proteins with viral Gag protein

(A) Genomic organization and protein coding capacity of HTLV-1. The location and sequence of the SNX SH3 domain target site is indicated.

(B) Western blot analysis of wild-type (WT) and mutant (A>L and AxxA) HTLV-1 Gag protein co-precipitation with SNX9, SNX33, or SNX18 from HEK293 cells (SNX pull-down). Uniform expression of these proteins in the co-transfected cells is shown in the two bottom panels (total cell lysate).

surface of the SH3 domain (Tossavainen et al., 2016). Electrostatics is also the key determinant in the high-affinity interaction of HCV NS5A with the same target SH3 domain, and a 17 times higher affinity compared with the cellular ligand (Aladag et al., 2014). In the herpesviral Tip-Lyn SH3 interaction, the core flanking residues provide stabilizing hydrophobic contacts instead, and a 15-fold stronger interaction (Bauer et al., 2005). $EspF_U$, a bacterial effector from enterohemorrhagic *E. coli*, triggers pathogen-driven actin polymerization by replacing an alanine, found in cellular ligands of IRTKS SH3 between tandem PxxP motifs, with a tryptophan to yield a 60-fold increase in affinity (Aitio et al., 2010, 2012). The alanine-tryptophan swapping is entropically unfavorable due to *cis* to *trans* isomerization of the proline succeeding the tryptophan upon binding. This is, however, positively compensated by a larger gain in ΔH resulting in higher affinity for $EspF_U$ (Aitio et al., 2012). Here, on the contrary, we have shown that *in vitro* EEEV nsP3 proline-rich peptide binds SNX9 SH3 with an affinity 28 times higher than its binding to dynamin 1, but the data suggest that the overall strategy is to reduce the entropic penalty associated with the reduction of ΔS_{conf} upon SNX9 SH3 binding. This reduction is at least partially achieved by increasing the number of prolines in the core binding motif; i.e., by mimicking the binding (PPII helix) conformation and hence minimizing conformational readjustment upon binding. Therefore, the overall ΔS contributes favorably to the Gibbs' free energy of binding, resulting in a stronger affinity compared with those of the cellular ligands. This interaction confirms the role of the SH3 domain in the previously observed strong interaction between EEEV and SNX9 (Frolov et al., 2017). By superseding cellular dynamin, EEEV could then exploit SNX9 BAR-domain-mediated membrane remodeling in a manner reminiscent of that of other alphaviruses using amphiphysin BAR domain in replication (Neuvonen et al., 2011).

SNX9 SH3 has a unique specificity for alanine in ligand position −1. Alanine is non-interchangeable with leucine, the typical residue at this position in canonical SH3 ligands, as demonstrated by the 92 times weaker complex formed with EEEV nsP3 $A^{-1}L$ mutant peptide. It appears that alanine in peptide position −1 and glycine as the last residue in the n-Src loop allow for close interaction of peptide and the n-Src loop, resulting in an H-bond between backbone atoms. Concurrently, SH3 tryptophan side chain is able to form a second H-bond between protein and ligand, and R^{-3} side chain is optimally positioned to participate in a network of H-bonds with E14, N17, and E19 in the RT loop. Formation of H-bonds translates into favorable binding enthalpy and affinity. However, this is an oversimplified view of the basis of alanine preference. The measured binding affinity is the result of an intricate combination of local favorable and unfavorable enthalpic and entropic effects, encompassing contributions not only from both solutes but also the solvent. Moreover, as shown by the comparison of chemical shifts, exchange protection, and dynamics data between the free and bound forms of SNX9 SH3, EEEV nsP3 binding-induced structural and dynamical changes are transmitted across the domain, which is also reflected in binding thermodynamics.

In the SNX9, -18, and -33 subgroup of SNX proteins SNX9 SH3 domain amino acid sequence is ∼37% identical with SNX18 and SNX33 SH3 domain sequences, whereas SNX18 and SNX33 sequences are ∼59% identical. According to amino acid sequence alignment (Figure S1), the interaction interface residues Y9, F11, E19, G38, W39, G50, L51, P53, and Y56 are strictly conserved in the three SNXs. On the other hand, the amino acid composition and spacing of the n-Src loop and the tip of the RT loop clearly differ, suggesting differences in loop conformations and in intricacies of protein-ligand interactions.

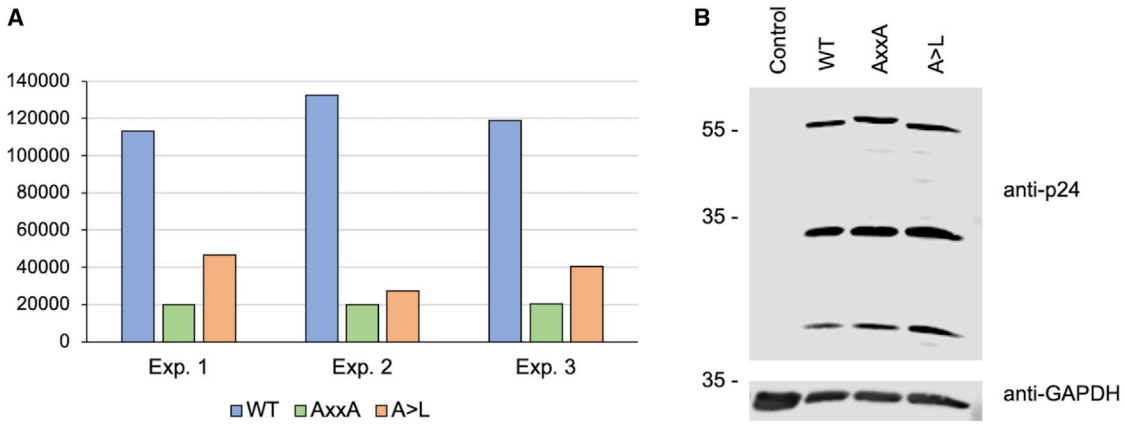


Figure 5. Infectivity of firefly luciferase-encoding HTLV-1 pseudoviruses

(A) Infectivity of firefly luciferase-expressing HTLV-1 pseudoviruses containing WT or SH3 binding site-mutated (AxxA or A>L) Gag proteins. Shown are firefly luciferase activity values in the infected HEK293 cells from three independent experiments. Each bar represents an average value of transfection/infection experiments performed in duplicate cultures in six-well plates.

(B) Western blot analysis of uniform expression of HTLV-1 Gag pr53 and its proteolytically processed forms in the transfected cells of a representative experiment, experiment 1 in (A). Untransfected cells are shown as control for anti-p24. Uniform transfection efficiency was further controlled using a co-transfected Renilla luciferase-encoding vector, as explained in the text.

Nevertheless, as indicated by binding of WT HTLV-1 Gag but not the Gag-A>L mutant to SNX9, -18, and -33, the SH3 domains of all these three related SNX proteins bind their ligands in an A^{-1}-specific manner. Strong A^{-1} preference was also evident for SNX9 and SNX33 SH3 domains in the study by Teyra et al. (2017), which investigated peptides selected by individual human SH3 domains from a large phage-displayed library of random peptides, whereas the target peptides favored by SNX18 SH3 in this experimental setup showed a more mixed pattern. Perhaps related to this, in a proteomics study to identify host cell co-factors of EEEV, Frolov et al. (2017) could identify SNX9 and SNX33 but not SNX18 as nsP3-associated proteins by mass spectroscopy.

While the detailed function of A^{-1}-dependent SNX recruitment in the replicative cycle of EEEV and HTLV-1 remain to be established, our studies clearly showed that this interaction is important for efficient spread of HTLV-1 infection. Considering that two viruses as divergent as EEEV and HTLV-1 have evolved to exploit SNX binding in their host cell subversion, it seems likely that other viruses may utilize this strategy as well. Further studies on SH3-mediated hijacking of SNX functions could therefore reveal important new insights into viral cell biology as well as novel targets for antiviral drug development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacterial cell culture
- METHOD DETAILS
  - Protein expression and purification
  - NMR spectroscopy
  - MD simulations
  - Cells and plasmids
  - Co-precipitation studies
  - HTLV-1 viral transmission
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2022.03.006.

ACKNOWLEDGMENTS

We acknowledge CSC-IT Center for Science for the allocation of computational resources. We thank Laura Pitkänen for technical support. This work is supported by the Jane and Aatos Erkko Foundation and grants from the Academy of Finland (grant 323435 to P.P.) and Helsinki University Central Hospital Research Council (grant TYH2020239 to K.S.). H.U. has been supported by the Turkish Ministry of National Education.

AUTHOR CONTRIBUTIONS

Conceptualization, K.S. and P.P.; investigation, all authors; writing – original draft, H.T., H.U., K.S., and P.P.; writing – review & editing, all authors; supervision, K.S. and P.P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 12, 2021
Revised: December 22, 2021
Accepted: March 4, 2022
Published: March 30, 2022

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

REFERENCES

Aitio, O., Hellman, M., Kazlauskas, A., Vingadassalom, D.F., Leong, J.M., Saksela, K., and Permi, P. (2010). Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly. Proc. Natl. Acad. Sci. U. S. A. *107*, 21743–21748.

Aitio, O., Hellman, M., Skehan, B., Kesti, T., Leong, J.M., Saksela, K., and Permi, P. (2012). Enterohaemorrhagic Escherichia coli exploits a tryptophan switch to hijack host f-actin assembly. Structure *20*, 1692–1703.

Aladag, A., Hoffmann, S., Stoldt, M., Bösing, C., Willbold, D., and Schwarten, M. (2014). Hepatitis C virus NS5A is able to competitively displace c-Myc from the Bin1 SH3 domain in vitro. J. Pept. Sci. *20*, 334–340.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Alto, N.M., Weflen, A.W., Rardin, M.J., Yarar, D., Lazar, C.S., Tonikian, R., Koller, A., Taylor, S.S., Boone, C., Sidhu, S.S., et al. (2007). The type III effector EspF coordinates membrane trafficking by the spatiotemporal activation of two eukaryotic signaling pathways. J. Cell Biol. *178*, 1265–1278.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. *44*, W344–W350.

Bacarizo, J., and Camara-Artigas, A. (2013). Atomic resolution structures of the c-Src SH3 domain in complex with two high-affinity peptides from classes I and II. Acta Crystallogr. D Biol. Crystallogr. *69*, 756–766.

Bauer, F., Schweimer, K., Meiselbach, H., Hoffmann, S., Rösch, P., and Sticht, H. (2005). Structural characterization of Lyn-SH3 domain in complex with a herpesviral protein reveals an extended recognition motif that enhances binding affinity. Protein Sci. *14*, 2487–2498.

Beckwith, M.A., Erazo-Colon, T., and Johnson, B.A. (2021). RING NMR dynamics: software for analysis of multiple NMR relaxation experiments. J. Biomol. NMR *75*, 9–23.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Bhattacharya, A., Tejero, R., and Montelione, G.T. (2007). Evaluating protein structures determined by structural genomics consortia. Proteins *66*, 778–795.

Boyko, R., and Sykes, B. (University of Alberta). Xcrvfit: A Graphical X-Windows Program for Binding Curve Studies and NMR Spectroscopic Analysis, Developed by Boyko, R. And Sykes, B.D. (University of Alberta)

Breiten, B., Lockett, M.R., Sherman, W., Fujita, S., Al-Sayah, M., Lange, H., Bowers, C.M., Heroux, A., Krilov, G., and Whitesides, G.M. (2013). Water networks contribute to enthalpy/entropy compensation in protein–ligand binding. J. Am. Chem. Soc. *135*, 15579–15584.

Brown, A.M., and Zondlo, N.J. (2012). A propensity scale for type II polyproline helices (PPII): aromatic amino acids in proline-rich sequences strongly disfavor PPII due to proline-aromatic interactions. Biochemistry *51*, 5041–5051.

Case, D.A., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, T.E., III, Cruzeiro, V.W.D., Darden, D.A., Duke, R.E., Ghoreishi, D., Gilson, M.K., et al. (2018). AMBER 2018 (University of California).

Cordier, F., Wang, C., Grzesiek, S., and Nicholson, L.K. (2000). Ligand-induced strain in hydrogen bonds of the c-Src SH3 domain detected by NMR. J. Mol. Biol. *304*, 497–505.

Darby, J.F., Hopkins, A.P., Shimizu, S., Roberts, S.M., Brannigan, J.A., Turkenburg, J.P., Thomas, G.H., Hubbard, R.E., and Fischer, M. (2019). Water networks can determine the affinity of ligand binding to proteins. J. Am. Chem. Soc. *141*, 15818–15826.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Dosset, P., Hus, J.-C., Blackledge, M., and Marion, D. (2000). Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. J. Biomol. NMR *16*, 23–28.

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

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

Ferreon, J.C., and Hilser, V.J. (2003). The effect of the polyproline II (PPII) conformation on the denatured state entropy. Protein Sci. *12*, 447–457.

Fowler, N.J., Sljoka, A., and Williamson, M.P. (2020). A method for validating the accuracy of NMR protein structures. Nat. Commun. *11*, 6321.

Frolov, I., Kim, D.Y., Akhrymuk, M., Mobley, J.A., and Frolova, E.I. (2017). Hypervariable domain of eastern equine encephalitis virus nsP3 redundantly utilizes multiple cellular proteins for replication complex assembly. J. Virol. *91*, e00371–17.

Gorelik, M., and Davidson, A.R. (2012). Distinct peptide binding specificities of Src homology 3 (SH3) protein domains can be determined by modulation of local energetics across the binding interface. J. Biol. Chem. *287*, 9168–9177.

Grigsby, I.F., Zhang, W., Johnson, J.L., Fogarty, K.H., Chen, Y., Rawson, J.M., Crosby, A.J., Mueller, J.D., and Mansky, L.M. (2010). Biophysical analysis of HTLV-1 particles reveals novel insights into particle morphology and Gag stochiometry. Retrovirology *7*, 75.

Güntert, P., and Buchner, L. (2015). Combined automated NOE assignment and structure calculation with CYANA. J. Biomol. NMR *62*, 453–471.

Heikkinen, L.S., Kazlauskas, A., Melén, K., Wagner, R., Ziegler, T., Julkunen, I., and Saksela, K. (2008). Avian and 1918 Spanish influenza a virus NS1 proteins bind to Crk/CrkL Src homology 3 domains to activate host cell signaling. J. Biol. Chem. *283*, 5719–5727.

Howard, L., Nelson, K.K., Maciewicz, R.A., and Blobel, C.P. (1999). Interaction of the metalloprotease disintegrins MDC9 and MDC15 with two SH3 domain-containing proteins, endophilin I and SH3PX1. J. Biol. Chem. *274*, 31693–31699.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Karjalainen, M., Tossavainen, H., Hellman, M., and Permi, P. (2020). HACANCOi: a new Hα-detected experiment for backbone resonance assignment of intrinsically disordered proteins. J. Biomol. NMR *74*, 741–752.

Kazlauskas, A., Schmotz, C., Kesti, T., Hepojoki, J., Kleino, I., Kaneko, T., Li, S.S.C., and Saksela, K. (2016). Large-scale screening of preferred interactions of human Src homology-3 (SH3) domains using native target proteins as affinity ligands. Mol. Cell. Proteomics *15*, 3270–3281.

Kim, D.Y., Reynaud, J.M., Rasalouskaya, A., Akhrymuk, I., Mobley, J.A., Frolov, I., and Frolova, E.I. (2016). New World and Old World alphaviruses have evolved to exploit different components of stress granules, FXR and G3BP proteins, for assembly of viral replication complexes. PLoS Pathog. *12*, e1005810.

Kleino, I., Ortiz, R.M., Yritys, M., Huovila, A.-P.J., and Saksela, K. (2009). Alternative splicing of ADAM15 regulates its interactions with cellular SH3 proteins. J. Cell. Biochem. *108*, 877–885.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucl. Acids Res. *33*, W299–W302.

Laskowski, R.A., Jabłońska, J., Pravda, L., Vařeková, R.S., and Thornton, J.M. (2018). PDBsum: structural summaries of PDB entries. Protein Sci. *27*, 129–134.

Lee, C.H., Leung, B., Lemmon, M.A., 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*, 5006–5015.

Lee, C.H., Saksela, K., Mirza, U.A., Chait, B.T., and Kuriyan, J. (1996). Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell *85*, 931–942.

Lim, W.A., Richards, F.M., and Fox, R.O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. Nature *372*, 375–379.

Lipari, G., and Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J. Am. Chem. Soc. *104*, 4546–4559.

Lundmark, R., and Carlsson, S.R. (2009). SNX9 - a prelude to vesicle release. J. Cell Sci. *122*, 5–11.

Maciejewski, M.W., Schuyler, A.D., Gryk, M.R., Moraru, I.I., Romero, P.R., Ulrich, E.L., Eghbalnia, H.R., Livny, M., Delaglio, F., and Hoch, J.C. (2017). NMRbox: a resource for biomolecular NMR computation. Biophys. J. *112*, 1529–1534.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. *47*, W636–W641.

Mazurov, D., Ilinskaya, A., Heidecker, G., Lloyd, P., and Derse, D. (2010). Quantitative comparison of HTLV-1 and HIV-1 cell-to-cell infection with new replication dependent vectors. PLoS Pathog. *6*, e1000788.

Morgenstern, B. (2004). DIALIGN: multiple DNA and protein sequence alignment at BiBiServ. Nucleic Acids Res. *32*, W33–W36.

Neuvonen, M., Kazlauskas, A., Martikainen, M., Hinkkanen, A., Ahola, T., and Saksela, K. (2011). SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. PLoS Pathog. *7*, e1002383.

Palencia, A., Cobos, E.S., Mateo, P.L., Martínez, J.C., and Luque, I. (2004). Thermodynamic dissection of the binding energetics of proline-rich peptides to the Abl-SH3 domain: implications for rational ligand design. J. Mol. Biol. *336*, 527–537.

Permi, P., and Annila, A. (2004). Coherence transfer in proteins. Prog. Nucl. Magn. Reson. Spectrosc. *44*, 97–137.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Roe, D.R., and Cheatham, T.E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J. Chem. Theor. Comput. *9*, 3084–3095.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. *23*, 327–341.

Saksela, K., and Permi, P. (2012). SH3 domain ligand binding: what's the consensus and where's the specificity? FEBS Lett. *586*, 2609–2614.

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*, 484–491.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucl. Magn. Reson. Spectrosc. *34*, 93–158.

Schmotz, C., Uğurlu, H., Vilen, S., Shrestha, S., Fagerlund, R., and Saksela, K. (2019). MC159 of molluscum contagiosum virus suppresses autophagy by recruiting cellular SH3BP4 via an SH3 domain-mediated interaction. J. Virol. *93*, e01613–18.

Shen, Y., and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J. Biomol. NMR *56*, 227–241.

Shin, N., Lee, S., Ahn, N., Kim, S.-A., Ahn, S.-G., YongPark, Z., and Chang, S. (2007). Sorting nexin 9 interacts with dynamin 1 and N-WASP and coordinates synaptic vesicle endocytosis. J. Biol. Chem. *282*, 28939–28950.

Shunaeva, A., Potashnikova, D., Pichugin, A., Mishina, A., Filatov, A., Nikolaitchik, O., Hu, W.-S., and Mazurov, D. (2015). Improvement of HIV-1 and human T cell lymphotropic virus type 1 replication-dependent vectors via optimization of reporter gene reconstitution and modification with intronic short hairpin RNA. J. Virol. *89*, 10591–10601.

Søreng, K., Munson, M.J., Lamb, C.A., Bjørndal, G.T., Pankiv, S., Carlsson, S.R., Tooze, S.A., and Simonsen, A. (2018). SNX18 regulates ATG9A trafficking from recycling endosomes by recruiting Dynamin-2. EMBO Rep. *19*, e44837.

Teyra, J., Huang, H., Jain, S., Guan, X., Dong, A., Liu, Y., Tempel, W., Min, J., Tong, Y., Kim, P.M., et al. (2017). Comprehensive analysis of the human SH3 domain family reveals a wide variety of non-canonical specificities. Structure *25*, 1598–1610.e3.

Tossavainen, H., Aitio, O., Hellman, M., Saksela, K., and Permi, P. (2016). Structural basis of the high affinity interaction between the alphavirus nonstructural protein-3 (nsP3) and the SH3 domain of amphiphysin-2. J. Biol. Chem. *291*, 16307–16317.

Trbovic, N., Cho, J.-H., Abel, R., Friesner, R.A., Rance, M., and Palmer, A.G. (2009). Protein side-chain dynamics and residual conformational entropy. J. Am. Chem. Soc. *131*, 615–622.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687–696.

Wagner, G., Pardi, A., and Wuethrich, K. (1983). Hydrogen bond length and proton NMR chemical shifts in proteins. J. Am. Chem. Soc. *105*, 5948–5949.

Wang, C., Pawley, N.H., and Nicholson, L.K. (2001). The role of backbone motions in ligand binding to the c-Src SH3 domain. J. Mol. Biol. *313*, 873–887.

Worby, C.A., and Dixon, J.E. (2002). Sorting out the cellular functions of sorting nexins. Nat. Rev. Mol. Cell Biol. *3*, 919–931.

Yang, D., and Kay, L.E. (1996). Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. J. Mol. Biol. *263*, 369–382.

Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W., and Schreiber, S.L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. Cell *76*, 933–945.

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-HTLV-1 p24	Abcam	Cat#ab9081; RRID: AB_306989
IRDye 680RD	LI-COR Biosciences	Cat#925-68070; RRID: AB_2651128
IRDye 800CW Streptavidin	LI-COR Biosciences	Cat#926-32230
GAPDH (14C10) Rabbit mAb	Cell Signaling Technology	Cat#2118; RRID: AB_561053
IRDye 800CW	LI-COR Biosciences	Cat#926-32211; RRID: AB_621843
Bacterial and virus strains		
E. coli: BL21(DE3) competent cells	Invitrogen	Cat#EC0114
Chemicals, peptides, and recombinant proteins		
Recombinant protein: SNX9 SH3	(Karjalainen et al., 2020)	N/A
Recombinant protein: c-Src SH3	This paper	N/A
Peptide: EEEV nsP3 AERLIPRRPAP^{0}PVPVPARIPSPR	GenScript	N/A
Peptide: A^{-1}L EEEV nsP3; AERLIPRRP**L**P^{0}PVPVPARIPSPR	GenScript	N/A
Peptide: R^{-8}A EEEV nsP3; AE**A**LIPRRPAP^{0}PVPVPARIPSPR	GenScript	N/A
Peptide: L^{-7}S EEEV nsP3; AER**S**IPRRPAP^{0}PVPVPARIPSPR	CASLO ApS	N/A
Peptide: L^{-7}Q EEEV nsP3; AER**Q**IPRRPAP^{0}PVPVPARIPSPR	CASLO ApS	N/A
Peptide: P^{-5}A EEEV nsP3; AERLI**A**RRPAP^{0}PVPVPARIPSPR	CASLO ApS	N/A
Peptide: P^{-2}A EEEV nsP3; AERLIPRR**A**AP^{0}PVPVPARIPSPR	CASLO ApS	N/A
Peptide: P^{1}A EEEV nsP3; AERLIPRRPAP^{0}**A**VPVPARIPSPR	GenScript	N/A
Peptide: V^{4}S, P^{5}S EEEV nsP3; AERLIPRRPAP^{0}PVP**SS**ARIPSPR	GenScript	N/A
Peptide: hADAM9; QGNLIPARPAP^{0}APPLYSSLT	CASLO ApS	N/A
Peptide: hDYN1; TSSPTPQRRAP^{0}AVPPARPGS	CASLO ApS	N/A
Protease inhibitor	Thermo Fisher Scientific	Cat#A32963
M−280 Streptavidin Dynabeads	Invitrogen	Cat#11205D
Lysis reagent	Promega	Cat#E153A
Critical commercial assays		
Firefly Luciferase Assay Kit	Promega	Cat#E1501
Oligonucleotides		
Primers for generating RPAPPPP to EPAAPPA Gag mutation to the plasmid peYFP-HTLV-1 Sense: P5′-CAGCCCCCCCCGCCAGCTCGCCAACCCA-3′ Antisense: P5′-CAGGCTCAGAGGGAATCTGCGCC-3′	Integrated DNA Technologies BVBA, Leuven, Belgium	N/A
Primers for generating A to L (RPAPPPP > RPLPPPP) Gag mutation to the plasmid peYFP-HTLV-1 Sense: P5′-CTGCCCCCCCCCCCCAGCTCGCCA-3′ Antisense: P5′-AGGGCGAGAGGGAATCTGCGCCTGGGT-3′	Integrated DNA Technologies BVBA, Leuven, Belgium	N/A

(*Continued on next page*)

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primers for amplifying SNX9 from the plasmid pEGFP-SNX9 to subclone it into the vector pEBB-PP Sense: 5′-ACAGGATATCATGGCCACCAAGGCTCGGGTT ATGTAT-3′ Antisense: 5′-ACAGGCGGCCGCCTACATCACTGGAAAGC GGCTGAGG-3′	Integrated DNA Technologies BVBA, Leuven, Belgium	N/A
Primers for amplifying SNX18 from the plasmid myc-SNX18 to subclone it into the vector pEBB-PP Sense: 5′-ACAGGATATCATGGCGCTGCGCGCCCGG-3′ Antisense: 5′-ACAGGCGGCCGCTTAAACACTATCATAT TTGTGAAGAGCTTCTTCCAA-3′	Integrated DNA Technologies BVBA, Leuven, Belgium	N/A
Primers for amplifying SNX30 from the plasmid pEBB-mCherry-bd-SNX33 to subclone it into the vector pEBB-PP Sense: 5′-ACAGGATATCATGGCACTGAAAGGCCGAGCCCT-3′ Antisense: 5′-ACAGGCGGCCGCTCAGAGGTTGTCATACAT GCGCAGGG-3′	Integrated DNA Technologies BVBA, Leuven, Belgium	N/A
Deposited data		
Free SNX9 SH3	N/A	PDB: 2ENM
c-Src SH3 T98E in complex with VSL12 peptide	(Bacarizo and Camara-Artigas, 2013)	PDB: 4HVW
Yeast Nbp2p SH3 in complex with Ste20 peptide	(Gorelik and Davidson, 2012)	PDB: 2LCS
SNX9 SH3 + EEEV nsP3 peptide complex structure	This paper	PDB: 7OJ9
NMR chemical shifts of the complex	This paper	BMRB: 34628
Experimental models: Cell lines		
HEK293T cells	ATCC	CRL-16268
Recombinant DNA		
plasmid: pET15b	Novagen	Cat#69661; Sigma-Aldrich
plasmid: pET15b-SNX9 SH3	(Karjalainen et al., 2020)	N/A
plasmid: pET15b-c-Src SH3	This paper	N/A
plasmid: pEBB-PP-SNX9	This paper	https://www.addgene.org/22226/
plasmid: peYFP-HTLV-1	A gift from Louis Mansky at the University of Minnesota	N/A
plasmid: pEGFP-SNX9		N/A
plasmid: myc-SNX18	A gift from Sven Carlsson at the Umeå University	N/A
plasmid: pEBB-mCherry-bd-SNX33	(Kleino et al., 2009)	N/A
plasmid: pEBB-PP-SNX18	This paper	https://www.addgene.org/22226/
plasmid: pEBB-PP-SNX33	This paper	https://www.addgene.org/22226/
genomic plasmid: pCRU5-inLuc-mR HTLV-1	A gift from Dimitry Mazurov at the Institute of Gene Biology Russian Academy of Sciences; (Mazurov et al., 2010; Shunaeva et al., 2015)	N/A
packaging plasmid: pCMV-HT1-M	A gift from Dimitry Mazurov at the Institute of Gene Biology Russian Academy of Sciences; (Mazurov et al., 2010; Shunaeva et al., 2015)	N/A
Software and algorithms		
Bruker TopSpin 3.5	Bruker Corporation	https://bruker-labscape.store/products/topspin-for-processing
CcpNmr Analysis 2.4.2	(Vranken et al., 2005)	https://ccpn.ac.uk/software/version-2/

(*Continued on next page*)

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TALOS-N	(Shen and Bax, 2013)	https://spin.niddk.nih.gov/bax/software/TALOS-N/
CYANA 3.98.13	(Güntert and Buchner, 2015)	N/A
AMBER 18	(Case et al., 2018)	https://ambermd.org/GetAmber.php
xcrvfit 5.0.3	Boyko and Sykes, University of Alberta	http://www.bionmr.ualberta.ca/bds/software/xcrvfit/latest/index.html#s_download
Ring NMR Dynamics 1.0.4 (on NMRbox)	(Beckwith et al., 2021)	https://nmrbox.org/
Tensor 2.0	(Dosset et al., 2000)	https://www.ibs.fr/research/scientific-output/software/tensor/
ANSURR v1.2.1	(Fowler et al., 2020)	https://github.com/nickjf/ANSURR
UCSF Chimera 1.15	(Pettersen et al., 2004)	https://www.cgl.ucsf.edu/chimera/download.html
CPPTRAJ 4.25.6 (on AmberTools 20.09)	(Roe and Cheatham, 2013)	https://github.com/Amber-MD/cpptraj
VMD 1.9.3	(Humphrey et al., 1996)	https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD
MicroCal PEAQ-ITC Analysis Software v1.21	Malvern Panalytical Ltd	https://www.malvernpanalytical.com/en/support/product-support/microcal-range/microcal-itc-range/microcal-peaq-itc
MicroCal PEAQ-ITC Control Software v1.21	Malvern Panalytical Ltd	https://www.malvernpanalytical.com/en/support/product-support/microcal-range/microcal-itc-range/microcal-peaq-itc
Image Studio v3.1	LI-COR Biosciences	https://www.licor.com/bio/image-studio/resources
NCBI BLAST	(Altschul et al., 1990)	https://blast.ncbi.nlm.nih.gov/Blast.cgi
EMBL-EBI Clustal Omega	(Madeira et al., 2019)	https://www.ebi.ac.uk/Tools/msa/clustalo/
ConSurf	(Ashkenazy et al., 2016; Landau et al., 2005)	https://consurf.tau.ac.il/
DIALIGN	(Morgenstern, 2004)	https://bibiserv.cebitec.uni-bielefeld.de/dialign/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Perttu Permi (perttu.permi@jyu.fi).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Atomic coordinates for the SNX9 SH3–EEEV nsP3 complex have been deposited in the RCSB Protein Data Bank (PDB) under accession number PDB: 7OJ9. NMR assignments have been deposited to the Biological Magnetic Resonance Data Bank (BMRB) under accession number BMRB: 34628.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial cell culture

The *E. coli* strain BL21(DE3) was transformed with plasmid pET15b-SNX9 SH3 or pET15b-c-Src SH3 (generated previously (Karjalainen et al., 2020)) and inoculated into M9 minimal medium with 100 μg/mL ampicillin. In 1 L culture, 1 g of ^{15}N labelled NH_4Cl and 2 g of uniformly ^{13}C labelled D-glucose were used as a sole nitrogen and carbon sources, respectively. The bacteria were grown at 37°C until OD_{600} reached 0.4 and then temperature was decreased to 16°C. The protein expression was induced by the addition of IPTG to a final concentration of 1 mM at the OD_{600} of 0.6. Cells were harvested 16 h after induction.

METHOD DETAILS

Protein expression and purification

The gene encoding N-terminally His-tagged human SNX9 SH3 (residues 1–64, GenScript Inc, USA) was cloned to pET15b vector (Novagen) into the NdeI and XhoI sites. ^{15}N, ^{13}C labeled or unlabeled SNX9 SH3 was produced by transforming the plasmid into BL21(DE3) cells. Cells were grown in M9 minimal medium supplemented with 1 g/L of $^{15}NH_4Cl$ and 2 g/L ^{13}C-D-glucose as the sole nitrogen and carbon source, respectively, or in LB medium to obtain unlabeled protein. Cell culture was incubated at 37°C until OD of the cell culture reached 0.4. Temperature was then decreased to 16°C and as OD reached 0.6 protein production was induced with 1 mM IPTG. Cells were further incubated at 16°C for 16 h and collected by centrifugation. Cells were disrupted with sonication and the resulting supernatant was clarified by centrifugation with 30000×g. The clarified supernatant was applied to 1-mL His GraviTrap column (GE Healthcare) according to the manufacturer's instructions. Imidazole was removed from eluted protein by PD-10 (Ge Healthcare) before protease cleavage. The His-tag was removed by thrombin protease. Digestion mixture was applied to His GraviTrap column. Cleaved SNX9 SH3 eluted with flow-through was concentrated and applied into Superdex30 (16/60) gel filtration column. Column was equilibrated with NMR buffer (20 mM sodium phosphate pH 6.5, 50 mM NaCl). Fractions, containing purified proteins, were pooled and concentrated by Vivaspin2 (SartoriusStedim). All gel filtrations were performed by using the ÄKTA Purifier FLPC purification system (GE Healthcare). Production and purification of recombinant c-Src SH3 (residues 1–61, GenScript Inc) was performed using a method similar to that used for SNX9 SH3. Synthetic peptides were purchased from GenScript (Piscataway NJ) or CASLO ApS (Kongens Lyngby, Denmark).

NMR spectroscopy

The complex sample composed of 1 mM ^{15}N, ^{13}C labeled SNX9 SH3 and 1 mM unlabeled EEEV nsP3 peptide. The buffer in both samples contained 20 mM NaHPO4 at pH 6.5 and 50 mM NaCl. All spectra were recorded at 25°C on a Bruker AVANCE III HD 800 MHz spectrometer, equipped with a TCI $^{1}H/^{13}C/^{15}N$ cryoprobe.

For chemical shift assignment and structure determination of the SNX9 SH3–EEEV peptide complex, NOESY-^{13}C-HSQC for aliphatic and aromatic ^{13}C regions, NOESY-^{15}N-HSQC, X-filtered, ^{13}C-edited NOESY (SNX9 SH3 → peptide NOESY, mixing time 120 ms), F_1, F_2 ^{15}N, ^{13}C filtered TOCSY (intrapeptide TOCSY, mixing time 80 ms) and X-filtered NOESY (intrapeptide NOESY, mixing time 120 ms), were acquired. SNX9 SH3 chemical shifts were inferred from those in the SNX9 SH3–EspF complex (Tossavainen et al., *to be published*).

NMR data were processed with TopSpin 3.5 (Bruker Inc.) and analysed with CcpNmr Analysis (Vranken et al., 2005) version 2.4.2. Peaks were manually picked from the NOESY spectra. TALOS-N (Shen and Bax, 2013) dihedral restraints for SNX9 SH3 were derived from chemical shifts. CYANA (Güntert and Buchner, 2015) version 3.98.13 iterative automatic NOE peak assignment–structure calculation routine was used to generate 300 complex structures, 30 of which were further refined in explicit water using AMBER 18 (Case et al., 2018). Twenty structures with lowest restraint violations were chosen to represent the complex structures.

NMR titration studies were performed by adding increasing concentrations of peptide into a 0.1 mM SH3 domain sample, and acquiring a ^{1}H, ^{15}N HSQC spectrum at each titration point. The protein to peptide concentration ratios were 1:0, 1:0.5, 1:1, 1:2, 1:4 (SNX9 SH3 + EEEV nsP3), 1:0, 1:0.5, 1:1, 1:3, 1:6, 1:12 (SNX9 SH3+EEEV nsP3 $A^{-1}L$) and 1:0, 1:0.5, 1:1, 1:2, 1:4 (c-Src SH3+EEEV nsP3 $A^{-1}L$). Changes in peak positions between titration points were calculated as $\Delta\delta = (\Delta H^2+(0.154*\Delta N)^2)^{½}$. Curve fitting was performed with xcrvfit (Boyko and Sykes, University of Alberta).

^{15}N T_1, T_2 and heteronuclear NOE spectra were acquired for free SNX9 SH3 and the SH3 – EEEV nsP3 peptide complex. For T_1 relaxation delays of 20, 100, 200, 400 (x2), 600, 900 (x2), 1200, 1500, 1900, and 2400 ms were used. For T_2 the delays were multiples of loop length 16.96 ms, with the loop counter set to 1, 2, 3 (x2), 5, 6, 8 (x2), 9, 11, 13, and 16. Two time points were acquired in duplicate to allow estimation of error. The recycle delay was set to 2.5 s in the T_1 and T_2 experiments, and to 10 s in the hetNOE experiment. All spectra were acquired as pseudo 3D spectra and Fourier transformed with TopSpin. T_1, T_2 peak intensities were fitted to decaying exponential function as implemented in Ring NMR Dynamics (Beckwith et al., 2021), using NMRbox (Maciejewski et al., 2017). Heteronuclear NOE values were calculated as the intensity ratios of peaks from a pair of spectra measured with and without ^{1}H presaturation during the recycle delay as implemented in CcpNmr Analysis. The internal model-free parameters S^2, τ_e and R_{ex} were derived from the ^{15}N relaxation data using TENSOR2 (Dosset et al., 2000).

H/D exchange experiments were performed by first preparing the free SNX9 SH3 or its EEEV nsP3 peptide complex sample in H_2O-containing buffer and recording a reference ^{1}H, ^{15}N HSQC spectrum. Then the samples were lyophilized, and re-dissolved into D_2O.

^{1}H, ^{15}N HSQC spectra were then acquired at regular intervals up to ~1 h for the free, by the time all signals had disappeared, and up to ~22 h for the complex form. An aliphatic ^{1}H, ^{13}C HSQC acquired at the end of the study showed that both protein samples were intact after lyophilization.

MD simulations

Atomistic molecular dynamics simulations were performed as follows: MD simulations in explicit solvent were performed with AMBER 18 using the ff14SB force field. Ensemble structure 9, the best structure according to ANSURR (Fowler et al., 2020) was selected as the target of simulation of the wt complex, the A^{-1}L complex was created by swapping A^{-1} with a leucine in UCSF Chimera. The complexes were placed in a cubic box with a minimum solute-box distance of 10 Å, and solvated with TIP3P water molecules. Two Na^{+} ions were added to neutralize the system. After minimization, heating and equilibration of the system, the production 100-ns MD simulations were performed with periodic boundary conditions at 300 K. The temperature was maintained by using the Langevin thermostat, whereas the pressure was kept at 1 bar using the Berendsen barostat (Berendsen et al., 1984). The time step was set to 2 fs. Long-range electrostatic interactions were treated using the Particle Mesh Ewald method (Darden et al., 1993) with a cut-off of 10 Å. Bond lengths involving hydrogen atoms were constrained by SHAKE (Ryckaert et al., 1977). Analyses of the trajectories were carried out with CPPTRAJ (Roe and Cheatham, 2013), VMD (Humphrey et al., 1996) and UCSF Chimera (Pettersen et al., 2004).**ITC**

Synthetic peptides were prepared by first dissolving them to purified water (Milli-Q) and adjusting solution pH to 6.5. Next, the peptides were lyophilized and redissolved in 20 mM sodium phosphate, 50 mM NaCl, pH 6.5 buffer. Peptide concentrations varied between 0.5 and 2.0 mM. Peptides were titrated to sample cell containing SNX9 SH3 in 0.03–0.14 mM concentrations. The experiments were performed at 25°C using MicroCal PEAQ-ITC isothermal titration calorimeter (Malvern Instruments Ltd., UK). Experiments were performedtwo or three times. Parameter averages were calculated of all measurements. Data were analysed with MicroCal PEAQ-ITC Analysis software (Malvern Instruments Ltd., UK). The one set of sites fitting model was used and calculated constant control heat (Fitted Offset) was used as control.

Cells and plasmids

HEK293T cells were maintained in Dulbecco's Modified Eagle Medium (Sigma: D6546) supplemented with 4mM Ala-Gln (Sigma: G8541) and 10% Fetal Bovine Serum (Thermo Fisher Scientific: 10270-106). Cell cultures were incubated in +37°C with 5% CO2.

Template plasmids for SNX9 and SNX33 have been described previously (Kleino et al., 2009). SNX18 was a gift from Sven Carlsson (Søreng et al., 2018) and peYFP-HTLV1-coGAG producing codon optimized HTLV-1 GAG from Louis Mansky (Grigsby et al., 2010). SNX9, SNX18 and SNX33 were cloned into pEBB-PP plasmid and A to L and AxxA mutations were introduced by site-directed mutagenesis. pCRU5-inLuc-mR replication dependent HTLV-1 genomic plasmid and pCMV-HT1-M packaging plasmid were gifts from Dimitry Mazurov (Institute of Gene Biology Russian Academy of Sciences) and have been described previously (Mazurov et al., 2010; Shunaeva et al., 2015). pCMV-HT1-M A to L and AxxA mutations were generated from synthetic DNA fragments (Integrated DNA Technologies) and cloned into Not1 and Ade1 restriction sites of pCMV-HT1-M.

pCRU5-inLuc-mR plasmid contains a truncated HTLV-1 genome containing a firefly luciferase gene that is not expressed in cells transfected to produce HTLV-1 pseudoviruses and requires one round of reverse transcription to produce luciferase. This is because the luciferase gene in pCRU5-inLuc-mR has been placed in the antisense orientation and is disrupted by an intron in the sense orientation. Therefore, an open luciferase reading frame is present only in the new provirus in the infected cells generated via reverse transcription of the spliced mRNA product of pCRU5-inLuc-mR. This is important because HTLV-1 infection takes place mainly via cell-to-cell contacts, and HTLV-1 gene expression in the infected cells must be distinguished from expression of the transfected plasmids in the co-cultured HTLV-1 pseudovirus-producing cells.

pCMV-HT1-M packaging plasmid for HTLV-1 retroviral vector system expresses HTLV-1 genes from a CMV-enhancer, and is used to generate HTLV-1 pseudoviruses by co-transfecting it into producer cells together with a a truncated HTLV-1 proviral plasmid, such pCRU5-inLuc-mR. In order to generate "A to L" and "AxxA" mutations, synthetic gene fragments were purchased from Integrated DNA Technologies, Inc, USA, cloned between the Ade1 and Not1 restriction enzymes sites in pCMV-HT1-M.

Co-precipitation studies

3 × 10^{6} HEK293T cells were seeded on 10 cm plates one day before transfection. 4 μg peYFP-HTLV1-coGAG and 2 μg pEBB plasmid expressing PP-tagged SNX9, SNX18, or SNX33 were transfected using 18 μg PEI. 48 h after transfection cells were washed with PBS and lysed in 1 mL ice-cold lysis buffer (50 mM Tris [pH:7.4], 150 mM NaCl, 1 mM EDTA and %1 NP-40) containing phosphatase and protease inhibitors (Thermo Fisher Scientific, A32957 and A32963 respectively). After a 30 min incubation the lysates were centrifuged at 16 000 rcf at +4 for 20 min. Supernatants were mixed with 70 μg M−280 Streptavidin Dynabeads (Invitrogen, 11205D) and incubated at +4 with rotation. After 3–4 h of incubation beads were washed with lysis buffer 3 times and mixed with SDS-Loading buffer. Samples were analyzed with standard western blotting. Anti-HTLV-1 p24 antibody (Abcam ab9081) and goat anti mouse IRDye 680RD (LI-COR P/N: 925–68070) used for detecting GAG while IRDye 800CW Streptavidin (LI-COR P/N: 926–32230) was used to detect SNXs. GAPDH was detected with antibody from Cell Signaling Technology (#2118) using goat anti rabbit IRDye 800CW (LI-COR P/N: 926–32211) as a secondary antibody. Odyssey infrared imaging system and Image Studio v3.1 software (LI-COR) were employed to visualize the blots.

Please cite this article in press as: Tossavainen et al., Structure of SNX9 SH3 in complex with a viral ligand reveals the molecular basis of its unique specificity for alanine-containing class I SH3 motifs, Structure (2022), https://doi.org/10.1016/j.str.2022.03.006

HTLV-1 viral transmission

5×10^5 HEK293T cells were seeded on 6-well plates one day before transfection. For generating HTLV-1 viral vector expressing firefly Luciferase 500 ng of pCMV-HT1-M plasmid expressing HTLV-1 viral proteins and 1.5 μg pCRU5-inLuc-mR genomic plasmid were transfected using 6 μg PEI. 48 h after transfection the cells were washed with PBS and lysed with lysis reagent (Promega: E153A). Samples were analyzed with Firefly Luciferase Assay Kit (Promega: E1501). Expressions of viral proteins were analyzed by Western blotting using Anti-HTLV-1 p24 antibody as described above.

Image Studio v3.1 software (Li-Cor Biosciences) were employed to visualize the blots.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis details are found in the methods description and figure captions.