Published in final edited form as:

Structure. 2012 October 10; 20(10): 1692–1703. doi:10.1016/j.str.2012.07.015.

Enterohaemorrhagic *E. coli* (EHEC) exploits a tryptophan switch to hijack host F-actin assembly

Olli Aitio¹, Maarit Hellman¹, Brian Skehan², Tapio Kesti³, John M. Leong^{2,4}, Kalle Saksela³, and Perttu Permi^{1,*}

¹Program in Structural Biology and Biophysics, Institute of Biotechnology, University of Helsinki, FI-00014, Helsinki, Finland ²Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, MA 01655 ³Department of Virology, Haartman Institute, University of Helsinki and HUSLAB, University of Helsinki Central Hospital, FI-00014, Helsinki, Finland

SUMMARY

Intrinsically disordered protein (IDP)-mediated interactions are often characterized by low affinity but high specificity. These traits are essential in signaling and regulation that require reversibility. Enterohaemorrhagic *Escherichia coli* (EHEC) exploit this situation by commandeering host cytoskeletal signaling to stimulate actin assembly beneath bound bacteria, generating ‘pedestals’ that promote intestinal colonization. EHEC translocates into the host cell two proteins, EspF_U and Tir, which form a complex with the host protein IRTKS. The interaction of this complex with N-WASP triggers localized actin polymerization. We show that EspF_U is an IDP that contains a transiently α -helical N-terminus and dynamic C-terminus. Our structure shows that single EspF_U repeat is capable of forming a high-affinity trimolecular complex with N-WASP and IRTKS. We demonstrate that bacterial and cellular ligands interact with IRTKS SH3 in a similar fashion but the bacterial protein has evolved to outcompete cellular targets by utilizing a tryptophan switch that offers superior binding affinity enabling EHEC-induced pedestal formation.

INTRODUCTION

Intrinsically disordered proteins (IDPs) are ubiquitous proteins that are often involved in cell signaling. They do not possess a folded tertiary structure in native state, and typically rely on short motifs and transient but specific interactions to carry out their function (Vacic et al., 2007; Hazy & Tompa, 2009; Uversky, 2010; Babu et al., 2011). Disordered proteins bind to their targets at the expense of reduction in conformational entropy, which enables combining high specificity with modest affinity, and thus renders such interactions suitable for processes destined to be reversible (Dyson & Wright, 2005; Mittag et al., 2009). The high degree of regulation, typical for cellular processes, can be considered as an Achilles heel of these fine-tuned interactions and pathogens have evolved to exploit this vulnerability (Babu

© 2012 Elsevier Inc. All rights reserved.

*Corresponding author: Dr. Perttu Permi, Institute of Biotechnology/NMR Laboratory, University of Helsinki, P.O. Box 65, FI-00014, FINLAND, Tel: +358-9-191-589-40, Fax: +358-9-191-595-41, Perttu.Permi@helsinki.fi.

⁴Current address: Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston MA 02114

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

et al., 2011; Davey et al., 2010). Indeed, a common approach of pathogens is to copy a host protein's functionality and to produce mimetics of higher affinity (Davey et al., 2010).

EspF_U (also known as TccP) is a translocated bacterial effector enterohaemorrhagic *Escherichia coli* (EHEC) serotype O157:H7 that promotes the formation of actin 'pedestals' on mammalian cells beneath bound bacteria (Campellone et al., 2004; Garmendia et al., 2004; Campellone, 2010). To generate pedestals, EspF_U becomes localized in the host cell at sites of bacterial attachment, where it activates actin assembly. EspF_U is a 337-residue protein composed of an N-terminal sequence that promotes EspF_U translocation into the host cell via a bacterial type III secretion system, followed by multiple 47-residue highly conserved consecutive repeats (Campellone et al., 2004; Garmendia et al., 2004 & 2006) that possess dual activities. The N-terminal 20 residues of the repeat bind to the GBD (GTPase binding domain of neuronal Wiskott-Aldrich syndrome protein) of WASP/N-WASP, members of a family of nucleation promoting factors that regulate a central pathway of actin assembly. The EspF_U-GBD interaction disrupts an autoinhibitory interaction between the GBD and the WH2 (C-terminal WASP homology 2)/VCA (verpolin-connector-acidic) (Cheng et al., 2008; Sallee et al., 2008). In turn, this activated WASP/N-WASP stimulates the Arp2/3 actin nucleator complex. The 22 C-terminal residues of the EspF_U repeat contain a proline-rich sequence that binds to the SH3 (Src homology-3 domain) of IRTKS (insulin receptor tyrosine kinase substrate) or the related IRSp53 (Cheng et al., 2008; Sallee et al., 2008; Weiss et al., 2009; Vingadassalom et al., 2009; Aitio et al., 2010) (Fig. 1A). IRTKS and IRSp53 bind to a cytoplasmic sequence of EHEC effector protein Tir, which after translocation into host cells is localized at sites of bacterial attachment (Vingadassalom et al., 2009). Thus, EspF_U binding of the IRTKS/IRSp53 SH3 domain results in recruitment of the EspF_U:N-WASP:Arp2/3 complex and localized actin assembly.

We have recently shown that the IRTKS SH3 - EspF_U complex establishes a novel type of SH3 interaction that involves accommodation of two adjacent polyproline II (PPII) helical PxxP motifs by a single SH3 domain, representing one of the highest affinity SH3 interactions currently known (Aitio et al., 2010). Of note, a similar arrangement of tandem PxxP motifs is also found in the cellular ligands of the IRTKS/IRSp53 family SH3 domains suggesting that this interaction is evolutionary conserved (Aitio et al., 2010). While the mechanism of opening the autoinhibitory lock of N-WASP by EspF_U is well understood (Cheng et al., 2008; Sallee et al., 2008), the functional hijacking of IRTKS/IRSp53 SH3 by EspF_U has remained elusive

In this work we have used bioinformatics and biophysical tools, e.g. NMR (Nuclear Magnetic Resonance) spectroscopy and ITC (isothermal titration calorimetry) for structural characterization of EspF_U. We show that EspF_U is disordered in its native state. However, the N-WASP GBD binding domain in the N-terminus of EspF_U transiently populates an α -helical conformation, whereas the proline-rich IRTKS SH3 binding motif establishes a highly dynamic polypeptide. We also show that EspF_U undergoes disorder-to-order transition upon formation of a trimolecular complex with IRTKS SH3 and N-WASP GBD. Most importantly, we reveal the underlying structural mechanism by which EspF_U outcompetes cellular IRTKS SH3 binding ligands and firmly interconnects actin polymerization and membrane regulation machineries.

RESULTS

Structural and dynamical characterization of EspF_U R47₅ free in solution

We carried out sequence analysis of each repeat using several bioinformatics tools (e.g. IUPred, Disprot, PSIPred) available for predicting disordered regions in proteins based on their amino acid sequence. All these analyses suggested that repeats are disordered and

EspF_U belongs to a class of intrinsically disordered proteins, IDPs (Fig. 1B). Further analysis was carried out using NMR that has been shown to be an excellent tool for characterization of IDPs (Muckrasch et al., 2009; Hellman et al., 2011). Instead of well-dispersed ¹⁵N-¹H correlation spectrum of folded proteins, the ¹⁵N-HSQC spectrum of EspF_U R47₅ displayed a poorly dispersed correlation map reminiscent of disordered polypeptide chain (Fig. 1C). For more detailed characterization of EspF_U, we first carried out the assignment of main-chain ¹H, ¹³C, and ¹⁵N chemical shifts in EspF_U using the suite of H_α detected experiments that are very useful for proline-rich IDPs (Mäntylähti et al., 2010 & 2011). The ¹³C_α secondary chemical shift (SCS) is a reliable indicator of residual secondary structure in the polypeptide (Wishart et al., 1995; Kjaergaard & Poulsen, 2011). The N-terminal segment ³DVAQRLMQHLAEH¹⁵ shows clearly positive ¹³C_α SCSs, up to 1.3 ppm (Fig. 1D). This indicates that these residues fractionally populate α-helical conformation, up to 26.2% based on the secondary structure propensity score (Marsh et al., 2006). In contrast, the C-terminal part (residues 17–47), which includes the proline-rich segment ²⁷IPPAPNWPAPTPP³⁹ that harbors the tandem PxxP motifs responsible IRTKS SH3 binding, is highly disordered.

Further evidence of transient structural elements was gleaned by {¹H}-¹⁵N NOE data as well as ¹⁵N *T*₂ and *T*₁ relaxation times, which are reporters of ps-ns timescale dynamics (Supplementary Fig. S1). Spectral density mapping method was used for quantitative analysis of relaxation data (Farrow et al., 1995; Lefevre et al., 1996). Dynamics at three different frequencies, *J*(0), *J*(ω_N) and *J*(0.87ω_H) underscores increased rigidity for residues ⁵AQRLMQHLAEH¹⁵ that correspond to the transient α-helical region of R47₅ (Fig. 1E). Dissection of motional fluctuations in the proline-rich region pinpoints more distinctive local features. For both ²⁷IPPAP³¹ and ³⁵APTPP³⁹ (XPxXP) motifs significant contribution of high frequency motions was observed, *J*(0.87ω_H) ~ 20–27 ps/rad, indicating highly flexible polypeptide in this region. In contrast, the linker ³²NWP³⁴ between the PxxP motifs as well as the linker ²²NMAEH²⁶ that interconnects the N-terminal N-WASP GBD binding segment to the proline-rich region displays more restricted backbone mobility, *J*(0.87ω_H) ~ 10–15 ps/rad. The very C-terminal part of the R47₅ repeat is highly flexible with elevated supra-ps timescale dynamics as manifested by *J*(0.87ω_H) values up to 40 ps/rad. Altogether these data suggest that the tandem PxxP elements exhibit elevated local dynamics in ns-ps timescale when compared to their flanking regions.

A final point of interest concerns proline *cis-trans* isomerization, which has earlier been shown to play important role in signaling (Sarkar et al., 2011). We detected a second set of resonances corresponding to *cis* isomer of P34 in the linker between PxxP motifs as well as to the N- and C-terminal P2 and P47. The *cis* isomer content of P34 is approximately 30%. Kinetics of *cis-trans* isomerization was studied using the ¹⁵N exchange spectroscopy, but as no cross peaks between *cis* and *trans* conformers was observed, this process is likely to be very slow (*k*_{ex} < ¹⁵N *R*₁ ≈ 1 s). Taken together, while the N-terminal segment of free EspF_U R47₅ fractionally pre-exists in its bound conformation, it is unlikely that the proline-rich segments exist in a PPII conformation. Further support for this interpretation was obtained in terms of residual dipolar couplings (RDCs), which indicated α-helical tendency for the N-terminal residues of EspF_U, whereas the ideal PPII conformation found in the EspF_U:SH3 complex was clearly absent in the free EspF_U (Supplementary Fig. S2).

Trimolecular complex between EspF_U R47₅, N-WASP GBD and IRTKS SH3

To understand the structural details of the interaction of EspF_U R47 with N-WASP GBD and IRTKS SH3, we determined the structure of the ternary complex between N-WASP GBD, EspF_U R47₅ and IRTKS SH3 (Fig. 2A). It is composed of binary complexes between the N-terminal EspF_U and N-WASP GBD and the C-terminal EspF_U and IRTKS SH3, which are connected by a short six amino acid linker. Superposition of the N-WASP GBD

and the N-terminal R47₅ or IRTKS SH3 and the C-terminal R47₅ confirm that individual subunits of the complex are very well determined (Figure 2B and Table 1). The binary complexes are essentially similar to those reported previously (Cheng et al., 2008; Aitio et al., 2010), but subtle differences between the N-WASP:EspF_U and the WASP:EspF_U complexes (Cheng et al., 2008) can be recognized. This may arise from several differences between N-WASP/WASP GBD residues facing the EspF_U binding site. The GBD domain in both structures is highly similar (RMSD 1.02 Å for residues 216–262), but the length of α -helix in EspF_U and its orientation with respect to the GBD domain as well as the positioning of the extended arm deviate.

In N-WASP:EspF_U complex the EspF_U α -helix begins at P502, which is confirmed by the α -helix characteristic $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha N}(i,i+3)$ NOEs detected between P502 and A505 (making it 4 Å longer). P502 also makes hydrophobic contacts with M238 of the N-WASP GBD helix 1. Instead of a methionine WASP GBD has an arginine at this position, and the next residue is alanine contrary to the N-WASP cysteine 239. The hydrophobic contacts with P502 pull the N-terminus of EspF_U α -helix towards the N-WASP GBD helix 1, and on the other hand the larger van der Waals radius of C539, as compared to alanine in WASP, pushes it away. As a consequence the EspF_U α -helix makes a 12° angle with its WASP complex counterpart (Fig. 2C).

Despite subtle chemical shift perturbations observed in ¹⁵N-HSQC spectra of EspF_U R47₅ between binary and ternary complexes (Fig. 2D), the structure suggests that the N- and C-terminal regions of EspF_U function as independent units. To rule out putative reciprocal orientation between N-WASP GBD and IRTKS SH3 when bound to R47₅, we employed ¹⁵N-¹H RDCs measured in ¹⁵N-N-WASP GBD:EspF_U R47₅: ¹⁵N/¹³C-IRTKS SH3 complex (Supplementary Fig. S3). A simple isotropic motional model based on analysis of generalized degree of order for N-WASP GBD and IRTKS SH3 domains in the ternary complex indicated large amplitude interdomain motion up to $\Psi_{\text{cone}} = 75^\circ$ (Tolman et al., 2001).

Therefore, the simultaneous interaction of EspF_U with N-WASP GBD and IRTKS SH3 induces neither additional folding of EspF_U nor conformational changes within GBD and SH3 domains or interactions between them. The flexible linker is likely to have a role in the assembly and correct positioning of the domains during the formation of a multiprotein complex. This structure highlights several characteristic functional features of IDPs. IDPs often use for recognition short linear motifs that undergo disorder-to-order transition upon binding. The residual structure observed for the N-terminal part of EspF_U is implicated in molecular recognition. Also a considerable amount of disorder is maintained in the bound state.

To investigate the rigidity of EspF_U R47₅ in the complex, we measured heteronuclear steady state {¹H}-¹⁵N NOEs, and compared them with values measured from the unbound EspF_U R47₅. {¹H}-¹⁵N NOE plots for the ¹⁵N, ¹³C-labeled EspF_U R47₅ when bound in the ternary complex or free in solution are shown overlaid in Fig. 2E. Clearly, the N-WASP GBD binding epitope, encompassing residues ³Asp-Arg²¹, became more rigid upon binding to N-WASP as reported by increased heteronuclear NOEs > 0.7), whereas extensive disorder-to-order transition, which translates into substantial increase of heteronuclear NOEs from negative to large positive values, could be observed for the IRTKS SH3 binding region (²⁷IPPAPNWPAPTPP³⁹). The short linker region corresponding to residues 22–24 exhibits somewhat lower hetNOE values (0.3–0.6), but indicates stiffening of the mediating linker. Nevertheless, as confirmed by the RDC data (Supplementary Fig. S4), the linker still enables large amplitude motion of subunits and hence conformational readjustment of polypeptide upon binding to multiple targets. Similar motional averaging has been reported

for MBP_{145–165} - CaM complex (Nagulapalli et al., 2012). Also the very C-terminal part of R47₅, corresponding to residues 41–47 remains flexible in the complex as manifested by low positive or negative $\{^1\text{H}\}\text{-}^{15}\text{N}$ NOEs.

Finally, of the three prolines (P2, P34 and P47) that populate *cis/trans* conformations in free EspF_U, only P47 is found in equilibrium of *cis/trans* conformations, whereas P2 and P34 exist solely in *trans* conformation in the complex, hence undergoing conformational change upon binding to N-WASP GBD and IRTKS SH3, respectively

Thermodynamical characterization of N-WASP GBD - EspF_U interaction

IDP mediated molecular interactions are often characterized with low affinity but high specificity, which offer functional advantage over folded proteins e.g., by enabling association with multiple partners. Weak association stems from entropic cost for the Gibbs free energy (ΔG_{free}) as IDPs often undergo disorder-to-order transition upon binding. We utilized ITC data to glean information on nature and characteristics of EspF_U R47₅ - N-WASP GBD interaction (Table 2). These data show that interaction of EspF_U R47₅ with the GBD domain is strongly enthalpy driven ($\Delta H = -64.6$ kJ/mol) and counterbalanced with unfavorable entropy ($-\text{T}\Delta S = 23.7$ kJ/mol). Upon formation of ternary complex i.e. N-WASP GBD binding to EspF_U R47₅ - IRTKS SH3 complex, similar values are obtained, suggesting non-cooperative binding model (Table 2). Although the measured dissociation constants, K_{d} s $\sim 40\text{--}70$ nM, are similar to the value of 35 nM reported earlier between C-terminal region of N-WASP_C (residues 193–501 of N-WASP) and EspF_U R47₅ (Cheng et al., 2008), their thermodynamic fingerprints are very different, $\Delta H = -28.3$ kJ/mol and $-\text{T}\Delta S = -14.3$ kJ/mol. In case of N-WASP GBD - EspF_U R47₅ interaction, two disordered polypeptides undergo disorder-to-order transition upon binding, resulting in large entropic cost. In contrast, EspF_U R47₅ interaction with N-WASP_C involves disruption of its auto-inhibited conformation i.e. the process that includes large order-to-disorder transition, and yields entropically favorable binding.

A tryptophan in EspF_U linker is critical for high affinity

Similar to the pathogen protein EspF_U, the identified cellular ligands of IRTKS/IRSp53 SH3, namely Shank 1–3 and Eps8 also contain a tandem PxxP motif. However, a pathogen protein would be expected to have a higher SH3 binding affinity to displace cellular ligands in order to hijack IRTKS/IRSp53-mediated signaling pathways of the host cell. To test this assumption we compared the binding affinities obtained using NMR CSP (chemical shift perturbation) mapping and ITC of EspF_U R47₅ and a 20-residue peptide derived from Eps8 containing the tandem PxxP IRTKS recognition motif (see Table 3). The CSP mapping showed that these peptides bound to the same binding site on IRTKS SH3 as peptide addition induced chemical shift changes for the same set of NH correlations in the ^{15}N -HSQC spectrum of ^{15}N -labeled IRTKS SH3 (Fig. 3, upper panels). However, the binding affinities were different. EspF_U was clearly a strong binder as two sets of peaks were observed in subequimolar concentrations, while CSPs observed for the Eps8 peptide were in the regime of intermediate exchange in the NMR timescale, indicating lower affinity. In agreement with CSP, we obtained dissociation constants K_{d} (EspF_U) = 0.5 μM (Aitio et al., 2010) vs. K_{d} (Eps8) = 30.5 μM , using ITC. These data showed that the pathogenic EspF_U is able to usurp this host cell signaling pathway by superseding the cellular ligand with approximately 60 times higher affinity against its target.

As both EspF_U and Eps8 peptides carry a tandem PxxP motif some additional factor is needed to explain the substantially stronger binding of EspF_U. Comparison of EspF_U sequence with those of known and predicted cellular partners of IRTKS/IRSp53 SH3 demonstrates that only EspF_U contains a tryptophan residue in the linker between the two

PxxP motifs (Fig. 4A). We thus investigated whether this tryptophan could explain the enhanced affinity acquired by the pathogen. To this end, we made two peptides that carry mutations in position 33 according to EspF_U R47₅ numbering: An EspF_U peptide in which the tryptophan in ³²NWP³⁴ was replaced by alanine (yielding ³²NAP³⁴) and an Eps8 peptide in which the linker alanine in ³²RAP³⁴ was replaced by tryptophan (yielding ³²RWP³⁴). Again CSP mapping indicated that these peptides interacted with IRTKS SH3 through the same interface as the EspF_U and Eps8 wt peptides (Fig. 3, lower panels). W to A replacement in EspF_U peptide reduced the affinity substantially, K_d (EspF_U^{W33A}) = 22.3 μ M, and strikingly, the A to W replacement converted the Eps8 peptide into a strong binder K_d (Eps8^{A33W}) = 2.4 μ M. This clearly pinpointed the critical role of the linker tryptophan for high affinity, thus explaining the higher affinity of EspF_U as compared to the cellular Eps8 ligand.

In summary, the W33A mutation in EspF_U reduced the binding affinity by 50-fold, and the reciprocal A-to-W mutant of Eps8 bound 13-fold more tightly to IRTKS SH3 than the wt Eps8. Our IRTKS SH3:EspF_U R47₅ complex structure shows that EspF_U tryptophan W33 lies in a T-shaped edge-to-face arrangement above W378 of SH3 domain (Fig. 4B). EspF_U tryptophan also makes an intramolecular van der Waals contact with P31 in the bound form. These interactions are not possible with an alanine at this position. The W to A mutation in the Eps8 peptide provides the same inter- and intramolecular contacts when bound to IRTKS SH3 domain in the same overall conformation as EspF_U. We conclude that the intermolecular aromatic interaction between W33 and W378 as well as the intramolecular contacts stabilize the association and contribute to the higher affinity of EspF_U and Eps8^{A33W} mutant with respect to their lower affinity ligands. It is noteworthy that this is the first observation of a π - π interaction between two tryptophan residues involving an IDP.

Thermodynamics of binding of EspF_U and Eps8 peptides with IRTKS SH3

SH3-ligand interactions are typically characterized by a favorable ΔH contribution to ΔG counterbalanced by an unfavorable entropic penalty ($-T\Delta S$), which seems surprising considering the large hydrophobic interaction surface involved (Palencia et al., 2004; Wang et al., 2001). Similar to other SH3 ligand pairs, the IRTKS SH3 – ligand interactions are dominated by an enthalpic contribution (ΔH) (See Table 3). Enthalpy-driven hydrophobic complexation arises from poor solvation of the binding surface in unbound state (Bissantz et al., 2010), and the enthalpy gain results from stronger hydrogen bonds formed between water molecules released from the surface of the protein upon binding (Meyer et al., 2003). It is noteworthy that entropy change ($-T\Delta S$) contributes favorably to wt Eps8 (−9.1 kJ/mol), Eps8^{A33W} (−5.6 kJ/mol), and EspF_U^{W33A} (−16.3 kJ/mol) binding that according to our CSP mapping experiments interact with IRTKS SH3 in the similar manner as the wt EspF_U (see Fig. 3). In these peptides IRTKS SH3 recognizes two consecutive proline-rich motifs connected by a three-residue linker. The linker interacts, without making any polar contacts, with the specificity pocket, thus rendering the interaction particularly hydrophobic. This implies that although classical SH3 ligand interactions have predominantly hydrophobic character, even larger hydrophobic surfaces such as in the case of Eps8:IRTKS SH3 are required for a favorable entropic contribution.

However, binding is entropically unfavorable for the fifth EspF_U repeat (R47₅), as well as for the 20-residue peptide from the second EspF_U repeat (referred to as EspF_U R₂), which show the highest affinities for IRTKS SH3. Comparison of thermodynamic data shows that for both EspF_U R47₅ and Eps8 peptides changing of A to W increases affinity, which is accompanied by a gain in enthalpy and loss in entropy. The mechanisms underlying this frequently observed entropy-enthalpy compensation are not well understood. However, it is likely that while additional ligand-protein interactions provide a net gain in enthalpy, increased rigidity in a high affinity complex, i.e. reduction of backbone motion induced by

ligand binding, translates into decrease in entropy (Wang et al., 2003; Bissantz et al., 2010; Williams et al., 2004; Frederick et al., 2007). In addition, a specific edge-to-face orientation of tryptophan residues is likely to increase entropic cost of binding (Tatko & Waters, 2002).

Interestingly, Eps8^{A33W} and EspF_U R47₅ (or EspF_U R₂) have considerably different thermodynamic signatures, although both peptides contain the tandem PxxP motifs and the critical tryptophan in the linker. There is only a single amino acid difference in residues in direct contact with the SH3 domain. Eps8 as well as other cellular ligands have a proline instead of alanine at position 30 (according to EspF_U numbering). Both interactions are enthalpy-driven, but the change in entropy contributes unfavorably to EspF_U R47₅ binding, whereas it contributes favorably to Eps8^{A33W} binding. This might relate to higher *cis/trans* population of Eps8 ligands, but it is difficult to explain how these subtle structural differences between EspF_U and Eps8^{A33W} translate into difference in relative ΔH and $T\Delta S$ contributions (Bissantz et al., 2010).

W to A mutation disrupts the recruitment of EspF_U to sites of bacterial attachment

It has been shown previously that IRTKS (or IRSp53) binding is an essential activity for recruitment of EspF_U to the sites of clustered Tir (Weiss et al., 2009; Vingadassalom et al., 2009; Aitio et al., 2010). To test whether increased binding affinity and stability of the IRTKS:EspF_U complex, mediated by the tryptophan π - π interaction, translates into functional importance upon EHEC infection, we investigated the role of the W33A mutation in pedestal formation.

First, a proline-rich sequence of the EspF_U repeat that is recognized by the SH3 domain of IRTKS/IRSp53 has been identified (Weiss et al., 2009; Aitio et al., 2010), and deletion of the EspF_U repeat C-terminal 14 residues, which interrupts this sequence, blocked binding by IRTKS in yeast two-hybrid assays (Vingadassalom et al., 2009). Therefore, we designed an EspF_U construct that contained only the 23 residue proline-rich region ("P" in Fig. 4C) fused to the Gal4AD and co-transformed the reporter strain L40 with LexADB-IRTKS-SH3. The "P" construct closely approximated the reporter activity of a full repeat "HP" (Fig. 4C). Next, we tested alanine substitution of residue W33 in a yeast two-hybrid assay, and found that W to A mutation in "P_{W33A}" construct disrupted binding to IRTKS-SH3, and completely abrogated any activity in this assay (Fig. 4C).

To further explore the role of tryptophan switch in IRTKS recruitment and actin assembly, we generated HP*HP*, a simplified two-repeat EspF_U construct in which both of the two repeats contained the W33A mutation, and ectopically expressed GFP-HPHP or GFP-HP*HP* with wild type or mutant repeats, respectively, in mouse embryonic fibroblasts (MEFs). When MEFs expressing wild type GFP-HPHP fusion construct were infected with KC12, a modified *E. coli* capable of translocating Tir but not expressing EspF_U, the ectopically expressed GFP-HPHP was recruited to the sites of bacterial attachment, as shown by immunostaining the myc-tagged GFP-HPHP fusion, and induced actin pedestal formation (Fig. 4D). In contrast, the IRTKS binding-deficient mutant GFP-HP*HP* failed to be recruited to the sites of bacterial attachments, and consequently no actin assembly was observed, suggesting that IRTKS binding was an essential activity for EspF_U recruitment to sites of clustered Tir. To verify whether the defect of GFP-HP*HP* in actin assembly was due solely to its inability to be recruited to Tir, we tested whether the HP*HP* derivative of EspF_U could stimulate pedestal formation when artificially clustered by translational fusion to Tir. To this end, we devised an Tir Δ C-HP*HP*, in which the TirC-terminal cytoplasmic domain was replaced by the two repeat HP*HP* sequence, which lacks the ability to bind IRTKS due to W33A mutation. When Tir Δ C-HP*HP* clustered in the plasma membrane by infection with *E. coli* expressing intimin, robust actin pedestals were observed, indicating that the critical actin assembly defect in GFP-HP*HP* was due to its inability to be recruited

to Tir by IRTKS (Fig. 4E). Thus, we conclude that the enhanced IRTKS binding affinity provided by W33 plays a critical role in pedestal formation.

An engineered tryptophan switch promotes intracellular association of IRTKS with Eps8

Our results indicated the tryptophan residue 33 in the tripeptide linker between the two PxxP motifs of EspF_U provided it with a superior binding affinity compared to cellular ligands of IRTKS SH3. To further validate this concept we tested if binding to IRTKS by its cellular interaction partner Eps8 could be increased by introducing an EspF_U-like W-containing inter-PxxP linker into Eps8. As showed in Fig. 4F this prediction indeed turned out to be correct. A dramatic increase in co-precipitation of Eps8 with IRTKS was observed in 293 cells transfected with the linker-modified Eps8 compared to wild-type Eps8. Thus, the tryptophan switch is not only critical for the pedestal formation, as demonstrated by the loss-of-function phenotype in the bacterial infection experiment shown in Fig. 4F, but can also be used to engineer a gain-of-function mutant of a cellular ligand that otherwise binds to IRTKS with a modest affinity.

DISCUSSION

In this work we have shown that the bacterial effector EspF_U is an intrinsically disordered protein. It contains two protein recognition motifs, which undergo disorder-to-order transition upon binding. Although molecular interactions involving IDPs are typically weak and transient due to their regulatory roles in cellular processes, EspF_U is able to establish a tight complex with two host proteins, N-WASP GBD and IRTKS SH3. Our study highlights that a tertiary structure is not a prerequisite for tight interactions, and pathogens are able to use bacterial IDPs to commandeer tightly regulated cellular processes. The N-WASP GBD binding (“H”) motif is clearly disordered when free in solution although it is significantly more rigid than the IRTKS SH3 binding (“P”) region of EspF_U. Indeed, our data show that the H motif transiently pre-exists in its bound conformation while the left-handed PPII conformation is clearly absent in the P motif when free in solution. The H motif then falls in the category of preformed structural element or molecular recognition element (Fuxreiter et al., 2004; Oldfield et al., 2005) although very high affinity complex is established with N-WASP, atypically for IDP interactions.

We have characterized and demonstrated the critical role of W33 in the EspF_U P motif for high affinity binding to IRTKS SH3 *in vitro* and for actin pedestal formation *in vivo*. Strikingly, a single correctly positioned residue in a bacterial effector that mimics its host counterpart is sufficient to deceive host signaling and to enable a hostile takeover by having a 60 times higher affinity than its cellular counterparts. Indeed, our data show that W33 plays a decisive role in IRTKS mediated recruitment of EspF_U to Tir upon bacterial clustering and actin pedestal formation *in vivo*.

Molecular characterization of the P motif echoes its IDP nature; it is highly disordered as evidenced by the nuclear spin relaxation and bioinformatics analyses. Yet, amino acid composition of the P motif in EspF_U and cellular ligands is more typical for IDPs than linear motifs (LMs are depleted in Ala, Gly and enriched in aromatic, Cys and Leu residues) (Fuxreiter et al., 2007). Indeed, our findings add an interesting detail to so-called Y-F-W conundrum (Uversky, 2011). Although aromatic residues are rare in IDPs, they are strategically positioned and often participate in protein interactions (Fuxreiter et al., 2004; Uversky, 2011). In contrast, aromatic residues are often found enriched in LMs or short molecular recognition elements termed/consensus sequences (Fuxreiter et al., 2007). Given that cellular ligands of IRTKS/IRSp53 SH3 that contain the tandem PxxP motif do not have any aromatic residues, the strategic positioning of an aromatic residue is certainly true for EspF_U. The information regarding aromatic π - π interactions involving IDPs is very limited

(Espinoza-Fonseca, 2012). Yet, occurrence of tryptophans (~10 %) in π - π interactions is low in comparison to Phe (~60 %) and Tyr (~25 %). Furthermore, no π - π interactions have been reported between two tryptophans at the molecular interfaces involving an IDP (Espinoza-Fonseca, 2012). Our thermodynamical data show that this unique intermolecular π - π interaction between tryptophan residues adds 9.3 kJ/mol to the Gibbs free energy of binding between IRTKS/IRSp53 and EspF_U.

While EspF_U undergoes a disorder-to-order transition upon binding to N-WASP and IRTKS, it also contains segments that remain flexible or disordered in the ternary complex. The ²²NMAE²⁵ linker and especially the C-terminal tail, ⁴¹Gln-Pro⁴⁷, remain flexible when bound to N-WASP and IRTKS. These linkers are completely conserved among the EspF_U repeats, and the C-terminal tail contains mostly polar residues, Gln, Asn, and Ser, as well as Pro. It is quite likely the intrinsic flexibility of N-WASP- and IRTKS-bound EspF_U enables a relatively unhindered spatial search by attached domains and conformational readjustment upon recruitment of multiple N-WASP and IRTKS ligands. Indeed, a recent survey has shown that IDPs often contain repeat regions that might have evolved via repeat expansion, and have a role in the assembly of macromolecular arrays (Davey et al., 2010 & 2012). Short linker sequences of similar composition, known as Q-linkers owing to a high proportion of polar residues in these segments (~70%), are found in a number of bacterial regulatory proteins (Dyson & Wrighth, 2005; Wootton & Drummond, 1989).

Classical SH3 ligands consist of a proline-rich motif flanked by a positively charged residue. The ligands adopt a left-handed PPII helical conformation as they bind a hydrophobic groove on the SH3 surface. The positively charged residue forms a salt bridge with aspartate or glutamate at the bottom of the specificity pocket. SH3 interactions with short peptide ligands are typically characterized with favorable enthalpic contribution to ΔG , which is counterbalanced by an unfavorable entropic penalty. This seems surprising considering the large hydrophobic interaction surface involved in binding (Palencia et al., 2004; Wang et al., 2001). IRTKS SH3 recognizes two consecutive proline-rich motifs of EspF_U connected by a three-residue linker. The linker interacts without making any polar contacts with the specificity pocket, thus rendering the interaction particularly hydrophobic. Notwithstanding, IRTKS SH3 – ligand interactions are dominated by an enthalpic contribution similar to other SH3/ligand complexes (Table 3). Enthalpy-driven hydrophobic complexation is proposed to arise from poor solvation of the binding surface in the unbound state (Bissantz et al., 2010). The enthalpy gain results from stronger hydrogen bonds formed between water molecules released from the surface of the protein upon binding, known as the “nonclassical hydrophobic effect” phenomenon, and is characteristic for many previously described complexation processes (Meyer et al., 2003). We propose that this is also generally valid for SH3 ligands interactions.

Favorable entropy terms observed for wt Eps8, Eps8^{A33W} and EspF_U^{W33A} interactions suggest that larger hydrophobic surfaces than those found in classical SH3 - ligand interactions are a prerequisite for a favorable entropic contribution. On the contrary, EspF_U has a larger entropic penalty. Our structural and dynamic characterization of EspF_U reveals its highly disordered nature when unbound, and shows that binding to IRTKS is accompanied by substantial disorder-to-order transition resulting in decreased conformational entropy. Furthermore P34 has a significant *cis* population in the unbound state, and undergoes conformational change to *trans* upon binding to IRTKS. Both of these factors contribute unfavorably to entropy. Although differences in the thermodynamic signatures of EspF_U and Eps8^{A33W} remain to be unraveled, a similar thermodynamic profile is observed when comparing the binding of SLP-76 and HPK1 peptides to the C-terminal Mona/Gads SH3 domain (Lewitzky et al., 2004; Harkiolaki et al., 2003). These peptides

interact with similar binding sites, make similar hydrophobic and polar contacts, and yet their binding affinity, enthalpic, and entropic terms differ considerably.

In connection with SH3 mediated interactions it is often discussed whether or not simple PxxP motifs are sufficient to achieve specificity. Many SH3 domains have been reported to bind peptides without PxxP motifs, and also several examples indicate that regions beyond this motif are involved in recognition. For example p67^{phox} SH3 binds tightly to a p47^{phox} peptide consisting of a PxxP unit with a C-terminally flanking segment that forms a helix-turn-helix structure (Kami et al., 2002). On the other hand in the case of SH3 binding by HIV-1 Nef, the whole RT-loop of the SH3 domain contributes considerably to binding (Lee et al., 1996). Our study clearly demonstrates that the opposite scenario is also possible. Indeed, EspF_U contains only a single proline-rich sequence available for the SH3 interaction, one that enables it to outcompete its host counterpart whether or not that interaction involves interactions beyond PxxP.

EXPERIMENTAL PROCEDURES

Yeast Two-hybrid Analyses

The two-hybrid expression vectors pGAD424 and pBTM116, as well as reporter strain L40, were used to define the interaction between IRTKS, IRSp53, EspF_U and EHEC Tir as previously described (Cheng et al., 2008; Liu et al., 2002). ONPG assays were performed as previously described (Garmendia et al., 2004; Cheng et al., 2008). See also supplemental experimental procedures & Tables S1 and S2.

Mammalian cell infections and immunofluorescence microscopy

For microscopic analysis, mammalian cells were grown, infected with bacteria and processed as described previously (Garmendia et al., 2004; Cheng et al., 2008; Campellone & Leong, 2005). Cells were treated with mouse anti-HA tag mAb HA.11 (1:500; Covance), mouse anti-HA Alexa-488 (Invitrogen), or mouse anti-IRTKS mAb (1:100; Novus Biologicals).

NMR spectroscopy

All NMR spectra were measured at 25 °C, using either Varian INOVA 600 MHz or 800 MHz spectrometers, equipped with a 5 mm ¹⁵N/¹³C/¹H z-gradient triple-resonance coldprobes.

The spectra for the main-chain and side-chain resonance assignments as well as for measuring ¹⁵N dynamics and ¹H-¹⁵N RDCs were recorded at 800 MHz. The chemical shift perturbation mapping of ¹⁵N, ¹³C labeled IRTKS SH3 with unlabeled peptides was carried out at 600 MHz. (see also supplemental experimental procedures).

Resonance assignments were carried out both for free EspF_U R47₅, and for each subcomponent of the trimolecular N-WASP-GBD: IRTKS SH3: EspF_U R47₅ ternary complex. For the assignment and structure determination of ternary complex, three differentially labeled samples were made and mixed together in 1:1:1 ratio:

1. ¹⁵N, ¹³C N-WASP GBD: IRTKS SH3: EspF_U R47₅,
2. N-WASP GBD: ¹⁵N, ¹³C IRTKS SH3: EspF_U R47₅,
3. N-WASP GBD: IRTKS SH3: ¹⁵N, ¹³C EspF_U R47₅.

Structure calculation

Structure calculation of the EspF_U R475:N-WASP GBD:IRTKS SH3 complex was carried out automatically using the software package CYANA (Herrmann et al., 2002). Peaks were picked manually from ¹⁵N- and ¹³C NOESY spectra. The peak lists, together with the chemical shift assignments, were used as input for the iterative NOE assignment and structure calculations. During structure calculations the protein sequences were connected through a set of weightless non-interacting dummy atoms from C-terminus to N-terminus in the order N-WASP GBD, IRTKS SH3 and EspF_U R475. We generated 200 conformers in each of the seven cycles of the combined automated NOESY and structure calculation algorithm. The final 20 structures were energy-minimized, using CYANA-derived NOE restraints, in AMBER 8 (Case et al., 2005). One thousand iterations with the standard AMBER force field and generalized Born implicit solvent model were performed. Quality of structure was analyzed with PROCHECK-NMR (Laskowski et al., 1996), indicating that 81.4 %, 17.1 % 1.1 % and 0.4 % of the residues are in the most favored, additionally allowed regions, generously allowed and disallowed regions respectively, of the Ramachandran plot.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed at 25°C using a VP-ITC microcalorimeter (Microcal, Inc. Northampton, MA, USA). Eps8 WT, Eps8^{A33W} and EspF_U^{W33A} peptides were dissolved in ddH₂O and pH was adjusted to 7 with NaOH, lyophilized and dissolved in NMR buffer for final concentration of 0.25 mM (Eps8^{A33W}), 0.5 mM (Eps8 WT) or 1 mM (EspF_U^{W33A}). Peptides were titrated separately into the 20 µM (Eps8^{A33W} and Eps8 WT) or 65 µM (EspF_U^{W33A}) IRTKS SH3 solution in the sample cell. In addition, 0.22 mM EspF_U was titrated to 10 µM N-WASP, and 0.1 mM IRTKS SH3 - EspF_U complex to 10 µM N-WASP solution in the sample cell. Experiments were repeated twice. In order to measure heats of dilution, control experiments were performed by titrating peptide to buffer and subtracted from raw titration data. Thermodynamic profile of the IRTKS SH3 and peptide interactions, were obtained by nonlinear least square fitting of experimental data using a single-site binding model of the Origin® 7 software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Elina Ahovalo for excellent technical assistance. We thank Philip Simister and Stephen Feller for the ITC data for Mona/Gads – HPK1/SLP-76, and Hui-Chun Cheng and Michael Rosen for thermodynamical data on N-WASPC - EspF_U R475. This work was supported by the Academy of Finland grant 131144 to P.P and NIH R01AI46454 to J.M.L.

References

- Aitio O, Hellman M, Kazlauskas A, Vingadassalom DF, Leong JM, Saksela K, Permi P. Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly. *Proc Natl Acad Sci USA*. 2010; 107:21743–21748. [PubMed: 21098279]
- Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol*. 2011; 21:432–440. [PubMed: 21514144]
- Bissantz C, Kuhn B, Stahl M. A medicinal chemist's guide to molecular interactions. *J Med Chem*. 2010; 53:5061–5084. [PubMed: 20345171]
- Campellone KG, Robbins D, Leong JM. EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell*. 2004; 7:217–228. [PubMed: 15296718]

- Campellone KG, Leong JM. Nck-independent actin assembly is mediated by two phosphorylated tyrosines within enteropathogenic *Escherichia coli* Tir. *Mol Microbiol*. 2005; 56:416–432. [PubMed: 15813734]
- Campellone KG. Cytoskeleton-modulating effectors of enteropathogenic and enterohaemorrhagic *Escherichia coli*: Tir, EspF_U and actin pedestal assembly. *FEBS J*. 2010; 277:2390–2402. [PubMed: 20477869]
- Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. *J Comput Chem*. 2005; 26:1668–1688. [PubMed: 16200636]
- Cheng HC, Skehan BM, Campellone KG, Leong JM, Rosen MK. Structural mechanism of WASP activation by the enterohaemorrhagic *E. coli* effector EspF(U). *Nature*. 2008; 454:1009–1013. [PubMed: 18650809]
- Davey NE, Trave G, Gibson TJ. How viruses hijack cell regulation. *Trends Biochem Sci*. 2010; 36:159–169. [PubMed: 21146412]
- Davey NE, Van Roey KV, Weatheritt RJ, Toedt G, Uyar B, Altenberg B, Budd A, Diella F, Dinkel H, Gibson TJ. Attributes of short linear motifs. *Mol Biosyst*. 2012; 8:268–281. [PubMed: 21909575]
- Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. 2005; 6:197–208. [PubMed: 15738986]
- Espinoza-Fonseca LM. Aromatic residues link binding and function of intrinsically disordered proteins. *Mol Biosyst*. 2012; 8:237–246. [PubMed: 21863198]
- Farrow NA, Zhang O, Szabo A, Torchia DA, Kay LE. Spectral density function mapping using ¹⁵N relaxation data exclusively. *J Biomol NMR*. 1995; 6:153–162. [PubMed: 8589604]
- Frederick KK, Marlow MS, Valentine KG, Wand J. Conformational entropy in molecular recognition by proteins. *Nature*. 2007; 448:325–330. [PubMed: 17637663]
- Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol*. 2004; 338:1015–1026. [PubMed: 15111064]
- Fuxreiter M, Tompa P, Simon I. Structural disorder imparts plasticity on linear motifs. *Bioinformatics*. 2007; 23:950–960. [PubMed: 17387114]
- Garmendia J, Phillips AD, Carlier MF, Chong Y, Schüller S, Marches O, Dahan S, Oswald E, Shaw RK, Knutton S, et al. TccP is an enterohaemorrhagic *Escherichia coli* O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cell Microbiol*. 2004; 6:1167–1183. [PubMed: 15527496]
- Garmendia J, Carlier MF, Egile C, Didry D, Frankel G. Characterization of TccP-mediated N-WASP activation during enterohaemorrhagic *Escherichia coli* infection. *Cell Microbiol*. 2006; 8:1444–1455. [PubMed: 16922863]
- Goddard TD, Kneller DG. Sparky. Vol. 3. University of California; San Francisco: 2002.
- Harkiolaki M, Lewitzky M, Gilber RJC, Jones EY, Bourette RP, Mouchiroud G, Sondermann H, Moarefi I, Feller SM. Structural basis for SH3 domain mediated high affinity binding between Mona/Gads and SLP-76. *EMBO J*. 2003; 22:2571–2582. [PubMed: 12773374]
- Hazy E, Tompa P. Limitations of Induced Folding in Molecular Recognition by Intrinsically Disordered Proteins. *Chem Phys Chem*. 2009; 10:1415–1419. [PubMed: 19462392]
- Hellman M, Tossavainen H, Rappu P, Heino J, Permi P. Characterization of Intrinsically Disordered Prostate Associated Gene (PAGE5) at Single Residue Resolution by NMR Spectroscopy. *PLoS One*. 2011; 6:e26633. [PubMed: 22073178]
- Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol*. 2002; 319:209–227. [PubMed: 12051947]
- Kami K, Takeya R, Sumimoto H, Kohda D. Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. *EMBO J*. 2002; 21:4268–4276. [PubMed: 12169629]
- Kjaergaard M, Poulsen F. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *J Biomol NMR*. 2011; 50:157–165. [PubMed: 21604143]

- Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 1996; 8:477–486. [PubMed: 9008363]
- Lee CH, Saksela K, Mirza UA, Chait BT, Kuriyan J. Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell*. 1996; 85:931–942. [PubMed: 8681387]
- Lefèvre JF, Dayie KT, Peng JW, Wagner G. Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions. *Biochemistry*. 1996; 35:2674–2686. [PubMed: 8611573]
- Lewitzky M, Harkiolaki M, Domart MC, Jones EY, Feller SM. Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif R/KxxK, with a classical PxxP motif embedded in a polyproline type II (PPII) helix. *J Biol Chem*. 2004; 279:28724–28732. [PubMed: 15100220]
- Liu H, Radhakrishnan P, Magoun L, Prabu M, Campellone KG, Savage P, He F, Schiffer CA, Leong JM. Point mutants of EHEC intimin that diminish Tir recognition and actin pedestal formation highlight a putative Tir binding pocket. *Mol Microbiol*. 2002; 45:1557–1573. [PubMed: 12354225]
- Marsh JA, Singh VK, Jia Z, Forman-Kay JD. Sensitivity of secondary structure propensities to sequence differences between α - and γ -synuclein: Implications for fibrillation. *Protein Sci*. 2006; 15:2795–2804. [PubMed: 17088319]
- Meyer EA, Castellano RK, Diederich F. Interactions with aromatic rings in chemical and biological recognition. *Angew Chem Int Ed Engl*. 2003; 42:1210–1250. [PubMed: 12645054]
- Mittag T, Kay LE, Forman-Kay JD. Protein dynamics and conformational disorder in molecular recognition. *J Mol Recognit*. 2009; 23:105–114. [PubMed: 19585546]
- Mukrasch MD, Bibow S, Korukottu J, Jeganathan S, Biernat J, Griesinger C, Mandelkow E, Zweckstetter M. Structural polymorphism of 441-residue tau at single residue resolution. *PLoS Biol*. 2009; 7:e34. [PubMed: 19226187]
- Mäntylähti S, Aitio O, Hellman M, Permi P. HA-detected experiments for the backbone assignment of intrinsically disordered proteins. *J Biomol NMR*. 2010; 47:171–181. [PubMed: 20437194]
- Mäntylähti S, Hellman M, Permi P. Extension of the HA-detection based approach: (HCA)CON(CA)H and (HCA)NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. *J Biomol NMR*. 2011; 49:99–109. [PubMed: 21259120]
- Nagulapalli M, Parigi G, Yuan J, Gsponer J, Deraos G, Bamm VV, Harauz G, Matsoukas J, de Planque MRR, Gerothanassis IP, et al. Recognition pliability is coupled to structural heterogeneity: a calmodulin intrinsically disordered binding region complex. *Structure*. 2012; 20:522–533. [PubMed: 22405011]
- Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry*. 2005; 44:12454–12470. [PubMed: 16156658]
- Palencia A, Cobos ES, Mateo PL, Martinez JC, Lague I. Thermodynamic dissection of the binding energetic of proline-rich peptides to the Abl-SH3 domain: Implications for the rational ligand design. *J Mol Biol*. 2004; 336:527–537. [PubMed: 14757063]
- Sallee NA, Rivera GM, Dueber JE, Vasilescu D, Mullins RD, Mayer BJ, Lim WA. The pathogen protein EspF(U) hijacks actin polymerization using mimicry and multivalency. *Nature*. 2008; 454:1005–1008. [PubMed: 18650806]
- Sarkar P, Saleh T, Tzeng SR, Birge RB, Kalodimos CG. Structural basis for regulation of the Crk signaling protein by a proline switch. *Nat Chem Biol*. 2011; 7:51–57. [PubMed: 21131971]
- Tatko CD, Waters ML. Selective aromatic interactions in β -hairpin peptides. *J Am Chem Soc*. 2002; 124:9372–9373. [PubMed: 12167022]
- Tolman JR, Al-Hashimi HM, Kay LE, Prestegard JH. Structural and dynamic analysis of residual dipolar coupling data for proteins. *J Am Chem Soc*. 2001; 123:1416–1424. [PubMed: 11456715]
- Uversky VN. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *J Biomed Biotech*. 2010; 2010:1–14.
- Uversky VN. Intrinsically disordered proteins from A to Z. *Int. J Biochem Cell Biol*. 2011; 43:1090–1103.

- Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, Dunker AK. Characterization of molecular recognition features, MoRFs, and their binding partners. *J Proteome Res.* 2007; 6:2351–2366. [PubMed: 17488107]
- Vingadassalom DA, Kazlauskas A, Skehan B, Cheng HC, Magoun L, Robbins D, Rosen MK, Saksela K, Leong JM. Insulin receptor tyrosine kinase substrate links the *E. coli* O157:H7 actin assembly effectors Tir and EspF(U) during pedestal formation. *Proc Natl Acad Sci USA.* 2009; 106:6754–6759. [PubMed: 19366662]
- Wang C, Pawley NH, Nicholson LK. The role of backbone motions in ligand binding to the c-Src SH3 domain. *J Mol Biol.* 2001; 313:873–887. [PubMed: 11697910]
- Weiss SM, Ladwein M, Schmidt D, Ehinger J, Lommel S, Städing K, Beutling U, Disanza A, Frank R, Jänsch L, et al. IRSp53 links the enterohemorrhagic *E. coli* effectors Tir and EspFU for actin pedestal formation. *Cell Host Microbe.* 2009; 5:244–258. [PubMed: 19286134]
- Williams DH, Stephens E, O'Brien DP, Zhou M. Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. *Angew Chem Int Ed Engl.* 2004; 43:6596–6616. [PubMed: 15593167]
- Wishart DS, Bigam CG, Yao J, Abilgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J Biomol NMR.* 1995; 6:135–140. [PubMed: 8589602]
- Wittekind M, Mapelli C, Farmer BT II, Suen KL, Goldfarb V, Tsao J, Lavoie T, Barbacid M, Meyers CA, Mueller L. Orientation of peptide fragments from Sos proteins bound to the N-terminal SH3 domain of Grb2 determined by NMR spectroscopy. *Biochemistry.* 1994; 33:13531–13539. [PubMed: 7947763]
- Wootton JC, Drummond MH. The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Prot Eng.* 1989; 2:535–543.

Figure 1. Structural characterization of free EspFU

A) Amino acid sequence of EspFU fifth repeat (R475) along with N-WASP GBD and IRTKS/IRSp53 SH3 binding epitopes. B) Structural disorder prediction for R475 based on IUPred and Disprot algorithms. C) ^{15}N - ^1H correlation (HSQC) spectrum of ^{15}N , ^{13}C labeled EspFU R475, recorded at 800 MHz ^1H frequency. Narrow range of ^1H chemical shifts is a signature of disordered nature of EspFU R475. D) Analysis of ^{13}C and secondary chemical shifts in unbound EspFU R475. Deviations from residue specific random coil chemical shifts are shown, which take into account the nearest neighbor effects and temperature (Kjaergaard & Poulsen, 2011). E) Values of reduced spectral density functions at three frequencies $0.87\omega_{\text{H}}$, ω_{N} and 0 against the primary sequence of EspFU R475. Transiently populated α -helix as well as XPXP motifs are shown above histograms. See also Figures S1 and S2.

Figure 2. Structure and dynamics ternary complex

Structure of trimolecular complex between N-WASP GBD, EspF_U R47₅ and IRTKS SH3.

A) Ribbon presentation of the lowest energy conformation of the ternary complex. N-WASP GBD orange, EspF_U R47₅ green and IRTKS SH3 magenta. B) Superimposition of 20 lowest energy conformers of N-WASP GBD 212–270: EspF_U 502–521(left) and IRTKS SH3 343–400: EspF_U 527–540 (right). Same coloring as in A). C) Superimposition of GBD domains from N-WASP:EspF_U (red) and the WASP:EspF_U complexes (Cheng et al., 2008) (blue) for residues 216–262. M238, C239 from N-WASP GBD and the corresponding R35 and A36 from WASP are shown in stick presentation. D) Chemical shift perturbations observed ¹H-¹⁵N HSQC spectra between unbound EspF_U R47₅ (blue contours), in complex with IRTKS SH3 (red contours) and in complex with N-WASP GBD and IRTKS SH3 (green contours). E) Comparison of steady state heteronuclear {¹H}-¹⁵N NOEs for unbound EspF_U R47₅ (blue bars) and associated to trimolecular complex with N-WASP GBD and IRTKS SH3 (red bars). F) Color coding of observed {¹H}-¹⁵N NOEs in the complex on the structure of EspF_U R47₅ reflects the increased rigidity on ps-ns timescales for N-WASP binding residues 3–21 and IRTKS SH3 binding epitope H26-V40 (blue coloring). Sustained flexibility on ps-ns timescales (colored green and yellow) is observed for the linker residues 22–24 that connect GBD and SH3 binding domains. The C-terminal end of EspF_U R47₅ remains highly disordered also in the complex. See also Figure S3.

Figure 3. Chemical shift perturbation mapping of IRTKS SH3 upon addition of peptides from EspFU and Eps8

Full assignment of resonances is given for IRTKS SH3:EspFU R47₅ complex (upper left panel). Selected boxed resonances with assignments capture the embedded binding affinity and show that all the peptides bind to the same binding site on IRTKS SH3. The assignments corresponds to the saturated state with IRTKS SH3:peptide shown in blue. Saturated state for EspFU R47₅ and Eps8^{A33W} correspond to SH3:peptide molar ratio 1:1 and for EspFU^{W33A} and Eps8 to SH3:peptide molar ratio 1:3.25. Free IRTKS resonances are shown in red and other colors correspond to intermediate states between free and saturated states.

Figure 4. Structural and functional role of the W-switch

A) Alignment of tandem PxxP containing ligands. Delphinin is a protein that we predicted as a potential novel IRTKS ligand, but this has not been experimentally tested. B) EspF_U W33 establishes a T-shaped edge-to-face arrangement with IRTKS W378. C) Schematic showing a single repeat of EspF_UC. The N-WASP binding helix “H” (Cheng et al., 2008) and IRTKS binding “P” (Weiss et al., 2009; Vingadassolom et al., 2009) domains are indicated. Asterisk indicates the site of the W33A mutation. D) Plasmids encoding the LexA DNA binding domain-IRTKS_{SH3} and the indicated GAL4 AD-EspF_U fusions were co-transformed into a yeast two-hybrid reporter strain L40. “PW33A” indicates the mutant with the alanine substitution of residue W33. β-galactosidase activity was assessed as an average of three co-transformants in Miller Units (MU) with error bars indicating the standard deviation. Results are representative of at least three experiments. E) FLCs expressing myc-tagged GFP-EspF_U fusions were infected with EPEC KC12, which requires ectopic expression of EspF_U for pedestal formation. Red asterisks (and corresponding red stripes in schematic) indicate W33A mutations. Monolayers were stained with DAPI (blue), anti-myc antibody (green), and Alexa568-phalloidin (red). F) FLCs expressing HA-tagged Tir-EspF_U fusion protein carrying W33A mutations were infected with intimin-expressing *E. coli* K12. Monolayers were stained with DAPI (blue), anti-HA antibody (green), and Alexa568-phalloidin (red). F) An engineered tryptophan mutation in Eps8 dramatically enforces its intracellular association with IRTKS. Human 293T cells were transfected with an expression vector for IRTKS tagged with a biotin acceptor domain together with a vector for a Myc-tagged wild-type Eps8 (wt) or a mutated derivative (mut) containing an EspF_U-like tryptophan-containing linker between the PxxP motifs in the IRTKS SH3 domain binding region. IRTKS from lysates of these cells was precipitated using streptavidin-coated beads. Proteins precipitated from these lysates with streptavidin-coated beads were examined by Western blotting using anti-Myc antibodies and labeled streptavidin to detect Eps8 and IRTKS

proteins, respectively. Part of the total lysates was similarly analyzed for Eps8 and IRTKS expression without prior affinity selection, as indicated.

Table 1Statistics of structure calculation for N-WASP-GBD: EspF_U R475: IRTKS SH3 complex.

Distance constraints	N-WASP GBD:EspFU R475:IRTKS SH3
Total	4825
Short range $ i-j \leq 1$	2059
Medium-range $1 < i-j \leq 5$	986
Long-range $ i-j > 5$	1780
Number of constraints per residue	27.0
Structure Statistics	
Average AMBER energy (kcal/mol)	-6057±20
Violations	
Distance constraints violations >0.5Å	-
Deviations from idealized geometry	
Bond lengths (Å)	0.0098±0.0001
Bond angles (°)	2.319±0.017
Average RMSD from mean coordinates (Å)	
Backbone IRTKS SH3 343–400: EspF _U 527–540	0.27±0.06 *
Heavy IRTKS SH3 343–400: EspF _U 527–540	0.69±0.05 *
Backbone N-WASP GBD 212–270: EspF _U 502–521	0.24±0.04 *
Heavy N-WASP GBD 212–270: EspF _U 502–521	0.56±0.04 *
Ramachandran plot (%) **	81.4/17.1/1.1/0.4

* RMSD values are shown for the 20 calculated complex structures

** Residues in most favoured/additionally allowed/generously allowed/disallowed regions of the Ramachandran plot

Table 2

Thermodynamics of EspF_U - N-WASP interactions. See also Figure S4.

	ΔG	ΔH	$-T\Delta S$	K_d	n
N-WASP - EspF _U R47 ₅	-40.9±0.8	-64.6±4.0	23.7±4.9	0.07±0.02	1
N-WASP _C - EspF _U R47 ₅	-42.6 ^{a)}	-28.3 ^{a)}	14.3 ^{a)}	0.035 ^{b)}	1
N-WASP - EspF _U R47 ₅ +IRTKS SH3	-42.4	-66.3	23.9	0.04	1

ΔG , ΔH , $-T\Delta S$ are given in kJ/mol, n is stoichiometry of binding.

K_d is given in 10⁻⁶ M

^{a)} Data kindly provided by Drs. Hui-Chun Cheng and Michael Rosen

^{b)} From ref. (Cheng et al., 2008)

Table 3

Thermodynamics of SH3 interactions. See also Figure S4.

	ΔG	ΔH	$-T\Delta S$	K_d	n
IRTKS SH3 - EspF _U R47 ₅	-35.9	-44.0	8.1	0.5 ^{a)}	1
IRTKS SH3 - EspF _U R ₂	-34.2±0.6	-41.6±1.1	7.4±1.6	1.1±0.2	1
IRTKS SH3 - EspF _U W33A	-26.6±0.3	-10.2±0.6	-16.3±0.4	22.3±2.1	1
IRTKS SH3 - Eps8 WT	-25.8±0.1	-16.8±2.3	-9.1±2.2	30.5±1.7	1
IRTKS SH3 - Eps8 ^{A33W}	-32.1±0.1	-26.6±0.1	-5.6±0.2	2.4±0.1	1
Mona/Gads - HPK1 ^{b)}	-32.08	-29.47	-2.61	2.4	1
Mona/Gads - SLP-76 (P2) ^{b)}	-40.28	-49.35	9.07	0.09	1
Grb2 - SOS-A ^{c)}	-24.7	-40.6	15.6	54	1
Grb2 - SOS-E ^{c)}	-31.8	-54.4	22.3	3.5	1

ΔG , ΔH , $-T\Delta S$ are given in kJ/mol, n is stoichiometry of binding.

K_d is given in 10^{-6} M

^{a)} From ref. (Aitio et al., 2010)

^{b)} Data kindly provided by Drs. Philip Simister and Stephan Feller

^{c)} From ref. (Wittekind et al., 1994).