A pathogen-encoded signaling receptor mediating host-like interactions through intrinsic disorder

Marta F. M. Vieira[a], Guillem Hernandez[a], Tiago Veloso[a], Miguel Arbesú[b], Andreas Zanzoni[c] Tiago N. Cordeiro[a]

[a]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras Portugal

[b]Department of NMR-supported Structural Biology, Leibniz-Forschungsinstitut für Molekulare Pharmakologie, Robert-Rössle-strasse 10, 13125 Berlin, Germany

[c]Aix-Marseille Univ, Inserm, TAGC, UMR_S1090, Marseille, France

Corresponding author: tiago.cordeiro@itqb.unl.pt

Abstract

The translocated intimin receptor (Tir) is a central effector of Attaching and Effacing (A/E) pathogens responsible for worldwide foodborne disease cases. Upon delivery into host cells, Tir acts as a cell-signaling receptor, rewiring host cellular processes to assist infection. We found that this bacterial-encoded transmembrane protein comprises highly disordered intracellular domains bearing host-like motifs that bind host proteins. This unexpected trait was found prevalent in several other effectors secreted by A/E bacteria. We assessed Tir's intrinsic disorder by an integrative structural biophysics approach, unveiling that its intracellular side comprises a partially structured N-terminal dimer (N-Tir) and a disordered C-terminal tail (C-Tir). NMR analysis revealed that C-Tir has pre-existing transient structures at phosphorylation sites, including host-like immunoreceptor tyrosine-based inhibitory motifs (ITIMs). These ITIM-like sequences were found to bind lipid bilayers as previously observed for host T-cell receptor cytoplasmic disordered domains. C-Tir's membrane affinity is residue-specific and modulated by lipid composition, suggesting a regulation layer based on membrane composition. Using NMR, we also observed that the disordered C-Tir displays multisite tyrosine-phosphorylation sites that mediate promiscuous binding to the C-terminal SH2 domain of host SHP-1 in dynamic equilibrium. Together, these novel insights provide an updated picture of Tir's structural features and highlight Tir-mediated mimicry of host disordered membrane receptors as a molecular strategy for host cell subversion.

Keywords

Intrinsically disordered proteins | bacterial effectors | EPEC/EHEC | molecular mimicry and multivalency | Host-pathogen interactions

Summary

Tir is a cellular receptor secreted by life-threatening pathogens. Upon delivery into host cells, Tir inserts the host plasma membrane providing a means for these extracellular pathogens to control host intracellular processes. To prevent pathogens from relying on Tir, it is essential to understand its intracellular mechanics. This paper provides a coherent picture of the intracellular side of Tir, highlighting its ability to copycat the interactions of disordered intracellular domains of host immune receptors. This copycatting allows the bacterial pathogens to modulate critical host processes, allowing infection to spread further without triggering the immune system response. This work proposes that other bacterial secreted pathogenic proteins exploit intrinsic disorder to hijack human cells, suggesting a widespread host subversion mechanism.

Introduction

Intrinsically disordered proteins and regions (IDPs and IDRs, respectively) are widespread across all kingdoms of life (1). They do not adopt well-defined structures but exist as dynamic conformational ensembles that display unique properties complementary to globular proteins (2, 3). IDPs/IDRs play essential roles in virtually all signaling pathways in the cell (4), including cell cycle (5), circadian circuits (6), post-transcriptional regulation (7), and protein degradation (8). Given their ubiquitous relevance in cellular signaling and regulation, the onset of several human diseases, including cancer and neurodegeneration, are linked to dysfunctional disordered proteins (9). A common hallmark of IDPs/IDRs is the high occurrence of short linear motifs, also known as eukaryotic linear motifs (ELMs) (10), which are stretches of 3-10 contiguous residues mediating transient protein-protein interactions (e.g., post-translational modifications sites, targeting signals) (11). For instance, phosphorylation sites are often found in protein IDRs, where they can modify charge and hydrophobicity and modulate interactions with partners (7). The prevalence of such functional elements in IDRs provides versatility to cell interaction hubs by enabling flexibility and adaptability to multiple interaction interfaces (12, 13). Importantly, IDR sequence composition is biased towards low complexity and charged residues favoring electrostatic interactions at lipid bilayer surfaces (14), with lipid-binding proteins accounting for 15% of all disordered proteins (15). As a result, intrinsic disorder is also common in membrane receptors, particularly in their intracellular domains (16, 17).

Disordered proteins make up 30-50% of eukaryotes' proteins (18, 19). In contrast, they are remarkably less abundant in prokaryotes. However, increasing experimental evidence in the literature shows that several effector proteins secreted by pathogenic bacteria contain functionally relevant IDRs (20). A well-known case is the oncogenic *Helicobacter pylori* effector CagA, which promiscuously recruits host proteins to potentiate oncogenic signaling via a long IDR (22). Another illustrative example is the effector CyaA of *Bordetella pertussis*, the causative agent of whooping cough that interacts with host calmodulin via a disordered stretch of 75 amino acid residues (21). Other less-ordered effectors are also secreted by Enteropathogenic and Enterohemorrhagic *Escherichia coli* strains (EPEC and EHEC, respectively), such as EspF(U)/TccP (22) and EspB (23).

EPEC is a leading cause of child deaths worldwide (24), and EHEC causes hemolytic uremic syndrome (HUS), defined by acute kidney failure, low platelet count, and destruction of red blood cells (25). These bacteria attach to the gut mucosa, causing a histopathological condition known as the 'attaching and effacing' (A/E) lesions, characterized by actin-rich pedestal formation and microvilli degeneration. To establish a replicative niche on host cells, EPEC/EHEC delivers a vast repertoire of effectors (26, 27). In light of this, we ran a structural disorder analysis on the collection of effectors secreted by A/E-bacteria and searched for ELMs evidence. Among A/E-bacteria disordered effectors, we found the prominent translocated intimin receptor (Tir) to be a disordered membrane receptor with a relatively high density of host-like ELMs, particularly in its intracellular domains. Tir is the first effector secreted during infection (28). Once inside the host, it migrates to the plasma membrane to act as a receptor for intimin presented at the bacterial surface (29), anchoring A/E pathogens to targeted host cells. By mediating attachment, Tir also facilitates translocation of additional T3SS secreted effectors (30). Moreover, Tir promotes actin polymerization (31), suppresses autophagy (32) and immune response (33) to allow bacterial survival on the host cell's surface. It also triggers pyroptotic cell death (34) (35). As a membrane receptor, Tir adopts a hairpin topology with the external intimin-binding domain (IBD) (36) connecting two transmembrane domains, with both N- and C-terminus in the host cytoplasm, interacting with multiple intracellular eukaryotic proteins (37, 38). Those intracellular regions enable Tir to interact with at least 25 host intracellular targets (38), such as host tyrosine phosphatases SHP-1/2 to suppress pro-inflammatory cytokines signaling (33, 39) and TAK1-mediated immune response (40).

Despite all these findings, the structural mechanisms underlying Tir intracellular regions' function during infection are poorly understood. To gain further insights into Tir structural dynamics, we devised an integrative structure-based biophysical study to experimentally assess their intrinsic disorder and multivalency. In particular, we focus our attention on self-assembly, accessibility to multi-site phosphorylation, host protein- and membrane-binding of Tir intracellular regions. We found that N-Tir is a flexible dimer with a hybrid architecture of ordered and disordered regions, opening the question of whether host signaling activation involves structural changes of preformed dimers. Moreover, the present work shows that C-Tir can bind to lipid bilayers through short host-like motifs that, when phosphorylated, interact with the C-terminal SH2 domain (C-SH2) of host SHP-1 in dynamic equilibrium. Together these findings highlight a molecular mechanism through which Tir can interact with host components through mimicry and multivalency.

Results and Discussion

Structural disorder and short linear motifs are common features of A/E pathogen effectors

To quantify the prevalence of disorder propensity among A/E pathogens, we predicted the disorder content of A/E effectors by using DISOPRED3 (41) (**Fig. 1**) and IUPred 1.0 (42, 43)) (**Fig. S1**) structural disorder predictors. We observed that A/E effectors have a higher disorder propensity than their proteome counterparts in EHEC O157:H7, the closely related Enteropathogenic *E. coli* H127:H6 (EPEC), and *Citrobacter rodentium* (CR) — the standard small animal A/E pathogen model (**Fig. 1*A*, Table S1**). Subsequently, we classified them into five structural categories, based on disorder content (see **Material and Methods for details**). Our analysis shows that A/E pathogens have a structurally diverse repertoire of effectors, ranging from fully unstructured to ordered proteins (**Fig. 1*B***). While most proteins are folded in all prokaryotic collections as a whole and individually (**Fig. S2*A***), fully disordered effectors are predicted to be two to four-fold more frequent than the whole bacterial proteome (1.9% vs. 7.7 % on average, far from the 32.7 % of the human proteome) (**Fig. S1*B***). Partially Disordered Proteins (PDR) with long disordered regions occur in EPEC effectors similar to the human proteome. For instance, we classified as PDR the EspB effector, which is involved in cytoskeleton rearrangement and previously identified as an inherently less-ordered effector (23). An additional example of PDR is the NleH effector, a protein kinase with a disordered N-terminal domain that binds the ribosomal protein 3 (RPS3) to manipulate the NF-κB pathway (44, 45). In the order-to-disorder continuum, the effector EspF is akin to EspF(U) as a fully disordered protein (**Fig.1*C***). Both effectors contain multiple consecutive repeats of linear motif pairs able to bind critical host components for actin assembly, with EspF having three to five repetitions depending on the strain (46), and EspFu, only secreted by EHEC, five-and-a-half repeats (47) (**Fig. S2*B***).

Our analysis highlights Tir as a disordered effector (**Fig. 1*C*, Fig. S2*B***). Except for the transmembrane domains, Tir displays a high disorder propensity within its IBD and intracellular regions (**Fig. S1*D*)**. Increasing evidence shows that pathogens exploit host-like motifs as a strategy to interact with host components to subvert host cellular functions (48–50). On this basis, we searched the A/E effectors for eukaryotic linear motifs (ELMs) and found that aside from EspF and EspF(U), other A/E effectors bear multiple putative ELM instances (**Fig. 1*C***). Tir emerged again as having several predicted ELMs in their intracellular regions, including experimentally verified SH3- and SH2-domain binding motifs (51). Structural disorder and high ELM density support Tir's ability to interact with several host proteins (38). There is a correlation between disorder content and ELM density, being Tir, EspF(U), and EspF clustered with the highest disorder fraction and motif overall density (**Fig. 1*C***), similarly to eukaryotic IDPs (11, 52). This correlation highlights ELMs within disordered segments of bacterial effectors as a strategy to disrupt host networks. Disordered effectors, such as Tir, are an exception to the general trend that prokaryotic proteins are less disordered. Given the ubiquitous presence of IDPs as transcriptional factors (53) and, more generally, as hubs in host networks (54), bacterial effectors with host-like disordered proteins features are an efficient way to subvert host eukaryotic systems and promote infection. We selected Tir to study how disordered effectors control the host by validating the predicted structural disorder and molecular mimicry within its intracellular domains.

Tir bears an intrinsically disordered C-terminal tail

Tir secreted by A/E pathogens share substantial sequence-conservation and high frequency of disorder-promoting residues at the N- and C-terminal regions (N-Tir and C-Tir, respectively, both localized to the host intracellular side (**Fig. S3**). We assessed experimentally the structural disorder propensity of EPEC N-Tir and C-Tir by multiple biophysical methods, including small-angle X-ray scattering (SAXS), circular dichroism (CD), and liquid-state nuclear magnetic resonance (NMR). For this integrative study, we used the protein constructs encompassing the residues 1-233 and 388-550 from the N-terminal and C-terminal cytosolic regions of EPEC Tir, respectively (**Fig. S3**). Analytical size-exclusion chromatography (SEC) in combination with SAXS indicated a monomeric state for C-Tir in solution with a radius of gyration (*Rg*) of 38.8 ± 0.2 Å and a maximum distance (*Dmax*) of 128.0 ± 5.0 Å (**Table S4**). The *Rg*-values of IDPs approximately followed a power law of $Rg = 2.54N^{0.522}$ (55, 56) as a function of sequence length (*N*). Using this relationship, the predicted *Rg*-value is 37.4 Å for the 173-residue C-Tir construct ending with a StrepTag (8 extra residues). The similarity between the experimental and predicted values suggests that C-Tir might adopt IDP-like structures in solution. The SAXS-derived Kratky plot of C-Tir also shows characteristics of disordered or unfolded proteins; monotonically increasing without a well-defined maximum (**Fig. 2*A***). This feature is absent in globular proteins and implies conformational heterogeneity and flexibility for C-Tir (57). Likewise, the asymmetric pair distance distribution function, *P(r)*, obtained from the scattering data is compatible with a highly flexible protein sampling pairwise distances far exceeding those expected for a globular protein of the same molecular weight (56) (**Fig. S4**). The intrinsic disorder of C-Tir is also reflected in its CD and [^{1}H-^{15}N]-HSQC NMR spectra. C-Tir's CD profile has negative ellipticity at 200 nm and a shallow band in the 210–230 nm range (**Fig. 2*B***), indicating a high content of random coil with minimal ordered structural elements (58). The [^{1}H-^{15}N]-HSQC NMR spectrum for this construct displays a typical IDP fingerprint with low chemical shift dispersion (59) (**Fig. 2*C***). We observed a similar spectrum for the equivalent region of Tir from EHEC (**Fig. 2*C***). Although all A/E pathogens produce Tir, actin polymerization by EHEC Tir C-terminus requires Esp(U)/TccP (60), which is unneeded in EPEC to initiate actin pedestal formation in host cells (31). Besides the different mechanisms, C-Tir is disordered in both strains, thus highlighting the conservation and functional importance of structural disorder in this intracellular domain for A/E virulence.

Overall, the different biophysical measurements are consistent with a highly dynamic protein, unambiguously showing that the C-terminal half of Tir is an IDR that lacks a well-folded structure under native-like conditions. This inherent flexibility potentially enables Tir for multiple binding while working as a pathogenic scaffold/hub recruiting host proteins to perturb cellular functions.

C-Tir displays non-random structural preferences at sites of phosphorylation and host interactions

We employed NMR to dissect with further detail subtle structural features of C-Tir. The NMR chemical shift is the most readily accessible observable in NMR and sensitive to, even partial, secondary structural elements present in IDPs (59). To gain access to this information, we assigned the backbone resonances of C-Tir nuclei using standard triple-resonance spectra at high-field, together with reverse labeling (**Fig. S5**). Reverse labeling of selected amino-acids (61) proved to be a cost-efficient approach to assign this disordered protein by reducing the spectral ambiguity and should be readily applied to other IDPs (see SI **Methods and Materials**).

Provided with the assignment, we calculated the NMR chemical shift deviations ($\Delta\delta$) to random coil values of intrinsically disordered proteins (62). Although C-Tir has properties of a random coil, these so-called NMR secondary shifts allowed us to identify two regions with non-random structural preferences. In IDPs, transient secondary structural elements are often crucial in acting as molecular recognition units, playing significant roles in binding (63, 64). In figure **Fig. 3*A***, the consecutive positive $\Delta\delta C\alpha$–$\Delta\delta C\beta$ values indicate a tendency for α-helical structure in two distal segments of C-Tir: (I) D420-Q435; II) I507-A515. Both stretches include residues that are phosphorylated by host kinases, S434 and Y511, respectively. By taking into account all the assigned backbone resonances into a neighbor-corrected sequence structural propensity calculator (nsSPC) score (65), these apparent local α-helical regions became even more evident. New stretches also emerged with a putative α-helical propensity, including a sequence contiguous to Y511 and downstream tyrosine 483 and a few C-terminal residues (**Fig. 3*A***).

Structural signatures near phosphorylation sites can affect kinase selectivity to disordered substrates (66), and phosphorylation can modulate partially structured elements for regulation (67). Phosphorylation of Tir on serines S434 and S463 by host protein kinase A (PKA) triggers changes that affect its conformational and self-assembling properties (68, 69). So, the enhanced α-helical propensity near S434 might play a role in PKA-mediated phosphorylation effects on Tir structure and function (32, 70). Moreover, tyrosines Y483 and Y511 are core residues of Tir's host-like Immunoreceptor tyrosine-based inhibitory motifs (ITIMs). The intracellular signaling associated with Tir ITIMs relies on the phosphorylation of the central tyrosine by Src family kinases (SFKs), creating a binding site for SH2–containing proteins, which become activated upon recruitment. Yet, we found that unphosphorylated C-Tir can recognize host SH2-domains with low-intermediate affinity. Using the C-terminal SH2 domain (C-SH2) of phosphatase SHP-1, we identified, by NMR, that residues surrounding the unphosphorylated Y511 ITIM-like motif are involved in binding SH2 domains (**Fig. 3*B-C***). Binding to C-Tir caused a selective loss of intensities for the HSQC peaks, mainly from residues $A_{512}LLA_{515}$. Visible NMR signals retained low-dispersion, indicating that the protein remains mostly disordered and flexible in the complex. We estimated an apparent K_D of *ca.* 68.5μM for this interaction (**Fig. 3*D***). Thus, besides encoding for α-helical structures, the sequence surrounding Y511 (i+4) has a chemical signature for SH2-domain binding pre-phosphorylation. This observation highlights that those residues adjacent to Y511 can define a binding motif for C-SH2, likely serving as a precursor to a tighter binding upon phosphorylation. The presence of a pre-existing transiently helix at this region suggests a possible role for this structure as molecular recognition elements for C-Tir.

Remarkably, Tir ITIMs share with cytoplasmic tyrosine-based sequence motifs of host receptors the ability to adopt α-helical conformations. Indeed, host ITIMs and Immunoreceptor tyrosine-based activation (ITAMs) can form dynamic/transient α-helices, whether in free (71, 72) or membrane-bound state (73). Their dynamic binding to membranes is postulated to regulate tyrosine sites' access to phosphorylation, providing a switch between functional and non-functional conformations (74). The helical conformation in ITIMs/ITAMs is stabilized *in vitro* in the presence of 2,2,2-trifluoroethanol (TFE) or through binding to detergent micelles and lipids (72, 73, 75). In the presence of membrane-mimetic solvent TFE, we found that C-Tir exhibited a substantial increase in secondary structure, with the CD signature of a partial helical protein (**Fig. S6*A***). This conformational change is followed by chemical shift differences (**Fig. S6*B-C)*** and a differential decrease in peak intensities of the C-Tir amide proton resonances in the presence of increasing amounts of TFE, mainly around Y511 and Y483 and not at S434 (**Fig 3*E***). This effect supports an enhanced α-helical conformation propensity around the two tyrosines, and a putative role in membrane-binding, as postulated for some host tyrosine-based motifs that undergo TFE-induced structural changes (72).

Our NMR results show that unphosphorylated C-Tir host-like ITIM motifs can adopt pre-formed structures relevant to membrane binding and molecular recognition of SH2 domains of signaling proteins.

Disordered C-Tir binds lipids

The ability to form α-helices stabilized by TFE prompts us to further investigate lipid binding by C-Tir. To do so, we incubated ^{15}N-labelled C-Tir with different bicelles, which are small planar bilayers of long-chain lipids closed by curved micelle-like walls of short-chain lipids. These disk-like structures are membrane models extensively used to explore protein-membrane interactions, including those involving disordered proteins (76). As long-chain lipids, we used 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) or the acidic lipid 1,2-dimyristoyl-*sn*-glycero-3-phosphorylglycerol (DMPG) mixed with the short-chain lipid 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC), offering the possibility to create bicelles with different lipid charge density.

In the presence of bicelles containing DMPC lipids, we observed a significant residue-specific decrease in the ratio of NMR signals around Y511; yet more subtle, also around Y483, as previously induced with TFE (**Fig. 4**). This finding shows that C-Tir does interact with non-charged lipid bilayers predominantly via its Y511-motif, including mostly hydrophobic residues, such as $A_{512}LLA_{515}$. We further tested the influence of increasing negatively charged lipid head groups on C-Tir membrane association with bicelles containing DMPG lipids. Our NMR data show a more extended NMR signal attenuation in the presence of DMPG/DHPC bicelles, even including the first stretch of residues of this construct, which display some positively charged residues (e.g., R388 and K389) (**Fig. 4**). In the context of the full-length protein, these residues define the membrane-proximal region (**Fig. S3**). Given its proximity to the cell plasma membrane, it is reasonable to assume its involvement in membrane binding. Aside from Y511 surrounding residues, the sequence around Y483 is also notably affected by anionic lipid content, indicating that lipid-charge content modulates C-Tir membrane binding modes. Contrary, the residues 395-475 did not show any membrane interaction. This region has a net negative charge (pI~4.5), whereas the sequences displaying membrane ability are positive (pI~9.9), thus in line with an electrostatic model. In this residue range, C-Tir bears two experimental confirmed host protein binding sites:

a) NCK Src Homology 2 (SH2) domain binding motif (31); and b) NPY motif that interacts with the I-BAR domain of host IRSp53/IRTKS (77). With both bicelles, the last C-terminal residues of C-Tir also show NMR-signal attenuation indicative of membrane-binding (**Fig. 4*B***). These residues (A_{541}PTPGPVRFV_{550}) emerge as part of a potential new lipid-binding region, also affected by TFE (**Fig. 3*E*, *Fig. S6C***) and showing ncSPC values indicating α-helical propensity (**Fig. 3*A***). Our data thus suggest that C-Tir can undergo multivalent electrostatic interactions with lipid bilayers that might potentially fine-tune Tir's activity in the host cell.

Multi-phosphorylated C-Tir is a disordered multivalent cytosolic tail

In the host, Tir becomes phosphorylated by Src family protein tyrosine kinases (PTKs) at several tyrosine phosphorylation sites targeting SH2 domain-containing host proteins. To fully assess its ability to recruit SH2 domains, we reconstructed C-Tir's Tyr-phosphorylated state and evaluated its binding to C-SH2. To this end, we incubated C-Tir with the Src family PTK Fyn, and quantitatively monitored its phosphorylation by NMR (78). The sensitivity of chemical shifts to changes caused by phosphorylation allowed us to identify four phosphorylated tyrosine sites along the disordered C-Tir modified by Fyn (i.e., Y454, Y474, Y483, Y511). Phosphorylation caused small but noticeable changes in the chemical shifts of the tyrosines and adjacent residues. The [^{1}H-^{15}N]-HSQC spectrum displayed low amide proton dispersion, a diagnostic that the C-Tir remains disordered upon multisite phosphorylation. Moreover, the secondary chemical shifts of this 4-fold phosphorylated state (pC-Tir) did not reveal substantial changes in local structure propensity due to phosphorylation (**Fig. S7**), as similarly reported for other disordered proteins (79, 80).

To evaluate the interaction of phosphorylated C-Tir (pC-Tir) with an SH2 domain, we titrated unlabeled C-SH2 into a ^{15}N-labeled pC-Tir solution and monitored the binding interaction broadening of signals in NMR [^{1}H-^{15}N]-HSQC spectra (**Fig. 5*A,B***). With the resonance re-assignment of pC-Tir, we identified that upon phosphorylation, all tyrosine sites interact with C-SH2, and not exclusively the Y511-based motif as observed for unphosphorylated C-Tir. The interaction caused a drop in peak intensities around each site (i.e., i= pY454, pY474, pY483, pY511), including roughly residues i+4 and i-1, even at sub-stoichiometric conditions (**Fig. 5*B***). This suggests a "fuzzy" multivalent binding of C-SH2 to four xYx(x)ϕ motifs, where ϕ is hydrophobic (V/L/I) (**Fig. 5*C***), with multiple pYs on C-Tir interacting with C-SH2 in dynamic equilibrium (79). Except for these residues, backbone chemical shift changes were relatively small, with spectra displaying sharp line widths and low dispersion, both characteristics of a disordered protein. Thus, pC-Tir binds C-SH2 retaining its high level of disorder. The relative signal intensities at each phospho-tyrosine (pY) and surrounding residues approximately reflect the fraction of unbound sites, allowing us to estimate their apparent local binding affinity (**Fig. S8**). Phosphorylation enhanced the binding to Y511 ITIM by ~10 fold and enabled other tyrosine-based motifs to engage C-SH2. Among the four pYs, the resonances around pY454 were less broadened during the titration, suggesting lower local binding. The remaining pYs bind C-SH2 with similar strength (~3-9 μM). Overall, phosphorylation at multiple tyrosine sites provides various SH2 docking sites and reinforces the role of Tir as a scaffolding hub.

Tir's N-terminal intracellular tail is a dimer with order-disordered duality

In a similar fashion, we evaluated whether the N-terminal cytosolic region of Tir (N-Tir; **Fig S2**) was also intrinsically disordered. Yet, contrary to C-Tir, this region adopts a flexible but more defined structure. Based on SEC-SAXS data, the N-Tir is a non-globular protein with an *Rg* of 37.70 ± 0.10 Å and a *Dmax* of 140.0 ± 10.0 Å (**Fig. 6*A***) (**Table S4**), and a molecular weight compatible with a 52kDa dimer, in line with the observed gel filtration retention time. Like C-Tir, the respective *P(r)* curve is highly asymmetric but bimodal, with a peak at 21.9 Å and a prominent shoulder at 35.9 Å, suggesting that the dimer is elongated with spatially separated lobes. Its SAXS-derived Kratky plot is also not bell-shaped as expected for a non-globular and flexible arrangement (**Fig. 6*B***). The corresponding SAXS-driven *ab initio* reconstruction highlights an elongated S-shaped core for the dimeric N-Tir 1-233 polypeptide fused to a C-terminal Strep-Tag (**Fig. S4**). This extended dimer is largely disordered but contains stable secondary structural elements. Its far-UV CD profile reveals a partially folded protein, marked by a negative maximum at 205 nm due to high disorder content. The positive signal at 190 nm and the negative shoulder at 220 nm reflect ordered elements (**Fig. 6*C***). So, SAXS and CD data indicated that N-Tir is a partially folded dimer. To probe the order-disorder interplay within N-Tir at high-resolution, we have collected NMR data at different temperatures.

For globular proteins, the 52kDa dimer size is beyond the practical range amenable to traditional NMR spectroscopy in solution. The NMR signals of larger molecules relax faster, leading to line broadening, low spectral sensitivity, and eventually loss of NMR signals. This problem is even more significant at low temperatures. Nevertheless, we found that the 52kDa dimer has a [^{1}H-

^{15}N]-HSQC NMR spectrum at 5 ºC characterized by intense signals and low chemical shift dispersion akin to an IDP (**Fig. 6*D***). We assigned the resonances at 5 ºC mostly to the first 80 and last 35 residues of the N-Tir, meaning that those regions are flexible in solution. The missing resonances of the central residues became visible with increasing temperature due to faster tumbling. At 35 ºC, they adopted an NMR fingerprint of a folded-like protein (**Fig. 6*E***), reflecting a more well-structured region. With NMR, we could establish that N-Tir's central sequence has the broad NMR fingerprint of a well-folded protein, flanked by residues with poorly dispersed ^{1}H^{N} backbone resonances pointing disorder. A closer look at the structural disorder predictions computed for N-Tir also suggests an order-disorder duality, with flanking residues more flexible than those in the central region (**Fig. S9*A***). Such disordered segments harbor linear motifs that were experimentally identified to interact with host proteins (**FIg. S2B**) (51, 81).

The partial deletion of these disordered regions reduced the random-coil content reported by CD data, highlighting a more structured central part of N-Tir (N-Tir$_{60-200}$, hereafter NS-Tir). The difference between CD spectra resulted in a pure random-coil CD signature, reinforcing that the flanking regions are disordered, whereas the central part folded. NS-Tir's CD profile displays positive values below 200 nm and two negative bands at 210 and 222 nm, commonly associated with structured secondary conformations (**Fig. 6*C***).

We have employed SAXS to probe further the overall structure of NS-Tir. Our synchrotron SEC–SAXS data confirmed that NS-Tir is still a stable dimer in solution. A slightly less broadened Kratky profile reflects the absence of the disordered flanking regions (**Fig. 6*B***) but compatible with an elongated dimer (**Fig. 6*F***), retaining a similar overall conformation supported by NMR. We acquired [^{1}H-^{15}N]-HSQC data for NS-Tir and superimposed them to N-Tir spectra (**Fig. S9*B***). Over 75% of the visible resonances overlapped between the two constructs, with a low dispersion for the ^{1}H dimension at 5 ºC and a broader [^{1}H-^{15}N]-fingerprint at 35 ºC, indicating that NS-Tir maintained a similar folded isolated and in the full-length N-terminal region context. The reversible thermal denaturation of NS-Tir monitored by far-UV CD shows the dimer unfolding in two discrete steps, capturing both folding and dimerization events. It reveals a stable dimer that unfolds and dissociates by increasing temperature (82) (**Fig. S10**).

Overall, our results support a model of Tir spanning the host plasma membrane with flexible N-terminal intracellular domains assembled in a dimer form (**Fig 7**). In this model, N-Tir is anchored to the membrane by flexible segments and with dangling linear motifs located at its disordered N-termini.

Concluding Remarks

We observed that A/E pathogens have a structurally diverse repertoire of protein effectors enriched in disorder-prone residues. Some effectors are fully disordered proteins with a high density of putative host-like motifs that can mediate the interaction with human proteins. Among those highly disordered effectors emerged Tir, a 56 kDa membrane protein essential for virulence. Using multiple biophysical methods, we found that Tir-ICDs (i.e., N-Tir and C-Tir) are highly disordered with IDR features similar to those found in the cytoplasmic domains of host transmembrane proteins. NMR analysis showed that C-Tir is not exclusively disordered but can adopt partial secondary structural elements near phosphosites and motifs involved in versatile interactions with the host (i.e., protein and lipid-binding). The C-Tir displays cytoplasmic tyrosine-based motifs present in the disordered cytosolic tails of host receptors. The standard mode of action associated with these motifs is the phosphorylation of the central tyrosine by Src family PTKs, creating a binding site for SH2–containing proteins. In C-Tir, we found that one of those motifs, involving Y511, displays a pre-formed transient helical structure and binds the C-terminal SH2 domain of SHP-1 pre-phosphorylation, interacting as well with lipid bilayers (**Fig.3, Fig. 7**). These observations highlight a functional diversity for C-Tir and a possible cooperation between host proteins and the membrane in signaling processes inside the host (14). Some T-cell receptor cytoplasmic disordered domains associate with the plasma membrane via tyrosine-based motifs with a helical propensity, and such association regulates phosphorylation and downstream signaling (74). Like the CD3ε cytoplasmic domain of the T-cell receptor (73), C-Tir also has two tyrosines involved in lipid-binding and helical propensity. Our results suggest that the C-terminal half Tir can mimic this regulatory stratagem to fine-tune its ability to interact with human cell components, thereby interfering with normal cellular functions. Further studies are needed to corroborate this initial observation. Nevertheless, we show that the C-Tir can interact with membranes via ITMs and its last C-terminal residues. Importantly, our results established that this membrane affinity is residue-specific and modulated by lipid composition in a quantitative and site-resolved way, suggesting the existence of a regulation layer based on lipid composition. The flexibility and conformational plasticity of C-Tir make it readily accessible for phosphorylation. We found in C-Tir four *bona fide* tyrosine phosphorylation sites that render the ability for multivalent/promiscuous binding, supporting its ability to act as a signaling hub. Even though we only used protein fragments and explored one phosphorylation state, our results support the presence of four phospho-tyrosines (pY) in the C-Tir sequence coexisting simultaneously. All can bind the C-SH2 in the micromolar range as reported for other pY-containing motifs (84). Interestingly, the pY454 site, which displays the lowest affinity, is part of a conserved NPY motif (**Fig. S3**) that interacts with the I-BAR domain of host IRSp53/IRTKS, linking Tir to the actin polymerization machinery (77) in a phosphorylation-independent manner. This suggests that Y454 might act as a binding site for distinct and competing host proteins depending on its phosphorylation state.

The superposition of interplaying functional elements observed in C-Tir and their possible interplay mirror the mechanics of human intrinsically disordered domains. Transiently structured lipid-binding regions might bury tyrosine residues rendering them inaccessible, while their phosphorylation perturbs membrane anchoring (85). At the same time, these multiple modification sites promiscuously participate in fuzzy binding of host partners, their respective affinities being interdependent with the phosphorylation state. Future work using high-resolution structural studies targeting the interaction with multiple partners and different C-Tir's phosphorylation states will provide additional insights into Tir binding specificity and action mode.

We unraveled for the first time that the N-terminal region, yet partially disordered, forms a stable dimer with potential implications on Tir self-assembly and signaling. Clustering of Tir triggers host signaling events when binding to intimin. The extracellular intimin binding domain (IBD) of Tir binds intimin as a dimer (**Fig. 7**) (36) in a reticulating model (83). Thus, Tir's intracellular self-assembly opens the question of whether host signaling activation involves intimin-induced dimerization or structural changes of preformed dimers.

Together, these novel structural insights provide an updated picture of Tir's intracellular side (**Fig. 7**), laying the path towards illustrating how Tir hijacks host signaling. Importantly, yet less abundant in prokaryotes, the structural disorder of Tir and several other A/E effectors reinforce the idea of a positive evolutionary selection towards disordered proteins in the pool of secreted effectors by bacterial pathogens to target host cellular machinery (20).

Methods and Materials

Disorder prediction and short linear motif analysis

We collected the sequence of A/E pathogen effectors and corresponding reference proteomes from the UniprotKB database (86) (**Table S1**). Three effector sets were assembled for the following strains: EHEC, EPEC, and CR. The representative set of 20,365 human proteins was extracted from UniprotKB human reference proteome. To compute disorder propensity at the residue level, we used DISOPRED 3 (41) and IUPred 1.0 ("long" mode prediction) (42), structural disorder predictors. Disordered residues were defined as those with a propensity score equal to or above 0.5. We used this metric to calculate the fraction of disordered residues for each protein. The one-sided Mann-WhitneyU-test (87) was used to compare disorder fraction distribution of the effector collections and their corresponding proteomes (**SI Materials and Methods**). Both per-residue scores and aggregated disorder fractions were used to classify each protein according to the structural categories adapted from (88) (IDP: Intrinsically disordered proteins; PDR: Proteins with intrinsically disordered regions; FRAG: Proteins with fragmented-disorder; NDR: Not disordered proteins; ORD: Ordered Proteins). See **SI Materials and Methods** and **Table S2** for a detailed description of the criteria used to define the structural categories. The list of eukaryotic short linear motifs (ELMs) was downloaded from the ELM database (89) (version 1.4, May 2017). We searched for occurrences of putative ELMs in effector sequences using the ANCHOR tool (90). For each effector sequence, we calculated the motif density as the fraction between the amino acids belonging to putative ELMs predicted as disordered and the total number of amino acids predicted as disordered.

Protein expression and purification

All constructs (C-Tir, C-SH2, N-Tir, and NS-Tir) were subcloned into the pHTP8 plasmid (NZYTech), bearing a cleavable His6-tagged thioredoxin tag (TRX-His_6) for improved solubility and folding. They were successfully expressed using *E. coli* BL21 Star (DE3) pLysS and purified by affinity chromatography and Size-Exclusion gel filtration. More details in **SI Material and Methods.**

NMR spectroscopy

All NMR experiments were recorded on an 800 MHz Bruker Avance II+ spectrometer equipped with a cryoprobe. The backbone assignment was performed combinating standard triple-resonance experiments and reverse labeling (61). NMR is well suited to study such dynamic and transient interactions in solution at single-residue resolution. We extensively explored the sensitivity of NMR to changes in the chemical environment to monitor phosphorylation, protein- and lipid-binding. More details in **SI Material and Methods.**

Small-angle X-ray scattering (SAXS)

We employed synchrotron SAXS coupled with size exclusion chromatography (SEC) to probe the overall size and conformational properties of Tir intracellular regions. A comprehensive description of SAXS measurements and analysis is described in detail in **SI Material and Methods.**

Bicelles preparation

Bicelles with different charge density were prepared based on the protocols described in refs (76, 91). See **SI Material and Methods** for additional details.

Tyrosine phosphorylation

Fully phosphorylation of four tyrosines in C-Tir was achieved by incubating the protein with human recombinant Fyn Kinase overnight (71). Optimized experimental conditions are detailed on **SI Material and Methods.**

Data and software availability

The NMR chemical shifts of C-Tir and pC-Tir are available at the Biological Magnetic Resonance (BMR) data bank under the entry 50758, 50759, respectively. The SEC-SAXS data can be found at the Small Angle Scattering Biological Data Bank (SASBDB) under the project "SAXS studies on the intracellular region of the translocated intimin receptor". The accession codes are SASDKF8, SASDKG8 and SASDKH8. Proteome and effector collections, disorder predictions, derived aggregated data and figures are available for download https://osf.io/3mka9/. The associated code is available at https://osf.io/cxkjf/.

Acknowledgements

MFMV is funded by a MolBioS FCT PhD program fellowship (PD/BD/135482/2018). National funds financially support GH through the PT-NMR FCT PhD Program (PD/BD/147227/2019). TNC is recipient of the grant CEECIND/01443/2017. This work was funded by national funds through FCT– Fundação para a Ciência e a Tecnologia, I.P. (FCT), Project MOSTMICRO-ITQB with refs

UIDB/04612/2020 and UIDP/04612/2020 and FEDER Funds through COMPETE 2020 (0145-FEDER-007660). We acknowledge using the ESRF-BM29 (MX-2085-BAG) and DLS-B21 Bio-SAXS beamlines (MX20161-1, SM21035-177). We thank Miquel Pons, from UB, and Lígia Martins, and Adriano Henriques, from ITQB-NOVA, for many helpful discussions. We thank Isabel Pacheco, João Carita, Cristina Timoteo, and Teresa Silva (Research Facilities, ITQB-NOVA) for their technical assistance. MA thanks Prof. Hartmut Oschkinat and FMP-Berlin for the financial support. AZ acknowledges the financial support of the JPI HDHL-INTIMIC action cofunded by the Agence Nationale de la Recherche (ANR-17-HDIM-0001-03) and thanks Fabrice Lopez and Aurélie Bergon (TAGC, Marseille) for technical assistance. NMR data were acquired at CERMAX, ITQB-NOVA, Oeiras, Portugal, with equipment funded by FCT, project AAC 01/SAICT/2016.

Author contributions

Conceptualization: TNC

Methodology: MA, AZ, TNC

Investigation: MFMV, GH, TV, MA, AZ, TNC

Supervision: TNC

Writing—review & editing: MA, AZ, TNC

Competing interests

Authors declare that they have no competing interests.

References

1. Z. Peng, *et al.*, Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. *Cell. Mol. Life Sci.* **72**, 137–151 (2015).

2. M. M. Babu, R. W. Kriwacki, R. V. Pappu, Structural biology. Versatility from protein disorder. *Science* **337**, 1460–1461 (2012).

3. V. N. Uversky, Intrinsically disordered proteins and their "mysterious" (meta)physics. *Front. Phys.* **7** (2019).

4. P. E. Wright, H. J. Dyson, Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).

5. C. A. Galea, Y. Wang, S. G. Sivakolundu, R. W. Kriwacki, Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits.

Biochemistry **47**, 7598–7609 (2008).

6. J. F. Pelham, J. C. Dunlap, J. M. Hurley, Intrinsic disorder is an essential characteristic of components in the conserved circadian circuit. *Cell Commun. Signal.* **18**, 181 (2020).
7. A. Bah, J. D. Forman-Kay, Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *J. Biol. Chem.* **291**, 6696–6705 (2016).
8. R. van der Lee, *et al.*, Intrinsically disordered segments affect protein half-life in the cell and during evolution. *Cell Rep.* **8**, 1832–1844 (2014).
9. V. N. Uversky, Wrecked regulation of intrinsically disordered proteins in diseases: pathogenicity of deregulated regulators. *Front. Mol. Biosci.* **1**, 6 (2014).
10. N. E. Davey, *et al.*, Attributes of short linear motifs. *Mol. Biosyst.* **8**, 268–281 (2012).
11. K. Van Roey, *et al.*, Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem. Rev.* **114**, 6733–6778 (2014).
12. V. Csizmok, A. V. Follis, R. W. Kriwacki, J. D. Forman-Kay, Dynamic protein interaction networks and new structural paradigms in signaling. *Chem. Rev.* **116**, 6424–6462 (2016).
13. N. Jespersen, E. Barbar, Emerging Features of Linear Motif-Binding Hub Proteins. *Trends Biochem. Sci.* **45**, 375–384 (2020).
14. J. Cornish, S. G. Chamberlain, D. Owen, H. R. Mott, Intrinsically disordered proteins and membranes: a marriage of convenience for cell signalling? *Biochem. Soc. Trans.* **48**, 2669–2689 (2020).
15. E. Deryusheva, *et al.*, Does Intrinsic Disorder in Proteins Favor Their Interaction with Lipids? *Proteomics* **19**, e1800098 (2019).
16. J. Bürgi, B. Xue, V. N. Uversky, F. G. van der Goot, Intrinsic disorder in transmembrane proteins: roles in signaling and topology prediction. *PLoS ONE* **11**, e0158594 (2016).
17. M. Kjaergaard, B. B. Kragelund, Functions of intrinsic disorder in transmembrane proteins. *Cell. Mol. Life Sci.* **74**, 3205–3224 (2017).
18. W. Basile, M. Salvatore, C. Bassot, A. Elofsson, Why do eukaryotic proteins contain more intrinsically disordered regions? *PLoS Comput. Biol.* **15**, e1007186 (2019).
19. B. Xue, A. K. Dunker, V. N. Uversky, Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* **30**, 137–149 (2012).
20. M. Marín, V. N. Uversky, T. Ott, Intrinsic disorder in pathogen effectors: protein flexibility as an evolutionary hallmark in a molecular arms race. *Plant Cell* **25**, 3153–3157 (2013).
21. D. P. O'Brien, *et al.*, Calmodulin fishing with a structurally disordered bait triggers CyaA catalysis. *PLoS Biol.* **15**, e2004486 (2017).
22. O. Aitio, *et al.*, Enterohaemorrhagic Escherichia coli exploits a tryptophan switch to hijack host f-actin assembly. *Structure* **20**, 1692–1703 (2012).

23. D. Hamada, M. Hamaguchi, K. N. Suzuki, I. Sakata, I. Yanagihara, Cytoskeleton-modulating effectors of enteropathogenic and enterohemorrhagic Escherichia coli: a case for EspB as an intrinsically less-ordered effector. *FEBS J.* **277**, 2409–2415 (2010).

24. T. H. Hazen, *et al.*, Genomic diversity of EPEC associated with clinical presentations of differing severity. *Nat. Microbiol.* **1**, 15014 (2016).

25. H. Pennington, Escherichia coli O157. *Lancet* **376**, 1428–1435 (2010).

26. W. Deng, *et al.*, Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proc Natl Acad Sci USA* **101**, 3597–3602 (2004).

27. T. Tobe, *et al.*, An extensive repertoire of type III secretion effectors in Escherichia coli O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci USA* **103**, 14941–14946 (2006).

28. E. Mills, K. Baruch, X. Charpentier, S. Kobi, I. Rosenshine, Real-time analysis of effector translocation by the type III secretion system of enteropathogenic Escherichia coli. *Cell Host Microbe* **3**, 104–113 (2008).

29. B. Kenny, *et al.*, Enteropathogenic E. coli (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* **91**, 511–520 (1997).

30. S. E. Battle, M. J. Brady, S. K. Vanaja, J. M. Leong, G. A. Hecht, Actin pedestal formation by enterohemorrhagic Escherichia coli enhances bacterial host cell attachment and concomitant type III translocation. *Infect. Immun.* **82**, 3713–3722 (2014).

31. S. Gruenheid, *et al.*, Enteropathogenic E. coli Tir binds Nck to initiate actin pedestal formation in host cells. *Nat. Cell Biol.* **3**, 856–859 (2001).

32. Y. Xue, M. Du, H. Sheng, C. J. Hovde, M.-J. Zhu, Escherichia coli O157:H7 suppresses host autophagy and promotes epithelial adhesion via Tir-mediated and cAMP-independent activation of protein kinase A. *Cell Death Discov.* **3**, 17055 (2017).

33. D. Yan, X. Wang, L. Luo, X. Cao, B. Ge, Inhibition of TLR signaling by a bacterial protein containing immunoreceptor tyrosine-based inhibitory motifs. *Nat. Immunol.* **13**, 1063–1071 (2012).

34. P. J. Goddard, *et al.*, Enteropathogenic Escherichia coli Stimulates Effector-Driven Rapid Caspase-4 Activation in Human Macrophages. *Cell Rep.* **27**, 1008-1017.e6 (2019).

35. Q. Zhong, *et al.*, Clustering of Tir during enteropathogenic E. coli infection triggers calcium influx-dependent pyroptosis in intestinal epithelial cells. *PLoS Biol.* **18**, e3000986 (2020).

36. Y. Luo, *et al.*, Crystal structure of enteropathogenic Escherichia coli intimin-receptor complex. *Nature* **405**, 1073–1077 (2000).

37. A. R. C. Wong, *et al.*, Enteropathogenic and enterohaemorrhagic Escherichia coli: even more subversive elements. *Mol. Microbiol.* **80**, 1420–1438 (2011).

38. S. Blasche, *et al.*, The EHEC-host interactome reveals novel targets for the translocated

intimin receptor. *Sci. Rep.* **4**, 7531 (2014).

39. D. Yan, *et al.*, Enteropathogenic Escherichia coli Tir recruits cellular SHP-2 through ITIM motifs to suppress host immune response. *Cell. Signal.* **25**, 1887–1894 (2013).

40. R. Zhou, *et al.*, Enterohemorrhagic Escherichia coli Tir inhibits TAK1 activation and mediates immune evasion. *Emerg. Microbes Infect.* **8**, 734–748 (2019).

41. D. T. Jones, D. Cozzetto, DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **31**, 857–863 (2015).

42. Z. Dosztányi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433–3434 (2005).

43. B. Mészáros, G. Erdos, Z. Dosztányi, IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **46**, W329–W337 (2018).

44. A. M. Grishin, *et al.*, NleH defines a new family of bacterial effector kinases. *Structure* **22**, 250–259 (2014).

45. X. Gao, *et al.*, Bacterial effector binding to ribosomal protein s3 subverts NF-kappaB function. *PLoS Pathog.* **5**, e1000708 (2009).

46. A. Holmes, S. Mühlen, A. J. Roe, P. Dean, The EspF effector, a bacterial pathogen's Swiss army knife. *Infect. Immun.* **78**, 4445–4453 (2010).

47. N. A. Sallee, *et al.*, The pathogen protein EspF(U) hijacks actin polymerization using mimicry and multivalency. *Nature* **454**, 1005–1008 (2008).

48. T. Hagai, A. Azia, M. M. Babu, R. Andino, Use of host-like peptide motifs in viral proteins is a prevalent strategy in host-virus interactions. *Cell Rep.* **7**, 1729–1739 (2014).

49. A. Via, B. Uyar, C. Brun, A. Zanzoni, How pathogens use linear motifs to perturb host cell networks. *Trends Biochem. Sci.* **40**, 36–48 (2015).

50. N. E. Davey, G. Travé, T. J. Gibson, How viruses hijack cell regulation. *Trends Biochem. Sci.* **36**, 159–169 (2011).

51. B. Bommarius, *et al.*, Enteropathogenic Escherichia coli Tir is an SH2/3 ligand that recruits and activates tyrosine kinases required for pedestal formation. *Mol. Microbiol.* **63**, 1748–1768 (2007).

52. P. Tompa, N. E. Davey, T. J. Gibson, M. M. Babu, A million peptide motifs for the molecular biologist. *Mol. Cell* **55**, 161–169 (2014).

53. L. Staby, *et al.*, Eukaryotic transcription factors: paradigms of protein intrinsic disorder. *Biochem. J.* **474**, 2509–2532 (2017).

54. C. Haynes, *et al.*, Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol.* **2**, e100 (2006).

55. P. Bernadó, M. Blackledge, A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. *Biophys. J.* **97**, 2839–2845 (2009).

56. T. N. Cordeiro, *et al.*, Structural Characterization of Highly Flexible Proteins by Small-Angle Scattering. *Adv. Exp. Med. Biol.* **1009**, 107–129 (2017).

57. A. G. Kikhney, D. I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).

58. L. B. Chemes, L. G. Alonso, M. G. Noval, G. de Prat-Gay, Circular dichroism techniques for the analysis of intrinsically disordered proteins and domains. *Methods Mol. Biol.* **895**, 387–404 (2012).

59. E. Delaforge, T. N. Cordeiro, P. Bernadó, N. Sibille, "Conformational characterization of intrinsically disordered proteins and its biological significance" in *Modern Magnetic Resonance*, G. A. Webb, Ed. (Springer International Publishing, 2018), pp. 381–399.

60. J. Garmendia, *et al.*, TccP is an enterohaemorrhagic Escherichia coli O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cell. Microbiol.* **6**, 1167–1183 (2004).

61. C. Prasanna, A. Dubey, H. S. Atreya, Amino acid selective unlabeling in protein NMR spectroscopy. *Meth. Enzymol.* **565**, 167–189 (2015).

62. K. Tamiola, B. Acar, F. A. A. Mulder, Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* **132**, 18000–18003 (2010).

63. M. Fuxreiter, I. Simon, P. Friedrich, P. Tompa, Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol.* **338**, 1015–1026 (2004).

64. A. De Biasio, *et al.*, p15PAF is an intrinsically disordered protein with nonrandom structural preferences at sites of interaction with other proteins. *Biophys. J.* **106**, 865–874 (2014).

65. K. Tamiola, F. A. A. Mulder, Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. *Biochem. Soc. Trans.* **40**, 1014–1020 (2012).

66. M.-H. Cho, J. O. Wrabl, J. Taylor, V. J. Hilser, Hidden dynamic signatures drive substrate selectivity in the disordered phosphoproteome. *Proc Natl Acad Sci USA* **117**, 23606–23616 (2020).

67. N. Stanley, S. Esteban-Martín, G. De Fabritiis, Kinetic modulation of a disordered protein domain by phosphorylation. *Nat. Commun.* **5**, 5272 (2014).

68. A. Hawrani, *et al.*, Effect of protein kinase A-mediated phosphorylation on the structure and association properties of the enteropathogenic Escherichia coli Tir virulence protein. *J. Biol. Chem.* **278**, 25839–25846 (2003).

69. P. R. Race, A. S. Solovyova, M. J. Banfield, Conformation of the EPEC Tir protein in

solution: investigating the impact of serine phosphorylation at positions 434/463. *Biophys. J.* **93**, 586–596 (2007).

70. S. Backert, B. Kenny, R. Gerhard, N. Tegtmeyer, S. Brandt, PKA-mediated phosphorylation of EPEC-Tir at serine residues 434 and 463: A novel pathway in regulating Rac1 GTPase function. *Gut Microbes* **1**, 94–99 (2010).

71. J. Rosenlöw, L. Isaksson, M. Mayzel, J. Lengqvist, V. Y. Orekhov, Tyrosine phosphorylation within the intrinsically disordered cytosolic domains of the B-cell receptor: an NMR-based structural analysis. *PLoS ONE* **9**, e96199 (2014).

72. H. Cheng, *et al.*, Conformational Changes in the Cytoplasmic Region of KIR3DL1 upon Interaction with SHP-2. *Structure* **27**, 639-650.e2 (2019).

73. C. Xu, *et al.*, Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. *Cell* **135**, 702–713 (2008).

74. X. Xu, H. Li, C. Xu, Structural understanding of T cell receptor triggering. *Cell. Mol. Immunol.* **17**, 193–202 (2020).

75. E. Duchardt, A. B. Sigalov, D. Aivazian, L. J. Stern, H. Schwalbe, Structure induction of the T-cell receptor zeta-chain upon lipid binding investigated by NMR spectroscopy. *Chembiochem* **8**, 820–827 (2007).

76. Y. Pérez, *et al.*, Lipid binding by the Unique and SH3 domains of c-Src suggests a new regulatory mechanism. *Sci. Rep.* **3**, 1295 (2013).

77. J. C. de Groot, *et al.*, Structural basis for complex formation between human IRSp53 and the translocated intimin receptor Tir of enterohemorrhagic E. coli. *Structure* **19**, 1294–1306 (2011).

78. F.-X. Theillet, *et al.*, Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat. Protoc.* **8**, 1416–1432 (2013).

79. T. Mittag, *et al.*, Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci USA* **105**, 17772–17777 (2008).

80. A. L. Turner, *et al.*, Highly disordered histone H1-DNA model complexes and their condensates. *Proc Natl Acad Sci USA* **115**, 11964–11969 (2018).

81. A. Patel, *et al.*, Host protein interactions with enteropathogenic Escherichia coli (EPEC): 14-3-3tau binds Tir and has a role in EPEC-induced actin polymerization. *Cell. Microbiol.* **8**, 55–71 (2006).

82. C. Rat, J. C. Heiby, J. P. Bunz, H. Neuweiler, Two-step self-assembly of a spider silk molecular clamp. *Nat. Commun.* **9**, 4779 (2018).

83. T. Touzé, R. D. Hayward, J. Eswaran, J. M. Leong, V. Koronakis, Self-association of EPEC intimin mediated by the beta-barrel-containing anchor domain: a role in clustering of the Tir receptor. *Mol. Microbiol.* **51**, 73–87 (2004).

84. M. J. Wagner, M. M. Stacey, B. A. Liu, T. Pawson, Molecular mechanisms of SH2- and PTB-domain-containing proteins in receptor tyrosine kinase signaling. *Cold Spring Harb. Perspect. Biol.* **5**, a008987 (2013).

85. M. Arbesú, G. Iruela, H. Fuentes, J. M. C. Teixeira, M. Pons, Intramolecular fuzzy interactions involving intrinsically disordered domains. *Front. Mol. Biosci.* **5**, 39 (2018).

86. UniProt Consortium, UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).

87. H. B. Mann, D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18**, 50–60 (1947).

88. A. Giansanti, A. Deiana, S. Forcelloni, A. Porrello, New classification of intrinsic disorder in the human proteome. *BioRxiv* (2018) https:/doi.org/10.1101/446351.

89. H. Dinkel, *et al.*, ELM 2016--data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Res.* **44**, D294-300 (2016).

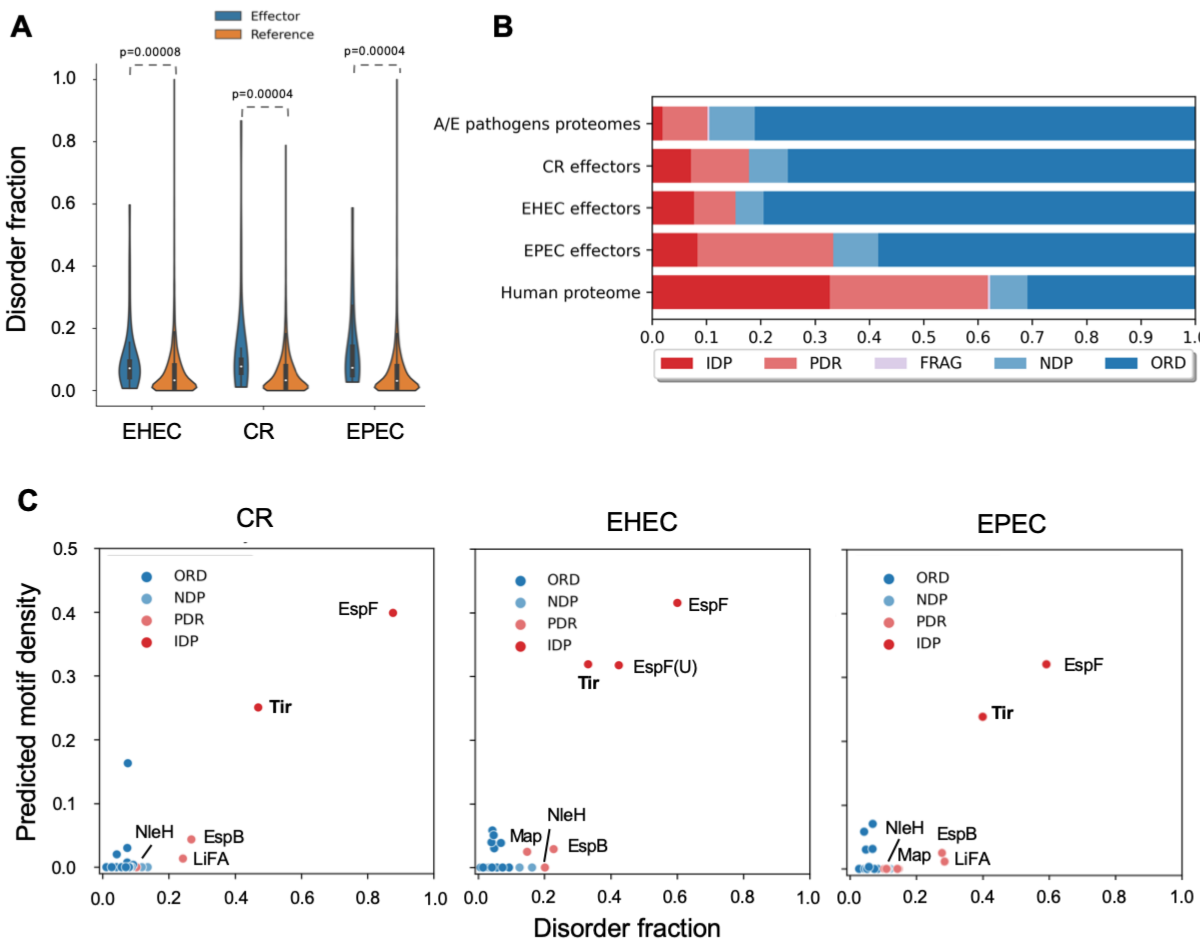
90. B. Mészáros, I. Simon, Z. Dosztányi, Prediction of protein binding regions in disordered proteins. *PLoS Comput. Biol.* **5**, e1000376 (2009).

91. M. Maffei, *et al.*, Lipid binding by disordered proteins. *Protoc. exch.* (2013) https:/doi.org/10.1038/protex.2013.094.

92. D. I. Svergun, M. V. Petoukhov, M. H. Koch, Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80**, 2946–2953 (2001).

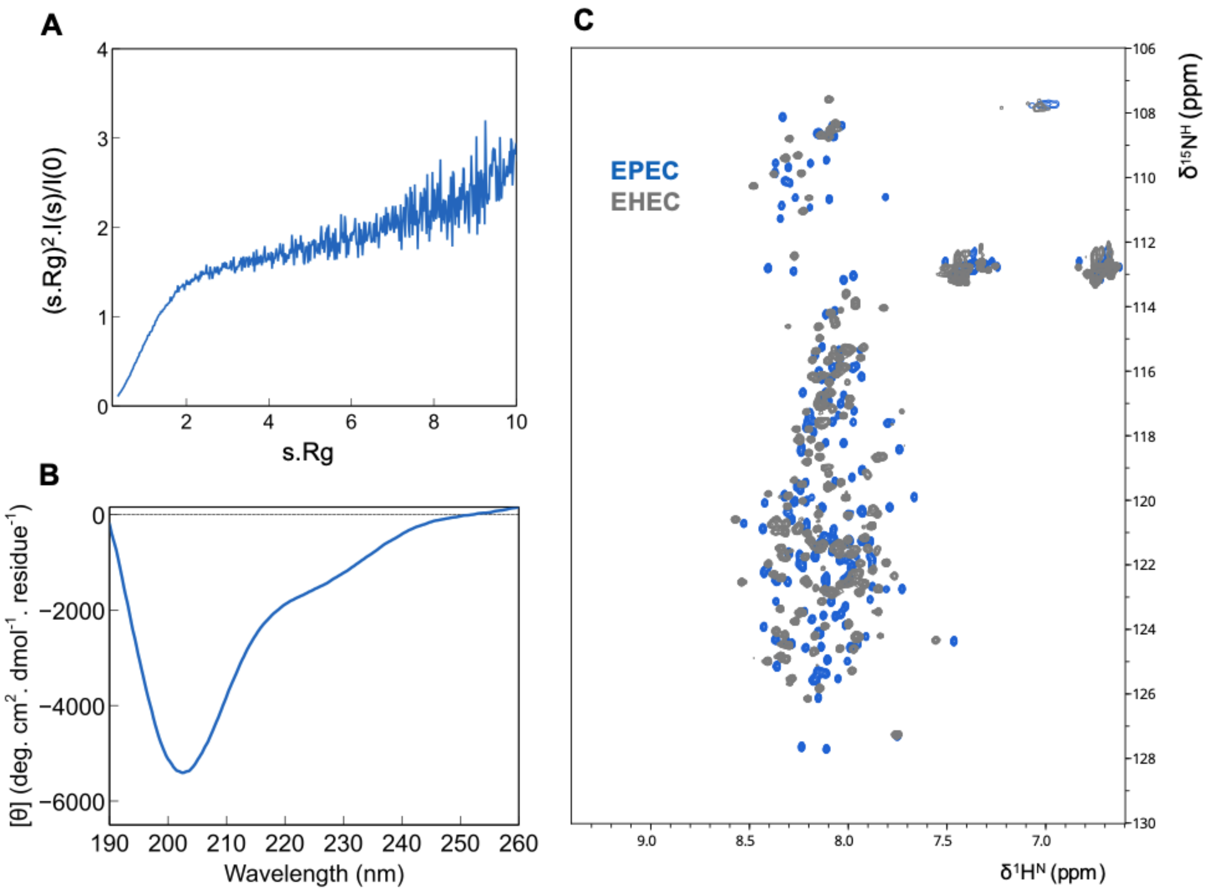
Figures

Fig. 1 - Predicted structural disorder in A/E pathogens.



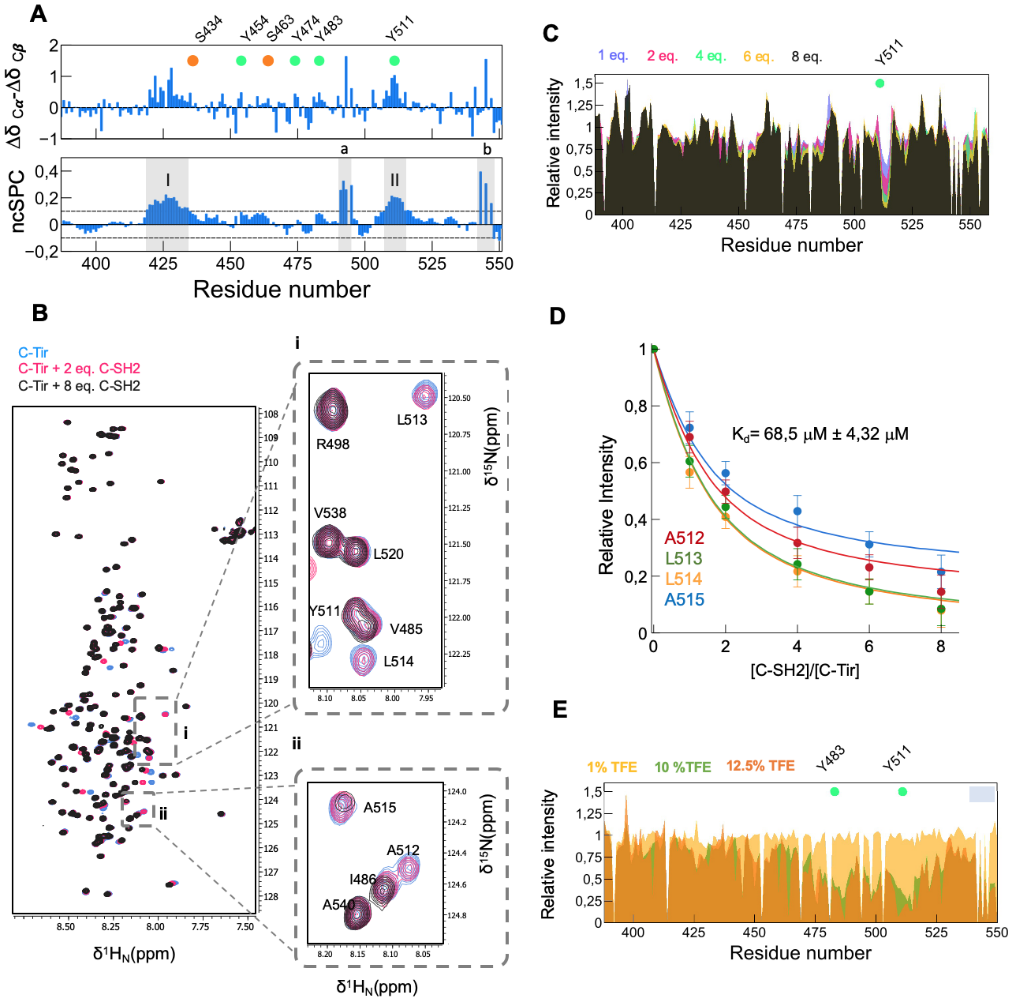
(*A*) Distribution of disorder fraction displayed as violin plots for effectors (blue) and full-proteomes (red) of A/E pathogens. (*B*) Accumulated fractions of the structural categories in terms of structural disorder. IDP: Intrinsically disordered proteins; PDR: Proteins with intrinsically disordered regions; FRAG: Proteins with fragmented-disorder; NDR: Not disordered proteins; ORD: Ordered Proteins. See full definitions in **Table S2** and the individual A/E pathogen proteomes in **Fig. S2**. (*C*) Fraction of ELMs vs. disorder fraction in A/E effectors. Tir, EspF, and EspF(U) display a high motif content within disordered regions. IDP and PDR effectors are labeled. See all data in **Table S3.**

Fig. 2 - C-Tir is an intrinsically disordered domain.



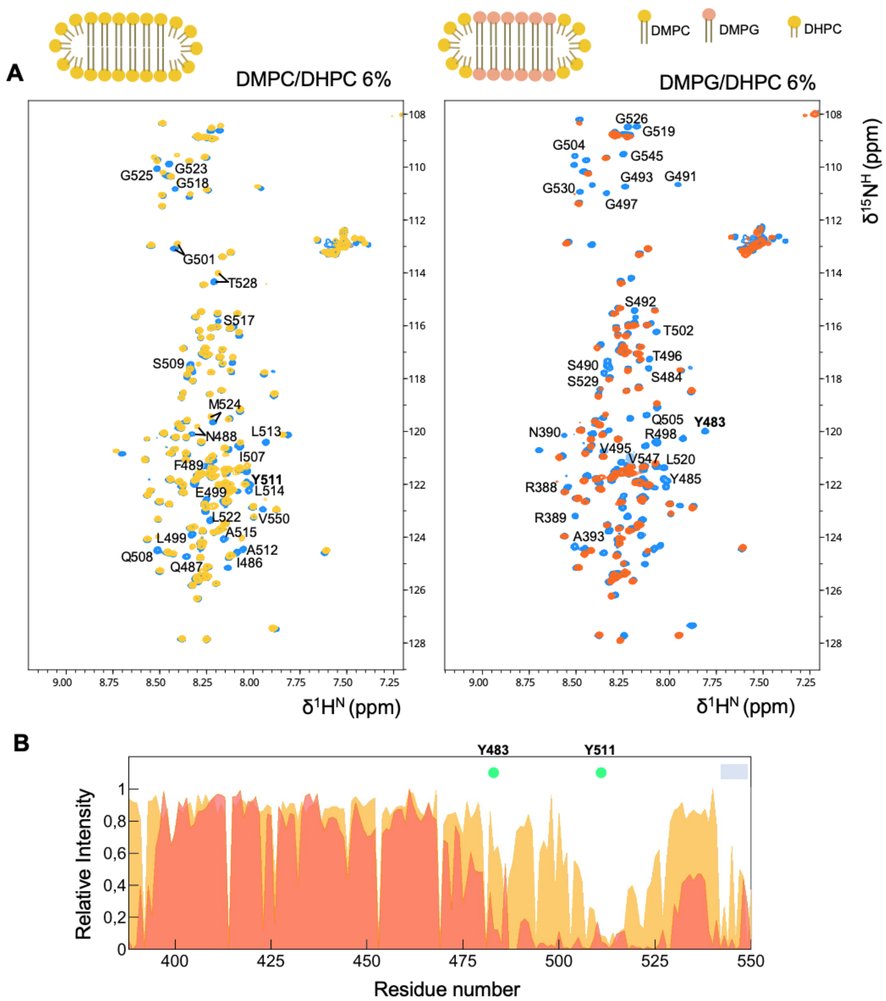
(*A*) Kratky representation of the SAXS profile measured for C-Tir indicating its lack of compactness. (*B*) Far-UV circular dichroism of C-Tir revealing the absence of high-populated structural secondary elements. (*C*) C-Tir EPEC (light blue) and C-Tir EHEC (grey) [^{1}H-^{15}N]-HSQC spectra reveal the narrow ^{1}H chemical shift dispersion characteristic of IDPs with ^{1}H amide backbone resonances clustering between 7.7 and 8.5 ppm. All together provide a definite diagnostic of protein disorder.

Fig. 3 - C-Tir is disordered with residual secondary structure at sites of host interaction.



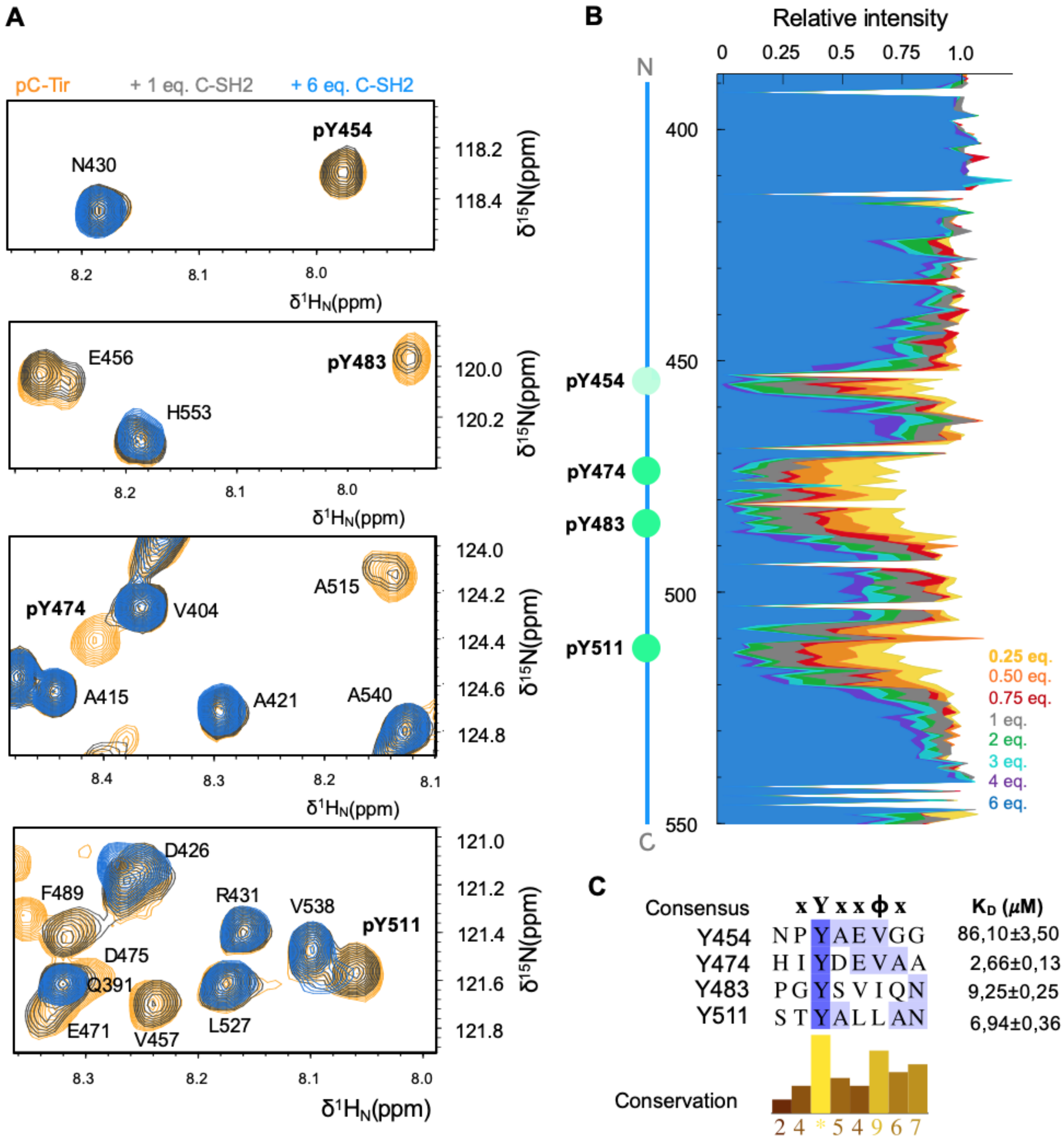
(*A*) Secondary chemical shifts ($\Delta\delta C\alpha$-$\Delta\delta C\beta$) (top) and structural propensity plot (bottom) for C-Tir where the dashed-lines depicts the random-coil threshold. Values above or below this threshold have α-helix or β-sheet propensities, respectively. Positive values (grey regions) show an increased α-helix tendency in regions D420-Q435 (I) and I507-A515 (II), as well as (a) F489-V495 and (b) T543-V547. Green and orange circles mark phosphorylation sites on C-Tir for tyrosine and serine residues, respectively. (*B*) Overlay of [^{1}H-^{15}N]-HSQC spectra of C-Tir in the absence (blue) and the presence of C-SH2 at 2.0 (pink) and 8.0 (dark grey) equivalents. Expansions (right panels) show the NMR signal of residues $A_{512}LLA_{515}$ with a significant intensity drop due to C-SH2 binding. (*C*) Relative [^{1}H-^{15}N]-peak intensities from the titration analysis. (*D*) Global fitting NMR quantification of the binding of C-SH2 to unphosphorylated C-Tir. (*E*) Residue-resolution mapping of the effect of TFE on C-Tir. TFE-induced NMR attenuation profiles, i.e., the ratio of peak intensity in the presence and absence of TFE, are plotted along the sequence. Green circles mark the position of Y483 and Y511. The grey bar highlights the C-terminal residues with NMR-signal attenuation.

Fig. 4 - Lipid-binding by disordered C-Tir.



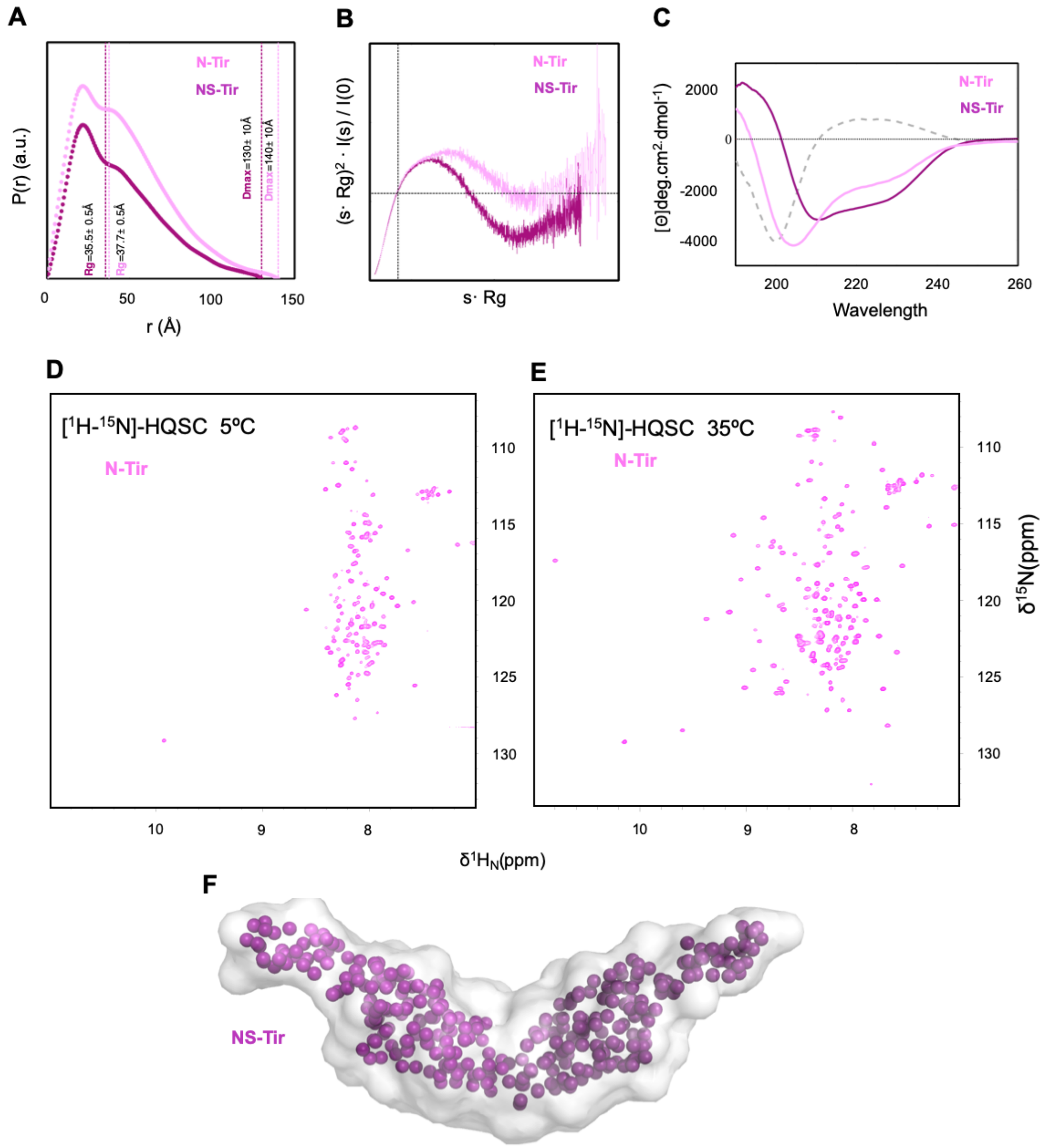
(*A*) NMR [^{1}H-^{15}N]-HSQC spectra of C-Tir in the absence (blue) and the presence of DMPC/DHPC (6% w/v), yellow) or DMPG/DHPC (6% w/v), orange) bicelles. C-Tir's HSQC data in bicelles superimpose well to free C-Tir, with identical chemical shifts for most visible resonances, pointing to intrinsic disorder. Residues with NMR signal attenuation upon addition of DMPC/DHPC bicelles are indicated in the spectra (left). The additional residues affected by DMPG/DHPC are also shown (right). (*B*) Lipid-induced NMR attenuation profiles are plotted against C-Tir primary sequence. Residues comprising the Y483 and Y511 ITIM motifs became less intense in lipids' presence, reinforcing their inherent ability to bind lipids. The grey bar highlights the C-terminal residues affected by lipids.

Fig. 5 - C-Tir contains multiple phosphorylated docking sites.



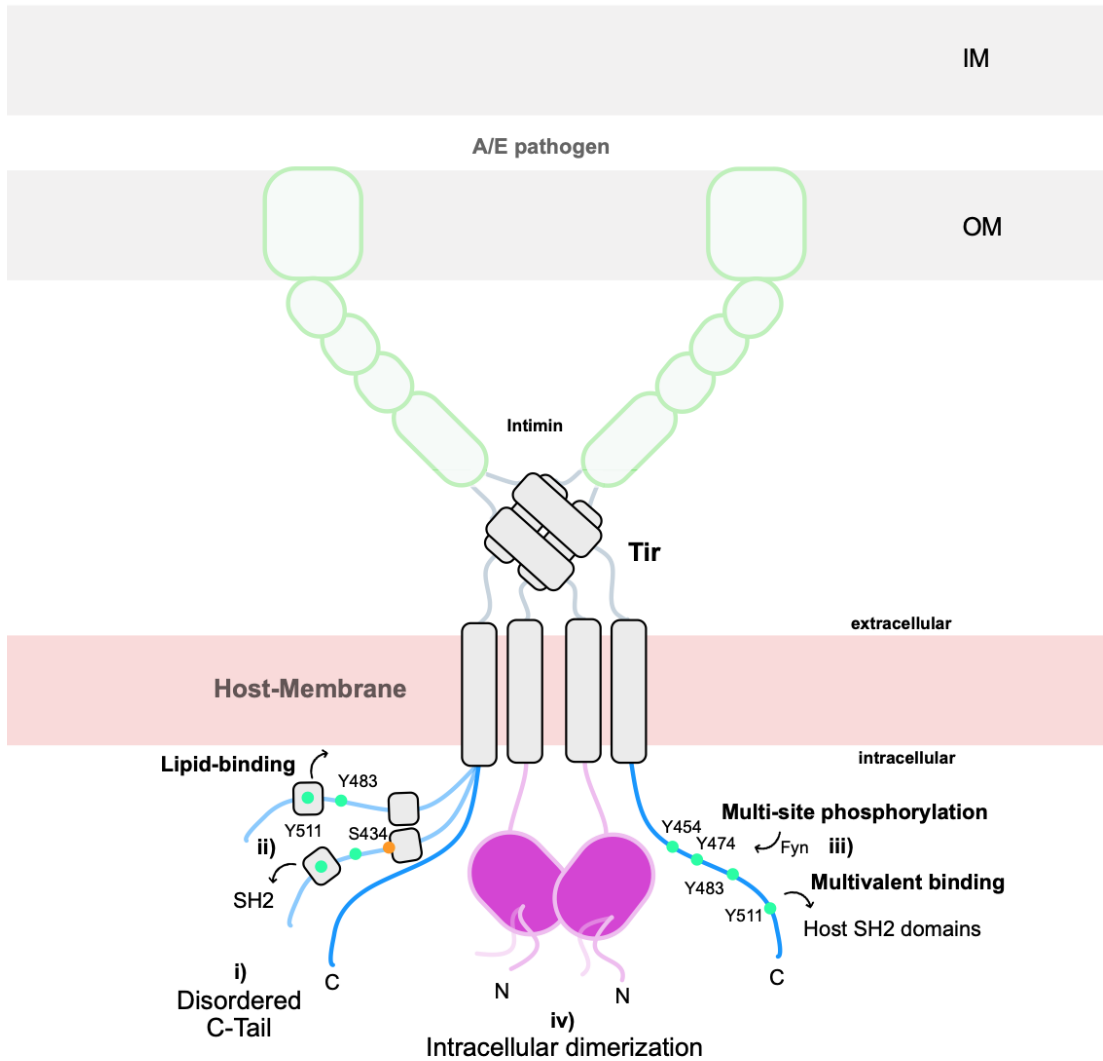
(*A*) NMR-detected binding of C-SH2 to phosphotyrosine pY454, pY474, pY483, pY511 before (orange) and after adding 1.0 and 6.0 equivalents of C-SH2 (blue and grey, respectively). (*B*) Site-specific NMR signal attenuation on pC-Tir due to C-SH2 binding, with local intensity drop around each phosphorylation site. (*C*) Sequence alignment of xYx(x)ϕ motifs of C-Tir interacting with C-SH2 and their respective apparent K_D values obtained by global fitting 1:1 model the local intensity drop around each phosphorylation site. Darker shades of blue refer to higher sequence identity. The conservation of physicochemical properties in each alignment position is reported in the corresponding barplot below the alignment. Green circles denote phosphorylation tyrosine-sites with color intensity reflecting the apparent relative affinity.

Fig. 6 - N-Tir is a partially disordered dimer.

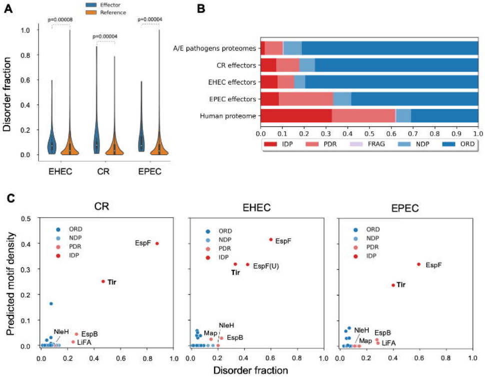


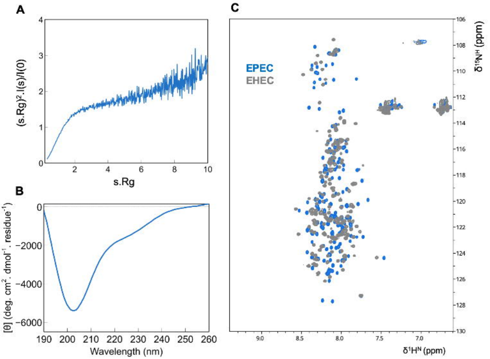
(*A*) Normalized pairwise distance distribution, P(r), computed from experimental SAXS curves of N-Tir (pink) and NS-Tir (purple) consistent with an elongated dimer. Dashed lines indicate the derived Rg and Dmax values. (*B*) Kratky plots of N-Tir (pink) and NS-Tir (dark-magenta), highlighting a high conformationally flexibility for N-Tir. (*C*) Far-CD spectra of N-Tir (pink) and NS-Tir (dark-magenta) and their difference (grey dashed line). (*D*) [^{1}H-^{15}N]-HSQC spectra of N-Tir at 5 ºC and (*E*) 35 ºC revealing two dynamically different regions within the protein. (*F*) NS-tir dimer at low-resolution driven from SAXS with GASBOR (92).

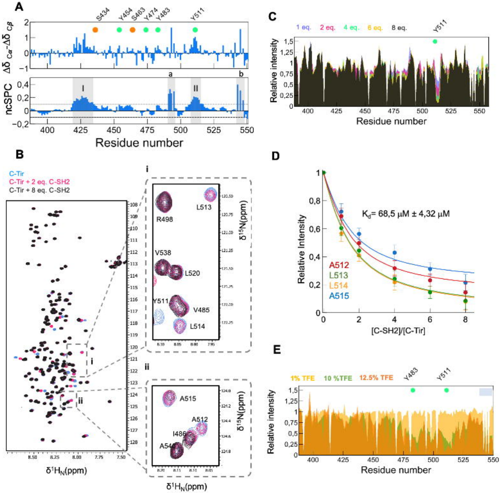
Fig. 7 - A proposed model of Tir structural organization.



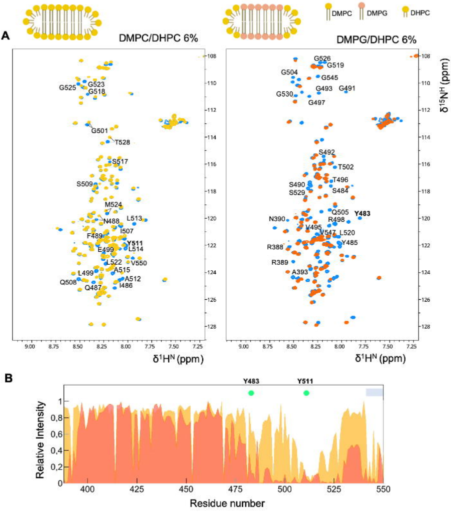
Tir in the host plasma membrane is depicted in a hairpin-topology while bound to intimin at the bacterial outer membrane. Our data support that (i) the C-terminal intracellular region of Tir (C-Tir) membrane receptor is disordered; (ii) with transient helical structural elements involved protein- and lipid-interactions; (iii) and host-like multi-phosphorylation sites that serve for docking various host SH2-domains. (iv) The intracellular N-terminal region of Tir (N-Tir) is partially disordered with a folded domain that self-assemble into dimer flanked by disordered residues, opening the question of whether host signaling activation involves intimin-induced dimerization or structural changes of preformed dimers. Green and orange circles denoted tyrosine and serine phosphorylation sites, respectively.


A
Effector
Reference
p=0.00008
p=0.00004
p=0.00004
Disorder fraction
EHEC
CR
EPEC
B
A/E pathogens proteomes
CR effectors
EHEC effectors
EPEC effectors
Human proteome
IDP
PDR
FRAG
NDP
ORD
C
CR
EHEC
EPEC
ORD
NDP
PDR
IDP
Predicted motif density
EspF
Tir
NleH
EspB
LiFA
EspF(U)
Map
Disorder fraction

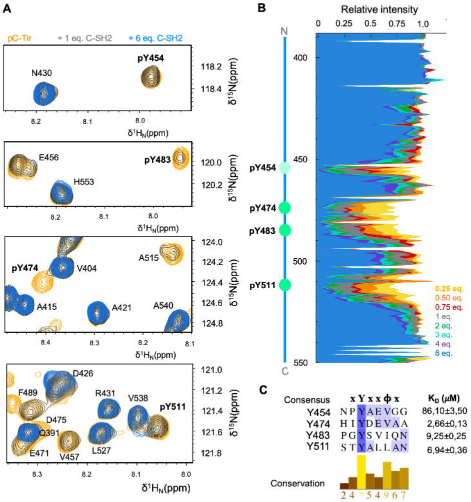

A
(s.Rg)².I(s)/I(0)
s.Rg
B
[θ] (deg. cm². dmol⁻¹. residue⁻¹)
Wavelength (nm)
C
EPEC
EHEC
δ¹⁵Nᴴ (ppm)
δ¹Hᴺ (ppm)

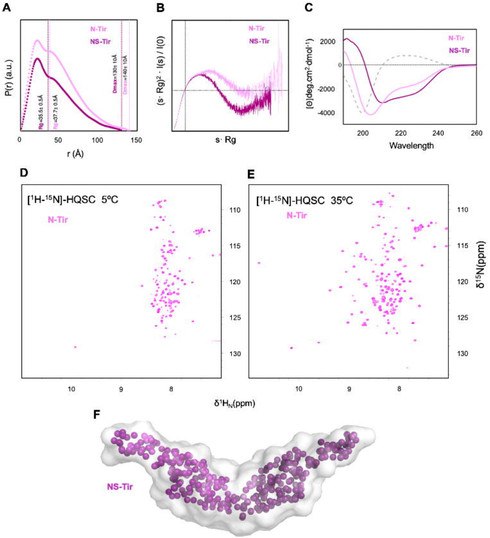


A
S434
Y454
S463
Y474
Y483
Y511
Δδ Cα-Δδ Cβ
ncSPC
I
II
a
b
Residue number
B
C-Tir
C-Tir + 2 eq. C-SH2
C-Tir + 8 eq. C-SH2
i
ii
R498
L513
V538
L520
Y511
V485
L514
A515
A512
I486
δ1HN(ppm)
δ15N(ppm)
C
1 eq.
2 eq.
4 eq.
6 eq.
8 eq.
Relative intensity
D
Kd= 68,5 μM ± 4,32 μM
A512
L513
L514
A515
Relative Intensity
[C-SH2]/[C-Tir]
E
1% TFE
10 %TFE
12.5% TFE
Y483

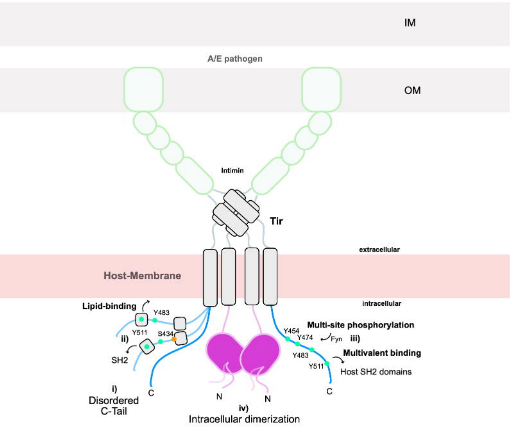


DMPC
DMPG
DHPC
A
DMPC/DHPC 6%
DMPG/DHPC 6%
δ¹⁵Nᴴ (ppm)
δ¹Hᴺ (ppm)
δ¹Hᴺ (ppm)
G525
G523
G518
G501
T528
S517
S509
M524
L513
N488
I507
F489
Y511
L514
E499
L522
V550
A515
L499
A512
Q508
Q487
I486
G526
G519
G504
G545
G493
G491
G530
G497
S492
T502
T496
S490
S529
S484
Q505
Y483
N390
R498
V495
V547
L520
Y485
R388
R389
A393
B
Y483
Y511
Relative Intensity
Residue number


A
pC-Tir
+ 1 eq. C-SH2
+ 6 eq. C-SH2
N430
pY454
E456
pY483
H553
pY474
V404
A515
A415
A421
A540
D426
F489
R431
V538
pY511
D475
Q391
E471
V457
L527
δ1HN(ppm)
δ15N(ppm)
B
Relative intensity
0
0.25
0.5
0.75
1.0
N
C
pY454
pY474
pY483
pY511
400
450
500
550
0.25 eq.
0.50 eq.
0.75 eq.
1 eq.
2 eq.
3 eq.
4 eq.
6 eq.
C
Consensus x Y x x ϕ x
KD (μM)
Y454 N P Y A E V G G 86,10±3,50
Y474 H I Y D E V A A 2,66±0,13
Y483 P G Y S V I Q N 9,25±0,25
Y511 S T Y A L L A N 6,94±0,36
Conservation
2 4 * 5 4 9 6 7



A
P(r) (a.u.)
N-Tir
NS-Tir
Rg=35.5± 0.5Å
Rg=37.7± 0.5Å
Dmax=130± 10Å
Dmax=140± 10Å
0
50
100
150
r (Å)
B
(s· Rg)² · I(s) / I(0)
N-Tir
NS-Tir
s· Rg
C
[Θ]deg.cm²dmol⁻¹)
2000
0
-2000
-4000
N-Tir
NS-Tir
200
220
240
260
Wavelength
D
[¹H-¹⁵N]-HQSC 5ºC
N-Tir
E
[¹H-¹⁵N]-HQSC 35ºC
N-Tir
110
115
120
125
130
δ¹⁵N(ppm)
10
9
8
δ¹H_N(ppm)
F
NS-Tir



IM
A/E pathogen
OM
Intimin
Tir
extracellular
Host-Membrane
intracellular
Lipid-binding
Y483
Y511
ii)
S434
SH2
i)
Disordered
C-Tail
C
N
N
iv)
Intracellular dimerization
Multi-site phosphorylation
Y454
Y474
Fyn
iii)
Y483
Multivalent binding
Y511
Host SH2 domains
C