The pathogen-encoded signaling receptor Tir exploits host-like intrinsic disorder to assist infection

Marta F. M. Vieira[a], Guillem Hernandez[a], Tiago Veloso[a], Hugo Monteiro[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 Université, Inserm, TAGC, UMR_S1090, Marseille, France

*Tiago N. Cordeiro
email: tiago.cordeiro@itqb.unl.pt

Keywords: Intrinsically disordered proteins | bacterial effectors | EPEC/EHEC | molecular mimicry | Host-pathogen interactions

Abstract

The translocated intimin receptor (Tir) is a central effector of Attaching and Effacing (A/E) pathogens responsible for worldwide foodborne disease cases. Tir acts as a cell-surface receptor in host cells, rewiring intracellular processes to assist infection by targeting multiple host proteins. We sought to understand the basis for Tir binding diversity in signaling. Here, we establish that Tir is a disordered protein with host-like binding motifs. A trait we find prevalent in several other effectors secreted by A/E bacteria. We disclose that Tir has a disordered C-terminal intracellular tail (C-Tir) with non-random structural preferences at phosphorylation sites, including host-like tyrosine-based motifs, with versatile lipid- and SH2 domain binding capability pre-phosphorylation. We show that multi-site tyrosine phosphorylation enables C-Tir to engage SH2 domains in a multivalent manner, consistent with Tir's scaffold/hub role for host proteins. Last, we uncover Tir's ability to dimerizes via its partially disordered N-terminal intracellular domain. Collectively, our findings provide an updated picture of Tir's intracellular side, highlighting its ability to mimic host disordered membrane receptors' versatility as a molecular strategy for host evasion.

Summary

Tir is a cell-surface 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 signaling pathways (4), including cell cycle (4), circadian circuits (5), post-transcriptional regulation (6), and protein degradation (7). Given their ubiquitous relevance in cellular signaling, the onset of several human diseases, including cancer and neurodegeneration, are linked to dysfunctional disordered proteins (8). A hallmark of IDPs/IDRs is the high occurrence of short linear motifs, also known as eukaryotic linear motifs (ELMs) (9). ELMs are stretches of 3-10 contiguous residues mediating transient protein-protein interactions (e.g., post-translational modifications sites, targeting signals) (10). For instance, phosphorylation sites are often found in protein IDRs, modifying charge and hydrophobicity and modulating interactions with partners(6). The prevalence of such functional elements in IDRs provides versatility to cell interaction hubs by enabling flexibility and adaptability to multiple interaction interfaces (11, 12). Moreover, IDR sequence composition is biased towards low complexity and charged residues favoring electrostatic interactions at lipid bilayer surfaces (13), with lipid-binding proteins accounting for 15% of all disordered proteins (14). As a result, intrinsic disorder is common in membrane signaling receptors, particularly in their intracellular domains (15, 16).

Disordered proteins make up 30-50% of eukaryotes' proteins (17, 18). 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 (19). 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 (20). Other less-ordered effectors are also secreted by Enteropathogenic and Enterohemorrhagic *Escherichia coli* strains (EPEC and EHEC, respectively), such as EspF(U)/TccP (21) and EspB (22).

EPEC is a leading cause of child deaths worldwide (23), and EHEC causes hemolytic uremic syndrome (HUS), defined by acute kidney failure, low platelet count, and destruction of red blood cells (24). These bacteria attach to the gut mucosa, causing a histopathological condition known as attaching and effacing (A/E) lesions, characterized by actin-rich pedestal formation and microvilli degeneration. To establish a replicative niche on host cells, EPEC and EHEC deliver a vast repertoire of effectors (25). Here, we sought to identify additional disordered effectors. So, 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) with a relatively high density of host-like ELMs, particularly in its intracellular domains. Tir is the first effector secreted during infection (26). Once inside the host, it migrates to the plasma membrane to act as a receptor for intimin presented at the bacterial surface (27), anchoring A/E pathogens to targeted host cells. Moreover, Tir promotes actin polymerization (28), suppresses autophagy (29) and immune response (30) to allow bacterial survival on the host cell's surface. It also triggers pyroptotic cell death (31) (32). As a double-pass membrane receptor, Tir adopts a hairpin topology with the external intimin-binding domain (IBD) (33) connecting two transmembrane domains, with both N- and C-terminus in the host cytoplasm (N-Tir and C-Tir, respectively) (**Fig. S1**). Those intracellular regions enable Tir to interact with at least 25 host intracellular targets (34), such as host tyrosine phosphatases SHP-1/2 to suppress pro-inflammatory cytokines signaling (30, 35) and TAK1-mediated immune response (36).

How Tir hijacks intracellular signaling by interacting with multiple host proteins to assist infection is poorly understood. To gain insights into Tir's detailed biophysics and biological function, we devised a structural biophysical study to assess its intrinsic disorder and interactions. We focus our attention

on stoichiometry, accessibility to multi-site phosphorylation, host protein- and membrane-binding of Tir intracellular regions. As a result, we confirm that Tir's intracellular side is structurally disordered. Like disordered cytosolic domains of eukaryotic cell-surface receptors, C-Tir is extensively phosphorylated. The present work shows that C-Tir adopts flexible but non-random structures around host-like phosphorylation sites implicated in Tir function and signaling. We found that some of those sites can bind to lipid bilayers, suggesting a regulation based on membrane association. Notably, we show that unphosphorylated C-Tir can recognize the C-terminal SH2 domain (C-SH2) of host SHP-1 via one host-like tyrosine-phosphorylation site. When phosphorylated, C-Tir then interacts with C-SH2 in dynamic equilibrium through any phosphorylation site. Here we highlight that the structural disorder and multi-site phosphorylation of the Tir as a molecular tool for obtaining multispecificity. Moreover, we found that N-Tir is a flexible dimer with a hybrid architecture of ordered and disordered parts, opening the question of whether Tir's signaling activation involves structural changes of preformed dimers.

Our findings establish an updated picture of Tir's intracellular side, highlighting a molecular mechanism through which Tir can interact with host components through mimicry of host disordered proteins' plasticity and versatility. This strategy may be abundant in pathogenic bacteria.

Results

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

EPEC and EHEC belong to the family of attaching and effacing (A/E) pathogens. This family also includes the mouse-specific pathogen *Citrobacter rodentium* (CR). To quantify the prevalence of disorder propensity among A/E pathogens, we estimated the disorder content of their effectors using DISOPRED3 (37) (**Fig. 1**) and IUPred 1.0 (38) (**Fig. S2**). As a result, we observed that A/E effector proteins have a higher disorder propensity than their proteome counterparts in EHEC O157:H7, the closely related EPEC H127:H6, and CR (**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 (**Fig. S3*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). 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 involved in cytoskeleton rearrangement and previously identified as an inherently less-ordered effector (22). 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 (39). 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 (40) and EspFu, only secreted by EHEC, five-and-a-half repeats (41) (**Fig. S3*B***).

Our analysis highlights Tir as a disordered effector (**Fig. 1*C***). Tir displays a high disorder propensity within its IBD and intracellular regions (**Fig. S3*B***). Pathogens can exploit host-like motifs as a strategy to interact with host components to subvert host cellular functions (42, 43) . 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 its intracellular regions, including experimentally verified SH3- and SH2-domain binding motifs (44). Structural disorder and high ELM density support Tir's ability

143 to interact with several host proteins (34). There is a correlation between disorder content and ELM
144 density, being Tir, EspF(U), and EspF clustered with the highest disorder fraction and motif overall
145 density (**Fig. 1*C***), similarly to eukaryotic IDPs (10, 45). This correlation highlights ELMs within
146 disordered segments of bacterial effectors as a strategy to disrupt host networks.

147 **Tir bears an intrinsically disordered C-terminal tail**

148 Tir secreted by A/E pathogens share substantial sequence-conservation (**Fig. S4**) and high
149 frequency of disorder-promoting residues (**Fig. S3*B***) at the N- and C-terminal regions (N-Tir and
150 C-Tir, respectively), both localized to the host intracellular side. We assessed the structural disorder
151 propensity of EPEC N-Tir and C-Tir by multiple biophysical methods, including small-angle X-ray
152 scattering (SAXS), circular dichroism (CD), and nuclear magnetic resonance (NMR). For this
153 integrative study, we used the protein constructs encompassing the residues 1-233 and 388-550
154 from the N-terminal and C-terminal cytosolic regions of EPEC Tir, respectively (**Fig. S1**). Analytical
155 size-exclusion chromatography (SEC) in combination with SAXS indicated a monomeric state for
156 C-Tir in solution with a radius of gyration (*Rg*) of 38.8 ± 0.2 Å and a maximum distance (*Dmax*) of
157 128.0 ± 5.0 Å (**Table S4**). The *Rg*-values of IDPs approximately followed a power law of $Rg = 2.54N^{0.522}$
158 (46, 47) as a function of sequence length (*N*). Using this relationship, the predicted *Rg*-
159 value is 37.4 Å for the 173-residue C-Tir construct ending with a StrepTag (8 extra residues). The
160 similarity between the experimental and predicted values suggests that C-Tir might adopt IDP-like
161 structures in solution. The SAXS-derived Kratky plot of C-Tir also shows characteristics of
162 disordered or unfolded proteins, monotonically increasing without a well-defined maximum (**Fig.
163 2*A***). This feature is absent in globular proteins and implies conformational heterogeneity and
164 flexibility for C-Tir. Likewise, the asymmetric pair distance distribution function, *P(r)*, obtained from
165 the scattering data, is compatible with a highly flexible protein sampling pairwise distances far
166 exceeding those expected for a globular protein of the same molecular weight (47) (**Fig. S5**). The
167 intrinsic disorder of C-Tir is also reflected in its CD and [^{1}H-^{15}N]-HSQC NMR spectra. C-Tir's CD
168 profile has negative ellipticity at 200 nm and a shallow band in the 210–230 nm range (**Fig. 2*B***),
169 indicating a high content of random coil with minimal ordered structural elements (48). The [^{1}H-
170 ^{15}N]-HSQC NMR spectrum for this construct displays a typical IDP fingerprint with low chemical
171 shift dispersion (49) (**Fig. 2*C***). We observed a similar spectrum for the equivalent region of Tir from
172 EHEC (**Fig. 2*C***). Although all A/E pathogens produce Tir, actin polymerization by EHEC Tir C-
173 terminus requires Esp(U)/TccP (50), which is unneeded in EPEC to initiate actin pedestal formation
174 in host cells (28). Yet, C-Tir is disordered in both strains, highlighting its structural disorder's
175 conservation and functional importance for A/E virulence.

176 Overall, the different biophysical measurements are consistent with a highly dynamic protein,
177 unambiguously showing that the C-Tir is an IDR that lacks a well-folded structure. Such inherent
178 flexibility will enable Tir for multiple binding while acting as a pathogenic scaffold/hub of intracellular
179 host proteins.

180 **C-Tir displays non-random structural preferences at phosphorylation sites**

181 We employed NMR to dissect with further detail subtle structural features of C-Tir. The NMR
182 chemical shift is the most readily accessible observable in NMR and sensitive to, even partial,
183 secondary structural elements present in IDPs (49). To gain access to this information, we assigned
184 the backbone resonances of C-Tir nuclei using standard triple-resonance spectra at high-field,
185 together with reverse labeling (**Fig. S6**). Reverse labeling of selected amino acids (51) proved to
186 be a cost-efficient approach to assign this disordered protein by reducing the spectral ambiguity

187 and should be readily applied to other IDPs. Provided with the assignment, we calculated the NMR
188 chemical shift deviations (Δδ) to random coil values of intrinsically disordered proteins (52).
189 Although C-Tir has properties of a random coil, these so-called NMR secondary shifts allowed us
190 to identify two regions with non-random structural preferences. In IDPs, transient secondary
191 structural elements are often crucial in acting as molecular recognition units, playing significant
192 roles in binding (53, 54). In figure **Fig. 3*A***, the consecutive positive ΔδCα–ΔδCβ values indicate a
193 tendency for α-helical structure in two distal segments of C-Tir: (I) D420-Q435; II) I507-A515. Both
194 stretches include residues that are phosphorylated by host kinases, S434 (55) and Y511 (30),
195 respectively. By considering all the assigned backbone resonances into a neighbor-corrected
196 sequence structural propensity calculator (nsSPC) score (56), these apparent local α-helical
197 regions became even more evident. New stretches also emerged with a putative α-helical
198 propensity, including a sequence contiguous to Y511 and downstream tyrosine 483 and a few C-
199 terminal residues (**Fig. 3*A***). Y483 and Y511 are core residues of Tir's host-like Immunoreceptor
200 tyrosine-based inhibitory motifs (ITIMs) that recruit tyrosine phosphatases SHP-1/2 in an ITIM
201 phosphorylation-dependent manner to inhibit host innate immune responses (30, 35).

202 We further probed C-Tir's non-random coil preferences using paramagnetic relaxation
203 enhancement (PRE) NMR experiments. To this end, five residues scattered along C-Tir's sequence
204 were single-point mutated to cysteine to attach a thiol-reactive nitroxide radical probe (i.e., MTSL).
205 The relaxation induced by this paramagnetic tag resulted in substantial peak broadening (**Fig. S7**).
206 In ideal random coils, this affects the residues in the vicinity of the MTSL tag, and other residues
207 experience less or no broadening. However, if long-range contacts are present within the
208 disordered chain, residues far away in the sequence can be spatially close to the tag and
209 experience the PRE effect. On this basis, we calculated the deviations between the experimental
210 PRE and those predicted from a random-coil model (i.e., ΔPRE) to detect non-random coil structural
211 preferences. We measured ΔPRE-values for each single-cysteine mutant. Figure **3*B*** shows, in a
212 grayscale, the ΔPRE-induced by MTSL at positions 409, 428,463, 509, and 536, respectively. All
213 consistent with compact conformations in the disordered C-Tir. The paramagnetic probe in
214 positions 405 and 428 showed noticeable long-range interactions (ΔPRE>0.2) with distant
215 downstream residues. The MTSL at S463C caused significant PRE (ΔPRE>0.6) in the residues
216 497–526, and when placed in position 509, it induced broadening in the region 438-479, including
217 position 463. The complementary long-range interactions observed between residues around
218 positions 463 and 509 are mutually self-consistent with compact conformations. Both areas include
219 tyrosine-based ELMs that are involved in assembling host components and intracellular signaling.
220 To better describe and visualize the non-random preferences of C-Tir, we integrated the PRE data
221 with SAXS data to determine the conformational ensemble of C-Tir (**Fig. 3*C***). The established sub-
222 ensembles agree with the experimental data (**Fig. 3*D*, Fig. S7**), revealing a bias for expanded and
223 compact structures. The ensemble-averaged distances between the Cα atoms for the
224 experimentally selected conformations ($\langle r_{ij} \rangle_{ens}$), normalized by the respective random coil (RC)
225 ensemble ($\langle r_{ij} \rangle_{RC}$), show an N-terminal region more extended relative to the RC. At the same time,
226 the C-terminal half is more compact (**Fig. 3E**). Compaction is clustered around the host-like ITIMs
227 implicated in interaction with host proteins.

228 **C-Tir can interact with the host SHP-1 pre-phosphorylation**

229 The intracellular signaling associated with Tir ITIMs relies on the phosphorylation of the central
230 tyrosine by Src family kinases (SFKs), creating a binding site for SH2–containing proteins, which
231 become activated upon recruitment. Yet, we found that unphosphorylated C-Tir can recognize host
232 SH2-domains with low-intermediate affinity. Using the C-terminal SH2 domain (C-SH2) of

233 phosphatase SHP-1, we identified, by NMR, that residues surrounding the unphosphorylated Y511
234 ITIM-like motif are involved in binding SH2 domains (**Fig. 4*A-B***). Binding to C-Tir caused a selective
235 loss of intensities for the HSQC peaks, mainly from residues $A_{512}LLA_{515}$. Visible NMR signals
236 retained low-dispersion, indicating that the protein remains mostly disordered and flexible in the
237 complex. We estimated an apparent K_D of *ca.* 68.5μM for this interaction (**Fig. 4*C***). Thus, besides
238 encoding for α-helical structures, the sequence surrounding Y511 (i+4) has a chemical signature
239 for SH2-domain binding pre-phosphorylation. This observation highlights that those residues
240 adjacent to Y511 can define a binding motif for C-SH2, likely serving as a precursor to a tighter
241 binding upon phosphorylation. The presence of a transient pre-formed helix with long-range
242 interactions at this region suggests a possible role for this structure as molecular recognition
243 elements for C-Tir.

244 Remarkably, Tir ITIMs share with cytoplasmic tyrosine-based sequence motifs of host receptors
245 the ability to adopt α-helical conformations. Indeed, host ITIMs and Immunoreceptor tyrosine-based
246 activation (ITAMs) can form dynamic/transient α-helices, whether in free (57, 58) or membrane-
247 bound state (59). Their dynamic binding to membranes is postulated to regulate tyrosine sites'
248 access to phosphorylation, providing a switch between functional and non-functional conformations
249 (60). The helical conformation in ITIMs/ITAMs is stabilized *in vitro* in the presence of 2,2,2-
250 trifluoroethanol (TFE) or through binding to detergent micelles and lipids (58, 59, 61). In the
251 presence of membrane-mimetic solvent TFE, we found that C-Tir exhibited a substantial increase
252 in secondary structure, with the CD signature of a partial helical protein (**Fig. S8*A***). This
253 conformational change is followed by chemical shift differences (**Fig.*S8B-C)*** and a differential
254 decrease in peak intensities of the C-Tir amide proton resonances in the presence of increasing
255 amounts of TFE, mainly around Y511 and Y483 and not at S434 (**Fig. S8*C***, bottom). This effect
256 supports an enhanced α-helical conformation propensity around the two tyrosines, and a putative
257 role in membrane-binding, as postulated for some host tyrosine-based motifs that undergo TFE-
258 induced structural changes (58). Our NMR results show that unphosphorylated C-Tir host-like ITIM
259 motifs can adopt pre-formed structures relevant to membrane binding and molecular recognition of
260 SH2 domains of signaling proteins.

261 **Disordered C-Tir binds lipids**

262 The ability to form α-helices stabilized by TFE prompts us to further investigate lipid binding by C-
263 Tir. To do so, we incubated ^{15}N-labelled C-Tir with different bicelles, which are small planar bilayers
264 of long-chain lipids closed by curved micelle-like walls of short-chain lipids. These disk-like
265 structures are membrane models extensively used to explore protein-membrane interactions,
266 including those involving disordered proteins (62). As long-chain lipids, we used 1,2-dimyristoyl-*sn*-
267 glycero-3-phosphocholine (DMPC) or the acidic lipid 1,2-dimyristoyl-*sn*-glycero-3-
268 phosphorylglycerol (DMPG) mixed with the short-chain lipid 1,2-dihexanoyl-*sn*-glycero-3-
269 phosphocholine (DHPC), offering the possibility to create bicelles with different lipid charge density.

270 In the presence of bicelles containing DMPC lipids, we observed a significant residue-specific
271 decrease in the ratio of NMR signals around Y511; yet more subtle, also around Y483, as previously
272 induced with TFE (**Fig. 5**). This finding shows that C-Tir does interact with non-charged lipid bilayers
273 predominantly via its Y511-motif, including mostly hydrophobic residues, such as $A_{512}LLA_{515}$. We
274 further tested the influence of increasing negatively charged lipid head groups on C-Tir membrane
275 association with bicelles containing DMPG lipids. Our NMR data show a more extended NMR signal
276 attenuation in the presence of DMPG/DHPC bicelles, even including the first stretch of residues of
277 this construct, which display some positively charged residues (e.g., R388 and K389) (**Fig. 5**). In

278 the context of the full-length protein, these residues define the membrane-proximal region (**Fig. S1,**
279 **S3**). Given its proximity to the cell plasma membrane, it is reasonable to assume its involvement in
280 membrane binding. Aside from Y511 surrounding residues, the sequence around Y483 is also
281 notably affected by anionic lipid content, indicating that lipid-charge content modulates C-Tir
282 membrane binding modes. Contrary, the residues 395-475 did not show any membrane interaction.
283 This region has a net negative charge (pI~4.5), whereas the sequences displaying membrane
284 ability are positive (pI~9.9), thus in line with an electrostatic model. In this residue range, C-Tir
285 bears two experimental confirmed host protein binding sites: a) NCK Src Homology 2 (SH2) domain
286 binding motif (28); and b) NPY motif that interacts with the I-BAR domain of host IRSp53/IRTKS
287 (63). With both bicelles, the last C-terminal residues of C-Tir also show NMR-signal attenuation
288 indicative of membrane-binding (**Fig. 5*B***). These residues (A_{541}PTPGPVRFV_{550}) emerge as part of
289 a potential new lipid-binding region, also affected by TFE (**Fig. S8C**) and showing ncSPC values
290 indicating α-helical propensity (**Fig. 3*A***) and pre-existing compaction (**Fig. 3*B***). Our data thus
291 suggest that C-Tir can undergo multivalent electrostatic interactions with lipid bilayers that might
292 potentially fine-tune Tir's activity in the host cell.

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

294 In the host, Tir becomes phosphorylated by Src family protein tyrosine kinases (PTKs) at several
295 tyrosine phosphorylation sites targeting SH2 domain-containing host proteins. To fully assess its
296 ability to recruit SH2 domains, we reconstructed C-Tir's Tyr-phosphorylated state and evaluated its
297 binding to C-SH2. To this end, we incubated C-Tir with the Src family PTK Fyn, and quantitatively
298 monitored its phosphorylation by NMR (64). The sensitivity of chemical shifts to changes caused
299 by phosphorylation allowed us to identify four phosphorylated tyrosine sites along the disordered
300 C-Tir modified by Fyn (i.e., Y454, Y474, Y483, Y511). Phosphorylation caused small but noticeable
301 changes in the chemical shifts of the tyrosines and adjacent residues. The [^{1}H-^{15}N]-HSQC spectrum
302 displayed low amide proton dispersion, a diagnostic that the C-Tir remains disordered upon
303 multisite phosphorylation. Moreover, the secondary chemical shifts of this 4-fold phosphorylated
304 state (pC-Tir) did not reveal substantial changes in local structure propensity due to
305 phosphorylation (**Fig. S9**), as similarly reported for other disordered proteins (65, 66).

306 To evaluate the interaction of phosphorylated C-Tir (pC-Tir) with an SH2 domain, we titrated
307 unlabeled C-SH2 into a ^{15}N-labeled pC-Tir solution and monitored the binding interaction
308 broadening of signals in NMR [^{1}H-^{15}N]-HSQC spectra (**Fig. 6*A*,*B***). With the resonance re-
309 assignment of pC-Tir, we identified that upon phosphorylation, all tyrosine sites interact with C-
310 SH2, and not exclusively the Y511-based motif as observed for unphosphorylated C-Tir. The
311 interaction caused a drop in peak intensities around each site (i.e., i= pY454, pY474, pY483,
312 pY511), including roughly residues i+4 and i-1, even at sub-stoichiometric conditions (**Fig. 6*B***).
313 This suggests a "fuzzy" multivalent binding of C-SH2 to four xYx(x)ϕ motifs, where ϕ is hydrophobic
314 (V/L/I) (**Fig. 6*C***), with multiple pYs on C-Tir interacting with C-SH2 in dynamic equilibrium (65, 66).
315 Except for these residues, backbone chemical shift changes were relatively small, with spectra
316 displaying sharp line widths and low dispersion, both characteristics of a disordered protein. Thus,
317 pC-Tir binds C-SH2 retaining its high level of disorder. The relative signal intensities at each
318 phospho-tyrosine (pY) and surrounding residues approximately reflect the fraction of unbound
319 sites, allowing us to estimate their apparent local binding affinity (**Fig. S10**). Phosphorylation
320 enhanced the binding to Y511 ITIM by ~10 fold and enabled other tyrosine-based motifs to engage
321 C-SH2. Among the four pYs, the resonances around pY454 were less broadened during the
322 titration, suggesting lower local binding. The remaining pYs bind C-SH2 with similar strength (~3-9

323 μM). Overall, phosphorylation at multiple tyrosine sites provides various SH2 docking sites and
324 reinforces the role of Tir as a scaffolding hub.

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

326 In a similar fashion, we evaluated whether the N-terminal cytosolic region of Tir (N-Tir; **Fig S1**) was
327 also intrinsically disordered. Yet, contrary to C-Tir, this region adopts a flexible but more defined
328 structure. Based on SEC-SAXS data, the N-Tir is a non-globular protein with a *Rg* of 37.70 ± 0.10
329 Å and a *Dmax* of 140.0 ± 10.0 Å (**Fig. 7*A***) (**Table S4**), and a molecular weight compatible with a
330 52kDa dimer, in line with the observed gel filtration retention time. Like C-Tir, the respective *P(r)*
331 curve is highly asymmetric but bimodal, with a peak at 21.9 Å and a prominent shoulder at 35.9 Å,
332 suggesting that the dimer is elongated with spatially separated lobes. Its SAXS-derived Kratky plot
333 is also not bell-shaped as expected for a non-globular and flexible arrangement (**Fig. 7*B***). The
334 corresponding SAXS-driven *ab initio* reconstruction highlights an elongated S-shaped core for the
335 dimeric N-Tir 1-233 polypeptide fused to a C-terminal Strep-Tag (**Fig. S5**). This extended dimer is
336 largely disordered but contains stable secondary structural elements. Its far-UV CD profile reveals
337 a partially folded protein, marked by a negative maximum at 205 nm due to high disorder content.
338 The positive signal at 190 nm and the negative shoulder at 220 nm reflect ordered elements (**Fig.**
339 **7*C***). So, SAXS and CD data indicated that N-Tir is a partially folded dimer. To probe the order-
340 disorder interplay within N-Tir at high-resolution, we have collected NMR data at different
341 temperatures.

342 For globular proteins, the 52kDa dimer size is beyond the practical range amenable to traditional
343 NMR spectroscopy in solution. The NMR signals of larger molecules relax faster, leading to line
344 broadening, low spectral sensitivity, and eventually loss of NMR signals. This problem is even more
345 significant at low temperatures. Nevertheless, we found that the 52kDa dimer has a [^{1}H-^{15}N]-HSQC
346 NMR spectrum at 5 ºC characterized by intense signals and low chemical shift dispersion akin to
347 an IDP (**Fig. 7*D***). We assigned the resonances at 5 ºC mostly to the first 80 and last 35 residues
348 of the N-Tir, meaning that those regions are flexible in solution. The missing resonances of the
349 central residues became visible with increasing temperature due to faster tumbling. At 35 ºC, they
350 adopted an NMR fingerprint of a folded-like protein (**Fig. 7*E***), reflecting a more well-structured
351 region. With NMR, we could establish that N-Tir's central sequence has the broad NMR fingerprint
352 of a well-folded protein, flanked by residues with poorly dispersed $^{1}H^{N}$ backbone resonances
353 pointing disorder. A closer look at the structural disorder predictions computed for N-Tir also
354 suggests an order-disorder duality, with flanking residues more flexible than those in the central
355 region (**Fig. S11*A***). Such disordered segments harbor linear motifs that were experimentally
356 identified to interact with host proteins **(Fig. S3*B***) (44, 67).

357 The partial deletion of these disordered regions reduced the random-coil content reported by CD
358 data, highlighting a more structured central part of N-Tir (N-Tir$_{60-200}$, hereafter NS-Tir). The
359 difference between CD spectra resulted in a pure random-coil CD signature, reinforcing that the
360 flanking regions are disordered, whereas the central part folded. NS-Tir's CD profile displays
361 positive values below 200 nm and two negative bands at 210 and 222 nm, commonly associated
362 with structured secondary conformations (**Fig. 7*C***).

363 We have employed SAXS to probe further the overall structure of NS-Tir. Our synchrotron
364 SEC−SAXS data confirmed that NS-Tir is still a stable dimer in solution. A slightly less broadened
365 Kratky profile reflects the absence of the disordered flanking regions (**Fig. 7*B***) but compatible with
366 an elongated dimer (**Fig. 7*F***), retaining a similar overall conformation supported by NMR. We

367 acquired [^{1}H-^{15}N]-HSQC data for NS-Tir and superimposed them to N-Tir spectra (**Fig. S11*B***). Over
368 75% of the visible resonances overlapped between the two constructs, with a low dispersion for the
369 ^{1}H dimension at 5 ºC and a broader [^{1}H-^{15}N]-fingerprint at 35 ºC, indicating that NS-Tir maintained
370 a similar folded isolated and in the full-length N-terminal region context. The reversible thermal
371 denaturation of NS-Tir monitored by far-UV CD shows the dimer unfolding in two discrete steps,
372 capturing both folding and dimerization events. It reveals a stable dimer that unfolds and dissociates
373 by increasing temperature (68) (**Fig. S12**).

374 Overall, our results support a model of Tir spanning the host plasma membrane with flexible N-
375 terminal intracellular domains assembled in a dimer form (**Fig. 8**). In this model, N-Tir is anchored
376 to the membrane by flexible segments and with dangling linear motifs located at its disordered N-
377 termini.

378 **Discussion**

379 In eukaryotic cells, protein disorder is dominant in protein hubs controlling the complex intracellular
380 networks that influence physiological responses. We observed that A/E pathogens have a
381 structurally diverse repertoire of protein effectors enriched in disorder-prone residues. Some
382 effectors are fully disordered proteins with a high density of host-like interacting motifs, challenging
383 the general trend that prokaryotic proteins are less disordered. Among such disordered effectors
384 emerged Tir, a 56 kDa cell surface receptor able to reshape host cellular behavior during infection.

385 In the present work, we examined the structure of Tir, finding that its intracellular domains (i.e., C-
386 Tir and N-Tir) are flexible with IDR features similar to those found in the cytoplasmic domains of
387 host transmembrane proteins. NMR analysis showed that C-Tir is not exclusively disordered but
388 can adopt partial secondary structural elements and non-random long-range contacts around
389 phosphorylation sites and motifs involved in versatile interactions (*i.e.*, host protein and lipid-
390 binding).

391 The C-Tir displays cytoplasmic tyrosine-based motifs present in the disordered cytosolic tails of
392 host receptors. The standard mode of action associated with these motifs is the phosphorylation of
393 the central tyrosine by Src family PTKs, creating a binding site for SH2–containing proteins. In C-
394 Tir, we found that one of those motifs, involving Y511, displays a preformed transient helical
395 structure and binds the C-SH2 pre-phosphorylation. This motif also interacts with lipid bilayers.
396 Notably, subsequent multi-site phosphorylation enables the transition from binary to multivalent
397 fuzzy complex with C-SH2. These observations highlight a functional diversity for C-Tir and possible
398 cooperation between host proteins, phosphorylation, and the membrane in signaling processes
399 inside the host (13). Some T-cell receptor cytoplasmic disordered domains associate with the
400 plasma membrane via tyrosine-based motifs with a helical propensity, and such association
401 regulates phosphorylation and downstream signaling (60). Like the CD3ε cytoplasmic domain of
402 the T-cell receptor (59), C-Tir also has two tyrosines involved in lipid-binding and helical propensity.
403 Our results suggest that the C-terminal half Tir can mimic this regulatory stratagem to fine-tune its
404 ability to interact with human cell components, thereby interfering with normal cellular functions.
405 Further studies are needed to corroborate this initial observation. Nevertheless, we show that the
406 C-Tir can interact with membranes via ITMs and its last C-terminal residues. Importantly, our results
407 establish that this membrane affinity is residue-specific and modulated by lipid composition in a
408 quantitative and site-resolved way, thus suggesting the existence of a regulation layer based on
409 lipid composition. Transiently structured lipid-binding regions might bury tyrosine residues
410 rendering them inaccessible, while their phosphorylation perturbs membrane anchoring (69).

411 The flexibility and conformational plasticity of C-Tir make it readily accessible for phosphorylation.
412 We found in C-Tir four *bona fide* tyrosine phosphorylation sites modified by Fyn that render the
413 ability for multivalent/promiscuous binding, supporting the idea of Tir acting as a signaling hub.
414 Even though we only used protein fragments and explored one phosphorylation state, our results
415 support the presence of four phospho-tyrosines (pY) in the C-Tir sequence coexisting
416 simultaneously. All can bind the C-SH2 in the micromolar range, as reported for other pY-containing
417 motifs (70). Interestingly, the pY454 site, which displays the lowest affinity, is part of a conserved
418 NPY binding motif (**Fig. S4**) for the I-BAR domain of host IRSp53/IRTKS, linking Tir to the actin
419 polymerization machinery (63) in a phosphorylation-independent manner. So, Y454 might act as a
420 binding site for distinct and competing host proteins depending on its phosphorylation state. The
421 superposition of interplaying functional elements observed in C-Tir mimics the mechanics of human
422 intrinsically disordered domains. Future work using high-resolution structural studies targeting the
423 interaction with multiple partners and different C-Tir's phosphorylation states will provide additional
424 insights into Tir binding specificity and action mode.

425 Last, we unraveled for the first time that the N-terminal region, yet partially disordered, forms a
426 stable dimer with potential implications on Tir self-assembly and signaling. Clustering of Tir triggers
427 host signaling events when binding to intimin. The extracellular intimin binding domain (IBD) of Tir
428 binds intimin as a dimer (**Fig. 8**) (33) in a reticulating model (71). Thus, Tir's intracellular self-
429 assembly opens the question of whether host signaling activation involves intimin-induced
430 dimerization or structural changes of preformed dimers.

431 We establish an updated picture of Tir's intracellular side (**Fig. 8**), laying the path towards illustrating
432 how Tir hijacks host signaling. Notably, yet less abundant in prokaryotes, the structural disorder of
433 Tir and several other A/E effectors reinforce the idea of a positive evolutionary selection towards
434 disordered proteins in the pool of secreted effectors by bacterial pathogens to target host cellular
435 machinery (19). Given the ubiquitous presence of IDPs as transcriptional factors (72) and, more
436 generally, as hubs in host networks (73), bacterial effectors with host-like disordered protein
437 features are an efficient way to subvert host eukaryotic systems and promote infection.

438 **Methods and Materials**

439 **Disorder prediction and short linear motif analysis**

440 We collected the sequence of A/E pathogen effectors and corresponding reference proteomes from
441 the UniprotKB database (74) (**Table S1**). Three effector sets were assembled for the following
442 strains: EHEC O157:H7, EPEC H127:H6, and CR. The representative set of 20365 human proteins
443 was extracted from UniprotKB human reference proteome. To compute disorder propensity at the
444 residue level, we used DISOPRED 3 (37) and IUPred 1.0 ("long" mode prediction) (38), structural
445 disorder predictors. Disordered residues were defined as those with a propensity score equal to or
446 above 0.5. We used this metric to calculate the fraction of disordered residues for each protein.
447 The one-sided Mann-WhitneyU-test (75) was used to compare disorder fraction distribution of the
448 effector collections and their corresponding proteomes (**SI Materials and Methods**). Both per-
449 residue scores and aggregated disorder fractions were used to classify each protein according to
450 the structural categories adapted from (76) (IDP: Intrinsically disordered proteins; PDR: Proteins
451 with intrinsically disordered regions; FRAG: Proteins with fragmented-disorder; NDR: Not
452 disordered proteins; ORD: Ordered Proteins). See **SI Materials and Methods** and **Table S2** for a
453 detailed description of the criteria used to define the structural categories. The list of eukaryotic
454 short linear motifs (ELMs) was downloaded from the ELM database (77) (version 1.4, May 2017).

455 We searched for occurrences of putative ELMs in effector sequences using the ANCHOR tool (78).
456 For each effector sequence, we calculated the motif density as the fraction between the amino
457 acids belonging to putative ELMs predicted as disordered and the total number of amino acids
458 predicted as disordered.

459 **Protein expression and purification**

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

464

465 **NMR spectroscopy**

466 All NMR experiments were recorded on an 800 MHz Bruker Avance II+ spectrometer equipped with
467 a TCI cryoprobe. The backbone assignment was performed combining standard triple-resonance
468 experiments and reverse labeling (51). NMR is well suited to study such dynamic and transient
469 interactions in solution at single-residue resolution. We extensively explored the sensitivity of NMR
470 to changes in the chemical environment to monitor phosphorylation, protein- and lipid-binding. More
471 details in **SI Material and Methods.**

472 **PRE sample preparation**

473 A 10-fold excess of MTSL was immediately added to DTT-free samples, just after elution from PD-
474 10 columns, and left reacting overnight at 4°C in the dark. Excess MTSL was removed using PD-
475 10 columns and samples exchanged into the NMR buffer. We obtained the reference diamagnetic
476 spectra by adding 5-fold excess of ascorbic acid (79). PRE analysis is described in detail in **SI**
477 **Material and Methods.**

478 **Small-angle X-ray scattering (SAXS)**

479 We employed synchrotron SAXS coupled with size exclusion chromatography (SEC) to probe the
480 overall size and conformational properties of Tir intracellular regions and their *ab initio* shape (N-
481 Tir) or ensemble (C-Tir) representations. A comprehensive description of SAXS measurements and
482 analysis is described in detail in **SI Material and Methods.**

483 **Structural ensembles of C-Tir**

484 We used the EOM approach (80) to select a subset of conformations from a large starting pool of
485 conformations to achieve agreement with SAXS and PRE data, i.e., in simultaneous agreement
486 with intramolecular distances from PREs and overall size/shape distribution encoded by the SAXS
487 data. This initial pool of structures was created with Flexible Meccano (81), followed by side-chain
488 modeling and *in-silico* incorporation of multiple spin-label dispositions at single-cysteines to
489 interpret the PRE data (82, 83). See **SI Material and Methods** for additional details.

490 **Bicelles preparation**

491 Bicelles with different charge density were prepared based on the protocols described in refs (62,
492 84). See **SI Material and Methods** for additional details.
493

494 **Tyrosine phosphorylation**

495 Fully phosphorylation of four tyrosine residues in C-Tir was achieved by incubating the protein with
496 human recombinant Fyn Kinase overnight (57). Optimized experimental conditions are detailed on
497 **SI Material and Methods.**

498 **CD spectroscopy**

499 We performed Far-UV CD spectroscopy in a J-815 spectrophotometer (Jasco), using a 1-mm
500 optical pathlength cuvette for high performance (QS) (Hellma). The protein concentration was 6,
501 10, or 11.9 μM for N-Tir, NS-Tir, or C-Tir samples, respectively. N-Tir and NS-Tir were in 0.4mM
502 Tris-HCl pH 8, 15.0 mM NaF, 0.1 mM TCEP. C-Tir was prepared in 2 mM phosphate pH 6.5, 15
503 mM NaF and 0.1 mM EDTA supplemented with 0, 5, 10, 20, and 50% 2,2,2-Trifluoroethanol (TFE).
504 CD spectra were obtained by averaging 5-10 accumulations collected in the 190 nm to 260 nm
505 wavelength range using a 0.1 nm data pitch and 2 nm bandwidth at a scan speed of 50 nm/min at
506 10°C. The thermal denaturation of NS-Tir was followed by monitoring temperature-dependent
507 stepwise changes in spectral features ranging from 5 to 95 ºC in 5 ºC steps. We also recorded the
508 CD signal at 222 nm probing for α-helix secondary structure by applying a temperature ramp of 1
509 ºC min^{-1}.

510 **Data and software availability**

511 The NMR chemical shifts of C-Tir and pC-Tir are available at the BMRB with accession codes
512 50758 (C-Tir), 50759 (pC-Tir). The SEC-SAXS data and models are available at SASBDB under
513 the project "SAXS studies on the intracellular region of the translocated intimin receptor". The
514 accession codes are detailed in Table S4. Proteomes and effector collections, disorder predictions
515 are available for download at https://osf.io/3mka9/. The associated code is available at
516 https://osf.io/cxkjf/.

517 **Acknowledgments**

518 MFMV is funded by a MolBioS FCT PhD program fellowship (PD/BD/135482/2018). GH thanks the
519 PT-NMR FCT PhD Program (PD/BD/147227/2019) for financial support. TNC is the recipient of the
520 grant CEECIND/01443/2017. This work was funded by national funds through FCT, Project
521 MOSTMICRO-ITQB with refs UIDB/04612/2020 and UIDP/04612/2020 and FEDER Funds through
522 COMPETE 2020 (0145-FEDER-007660). Financial support was provided by European EC
523 Horizon2020 TIMB3 (Project 810856). We acknowledge using the ESRF-BM29 (MX-2085-BAG)
524 and DLS-B21 Bio-SAXS beamlines (MX20161-1, SM21035-177). We thank Miquel Pons, from UB,
525 and Lígia Martins, and Adriano Henriques, from ITQB-NOVA, for many helpful discussions. We
526 thank the staff of ITQB-NOVA's Research Facilities for their technical assistance. MA thanks Prof.
527 Hartmut Oschkinat and FMP-Berlin for the financial support. AZ acknowledges the financial support
528 of the JPI HDHL-INTIMIC action co-funded by the Agence Nationale de la Recherche (ANR-17-
529 HDIM-0001-03) and thanks Fabrice Lopez and Aurélie Bergon (TAGC, Marseille) for technical
530 assistance. NMR data were acquired at CERMAX, ITQB-NOVA, Oeiras, Portugal, with equipment
531 funded by FCT, project AAC 01/SAICT/2016.

532 **Author Contributions**

533 Conceptualization: TNC
534 Methodology: MA, AZ, TNC
535 Investigation: MFMV, GH, TV, MA, AZ, TNC
536 Supervision: TNC
537 Writing—review & editing: MA, AZ, TNC
538

539 **Competing Interest**

540 The authors declare no competing interest.

541 **References**

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

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

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

548 4. C. A. Galea, Y. Wang, S. G. Sivakolundu, R. W. Kriwacki, Regulation of cell division by
549 intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits.
550 *Biochemistry* **47**, 7598–7609 (2008).

551 5. J. F. Pelham, J. C. Dunlap, J. M. Hurley, Intrinsic disorder is an essential characteristic of
552 components in the conserved circadian circuit. *Cell Commun. Signal.* **18**, 181 (2020).

553 6. A. Bah, J. D. Forman-Kay, Modulation of Intrinsically Disordered Protein Function by Post-
554 translational Modifications. *J. Biol. Chem.* **291**, 6696–6705 (2016).

555 7. R. van der Lee, *et al.*, Intrinsically disordered segments affect protein half-life in the cell
556 and during evolution. *Cell Rep.* **8**, 1832–1844 (2014).

557 8. V. N. Uversky, Wrecked regulation of intrinsically disordered proteins in diseases:
558 pathogenicity of deregulated regulators. *Front. Mol. Biosci.* **1**, 6 (2014).

559 9. N. E. Davey, *et al.*, Attributes of short linear motifs. *Mol. Biosyst.* **8**, 268–281 (2012).

560 10. K. Van Roey, *et al.*, Short linear motifs: ubiquitous and functionally diverse protein
561 interaction modules directing cell regulation. *Chem. Rev.* **114**, 6733–6778 (2014).

562 11. V. Csizmok, A. V. Follis, R. W. Kriwacki, J. D. Forman-Kay, Dynamic protein interaction
563 networks and new structural paradigms in signaling. *Chem. Rev.* **116**, 6424–6462 (2016).

564 12. N. Jespersen, E. Barbar, Emerging Features of Linear Motif-Binding Hub Proteins. *Trends*
565 *Biochem. Sci.* **45**, 375–384 (2020).

566 13. J. Cornish, S. G. Chamberlain, D. Owen, H. R. Mott, Intrinsically disordered proteins and
567 membranes: a marriage of convenience for cell signalling? *Biochem. Soc. Trans.* **48**, 2669–2689
568 (2020).

569 14. E. Deryusheva, *et al.*, Does Intrinsic Disorder in Proteins Favor Their Interaction with
570 Lipids? *Proteomics* **19**, e1800098 (2019).

571 15. J. Bürgi, B. Xue, V. N. Uversky, F. G. van der Goot, Intrinsic disorder in transmembrane
572 proteins: roles in signaling and topology prediction. *PLoS ONE* **11**, e0158594 (2016).

573 16. M. Kjaergaard, B. B. Kragelund, Functions of intrinsic disorder in transmembrane proteins.
574 *Cell. Mol. Life Sci.* **74**, 3205–3224 (2017).

575 17. W. Basile, M. Salvatore, C. Bassot, A. Elofsson, Why do eukaryotic proteins contain more
576 intrinsically disordered regions? *PLoS Comput. Biol.* **15**, e1007186 (2019).

577 18. B. Xue, A. K. Dunker, V. N. Uversky, Orderly order in protein intrinsic disorder distribution:
578 disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* **30**,
579 137–149 (2012).

580 19. M. Marín, V. N. Uversky, T. Ott, Intrinsic disorder in pathogen effectors: protein flexibility
581 as an evolutionary hallmark in a molecular arms race. *Plant Cell* **25**, 3153–3157 (2013).

582 20. D. P. O'Brien, *et al.*, Calmodulin fishing with a structurally disordered bait triggers CyaA
583 catalysis. *PLoS Biol.* **15**, e2004486 (2017).

584 21. O. Aitio, *et al.*, Enterohaemorrhagic Escherichia coli exploits a tryptophan switch to hijack
585 host f-actin assembly. *Structure* **20**, 1692–1703 (2012).

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

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

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

592 25. T. Tobe, *et al.*, An extensive repertoire of type III secretion effectors in Escherichia coli
593 O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci USA* **103**, 14941–
594 14946 (2006).

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

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

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

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

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

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

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

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

613 34. S. Blasche, *et al.*, The EHEC-host interactome reveals novel targets for the translocated
614 intimin receptor. *Sci. Rep.* **4**, 7531 (2014).

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

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

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

621 38. Z. Dosztányi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of
622 intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**,
623 3433–3434 (2005).

624 39. A. M. Grishin, *et al.*, NleH defines a new family of bacterial effector kinases. *Structure* **22**,
625 250–259 (2014).

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

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

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

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

634 44. B. Bommarius, *et al.*, Enteropathogenic Escherichia coli Tir is an SH2/3 ligand that recruits
635 and activates tyrosine kinases required for pedestal formation. *Mol. Microbiol.* **63**, 1748–1768
636 (2007).

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

639 46. P. Bernadó, M. Blackledge, A self-consistent description of the conformational behavior of
640 chemically denatured proteins from NMR and small angle scattering. *Biophys. J.* **97**, 2839–2845
641 (2009).

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

644 48. L. B. Chemes, L. G. Alonso, M. G. Noval, G. de Prat-Gay, Circular dichroism techniques
645 for the analysis of intrinsically disordered proteins and domains. *Methods Mol. Biol.* **895**, 387–404
646 (2012).

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

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

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

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

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

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

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

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

57. 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).

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

59. 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).

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

61. 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).

62. 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).

63. 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).

64. 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).

65. 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).

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

67. 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).

68. 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).

69. 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).

70. 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).

71. 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).

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

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

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

75. 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).

76. 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.

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

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

715 79. M. Arbesú, *et al.*, The unique domain forms a fuzzy intramolecular complex in src family
716 kinases. *Structure* **25**, 630-640.e4 (2017).

717 80. P. Bernadó, E. Mylonas, M. V. Petoukhov, M. Blackledge, D. I. Svergun, Structural
718 characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**,
719 5656–5664 (2007).

720 81. V. Ozenne, *et al.*, Flexible-meccano: a tool for the generation of explicit ensemble
721 descriptions of intrinsically disordered proteins and their associated experimental observables.
722 *Bioinformatics* **28**, 1463–1470 (2012).

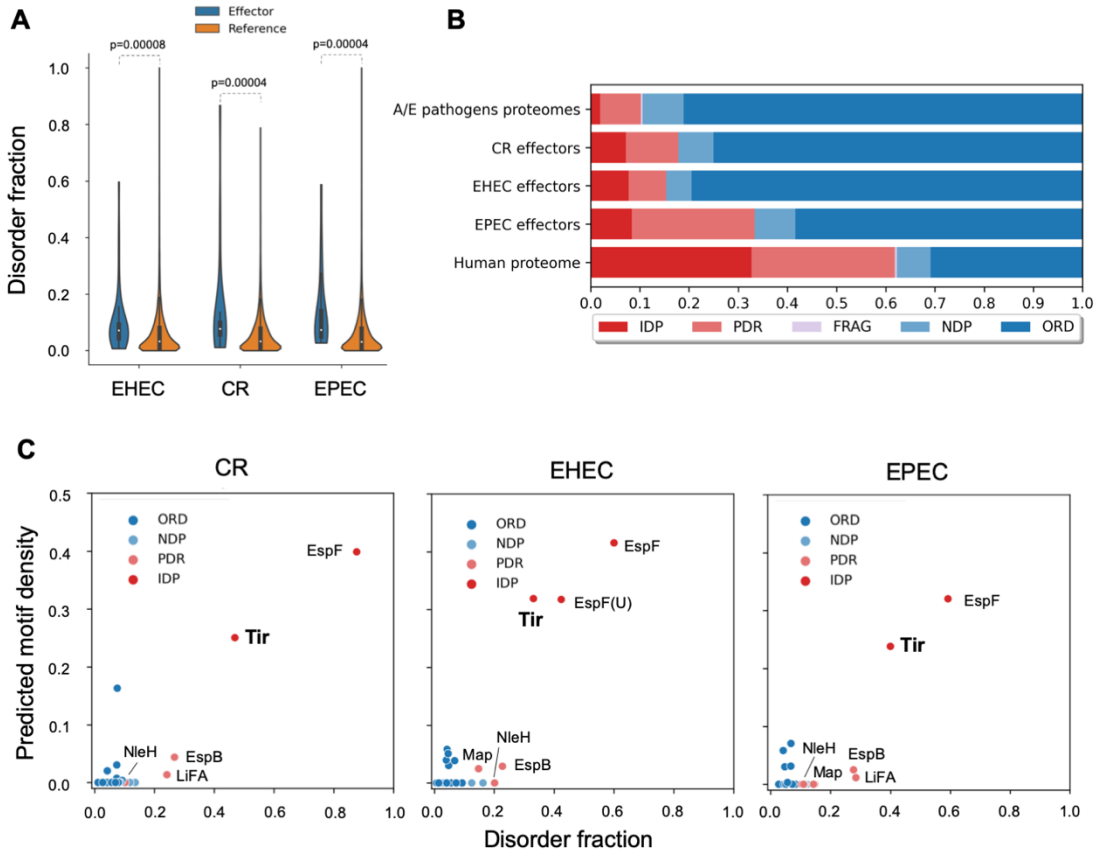
723 82. L. Salmon, *et al.*, NMR characterization of long-range order in intrinsically disordered
724 proteins. *J. Am. Chem. Soc.* **132**, 8407–8418 (2010).

725 83. T. N. Cordeiro, *et al.*, Interplay of protein disorder in retinoic acid receptor heterodimer and
726 its corepressor regulates gene expression. *Structure* **27**, 1270-1285.e6 (2019).

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

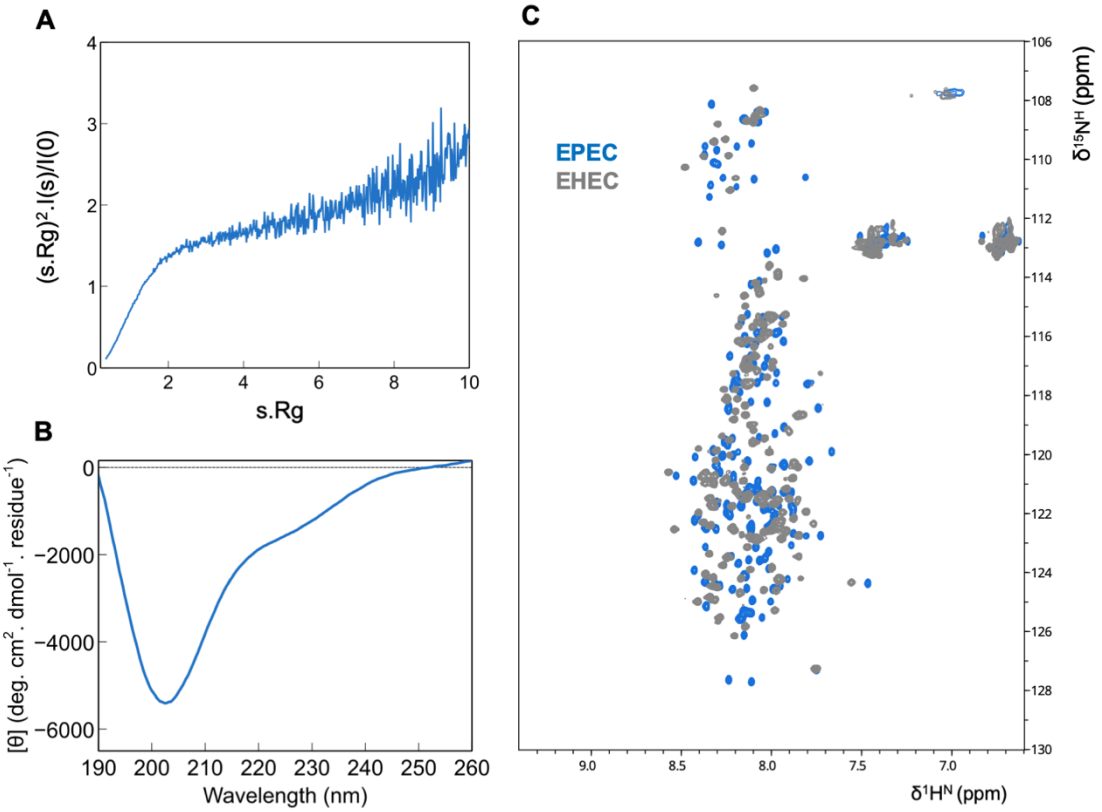
729

730 **Figures and Tables**
731
732



733
734 **Figure 1. Predicted structural disorder in A/E pathogens.** (*A*) Distribution of disorder fraction
735 displayed as violin plots for effectors (blue) and full-proteomes (red) of A/E pathogens. (*B*)
736 Accumulated fractions of the structural categories in terms of structural disorder. IDP: Intrinsically
737 disordered proteins; PDR: Proteins with intrinsically disordered regions; FRAG: Proteins with
738 fragmented-disorder; NDR: Not disordered proteins; ORD: Ordered Proteins. See definitions in
739 **Table S2.** (*C*) Fraction of ELMs vs. disorder fraction in A/E effectors. Tir, EspF, and EspF(U) display
740 a high motif and disorder content. IDP and PDR effectors are labeled. See all data in **Table S3.**
741

742



743

744 **Figure 2. C-Tir is an intrinsically disordered domain.** (*A*) Kratky representation of the SAXS
745 profile measured for C-Tir indicating its lack of compactness. (*B*) Far-UV circular dichroism of C-Tir
746 revealing the absence of high-populated structural secondary elements. (*C*) C-Tir EPEC (light blue)
747 and C-Tir EHEC (grey) [^{1}H-^{15}N]-HSQC spectra reveal the narrow ^{1}H chemical shift dispersion
748 characteristic of IDPs with ^{1}H amide backbone resonances clustering between 7.7 and 8.5 ppm.
749 All together provide a definite diagnostic of protein disorder.

750

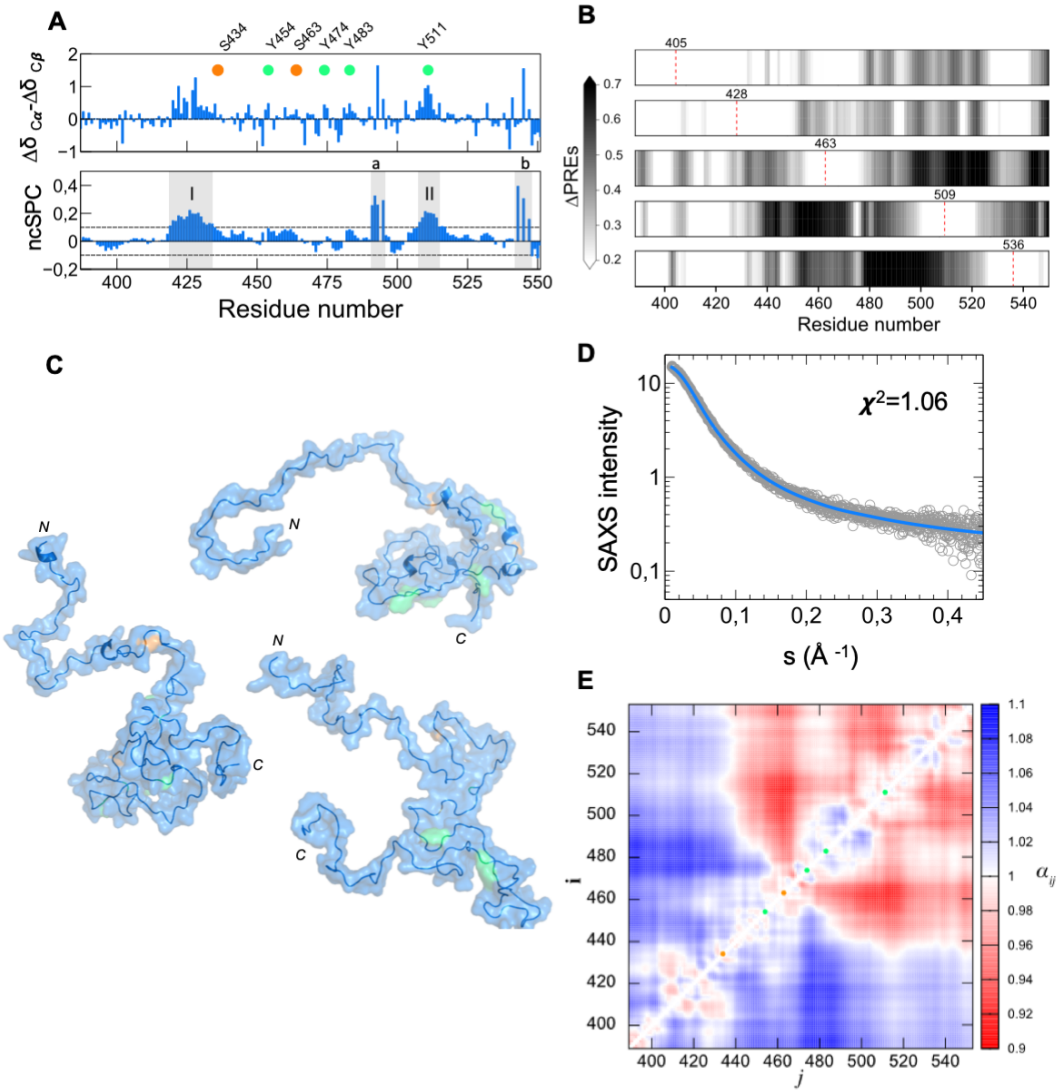


Figure 3. C-Tir is disordered with residual structure at sites of host interaction. (*A*) Secondary chemical shifts (ΔδCα-ΔδCβ) (top) and structural propensity plot (bottom) for C-Tir where the dashed-lines depict the random-coil threshold. Values above or below this threshold have α-helix or β-sheet propensities, respectively. Positive values (grey regions) show an increased α-helix tendency at segments D420-Q435 (I) and I507-A515 (II), as well as F489-V495 (a) and T543-V547 (b). (*B*) ΔPRE-values measured on V405C, S428C, S463C, S509C, and S536C C-Tir mutants as gray-scale heatmaps, from top to bottom, respectively. Color intensity reflects the relative deviations between the experimental PRE and those predicted from a random-coil model ($\Delta PRE = [I_{para}/I_{dia}]_{exp} - [I_{para}/I_{dia}]_{RC}$). Dashed red lines mark the position of MTSL tags. (*C*) Cartoon representations of 3 representative conformers of the C-Tir ensemble refined with PRE+SAXS data with serine- and tyrosine-phosphorylation sites colored in orange and green, respectively. (*D*) Logarithmic-scale representation of scattering intensity, *I(s)*, as a function of the momentum transfer, *s*, measured for C-Tir (empty gray circles). The solid line is the averaged back-calculated curve derived from PRE+SAXS-based ensemble of C-Tir (blue). (*E*) Ensemble-averaged contact map normalized to random-coil distances ($\alpha_{ij} = \langle r_{ij}\rangle_{ens}/\langle r_{ij}\rangle_{RC}$). Green and orange circles mark phosphorylation sites on C-Tir for tyrosine and serine residues, respectively.

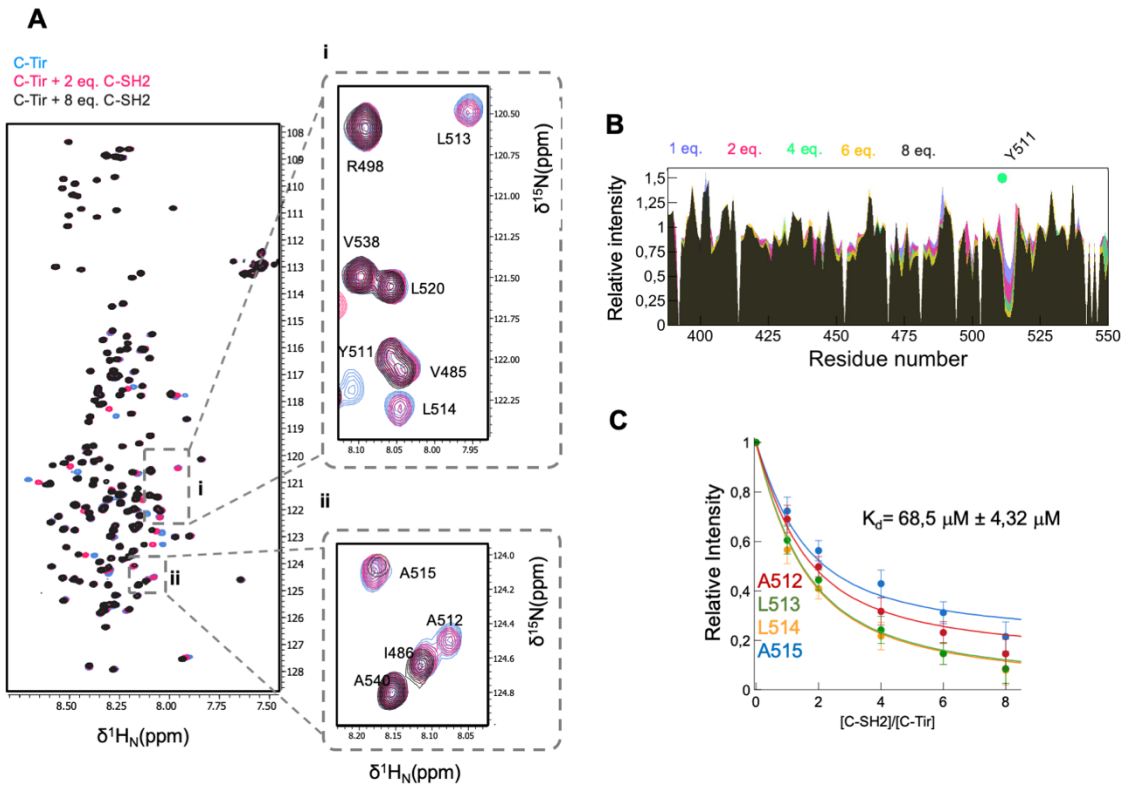


Figure 4. C-Tir can bind host C-SH2 domain of SHP-1 pre-phosphorylation. (*A*) Overlay of [^{1}H-^{15}N]-HSQC spectra of C-Tir in the absence (blue) and the presence of C-SH2 at 2.0 (pink) and 8.0 (dark grey) equivalents. Expansions (right panels) show the NMR signal of residues $A_{512}LLA_{515}$ with a significant intensity drop due to C-SH2 binding. (*B*) Relative [^{1}H-^{15}N]-peak intensities from the titration analysis. The green circle marks the tyrosine Y511 position. (*C*) Global fitting NMR quantification of the binding of C-SH2 to unphosphorylated C-Tir.

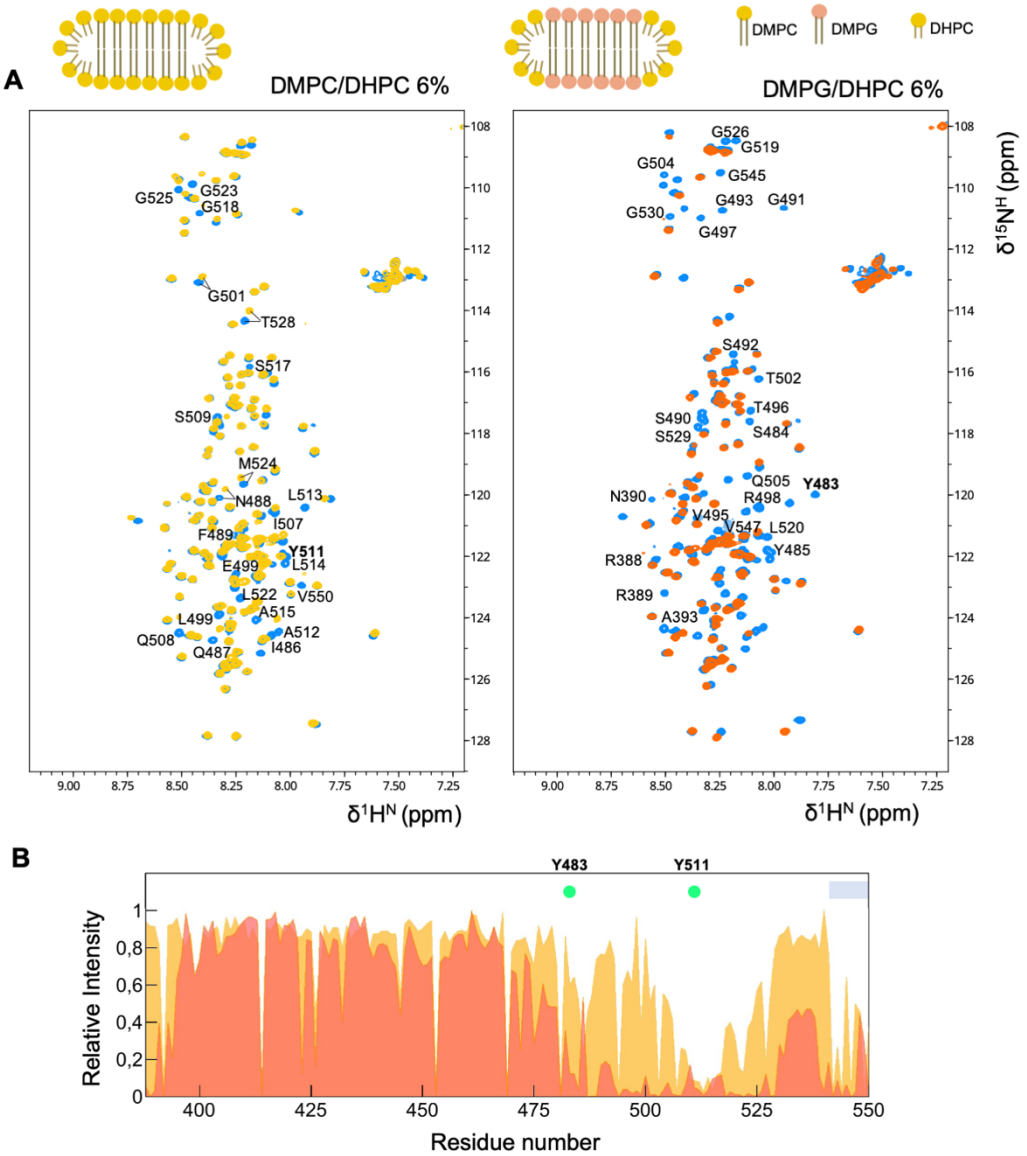


Figure 5. Lipid-binding by disordered C-Tir. (*A*) NMR [^{1}H-^{15}N]-HSQC spectra of C-Tir in the absence (blue) and the presence of DMPC/DHPC (6% w/v), yellow) or DMPG/DHPC (6% w/v), orange) bicelles. C-Tir's HSQC data in bicelles superimpose well to free C-Tir, with identical chemical shifts for most visible resonances, pointing to intrinsic disorder. Residues with NMR signal attenuation upon the 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.

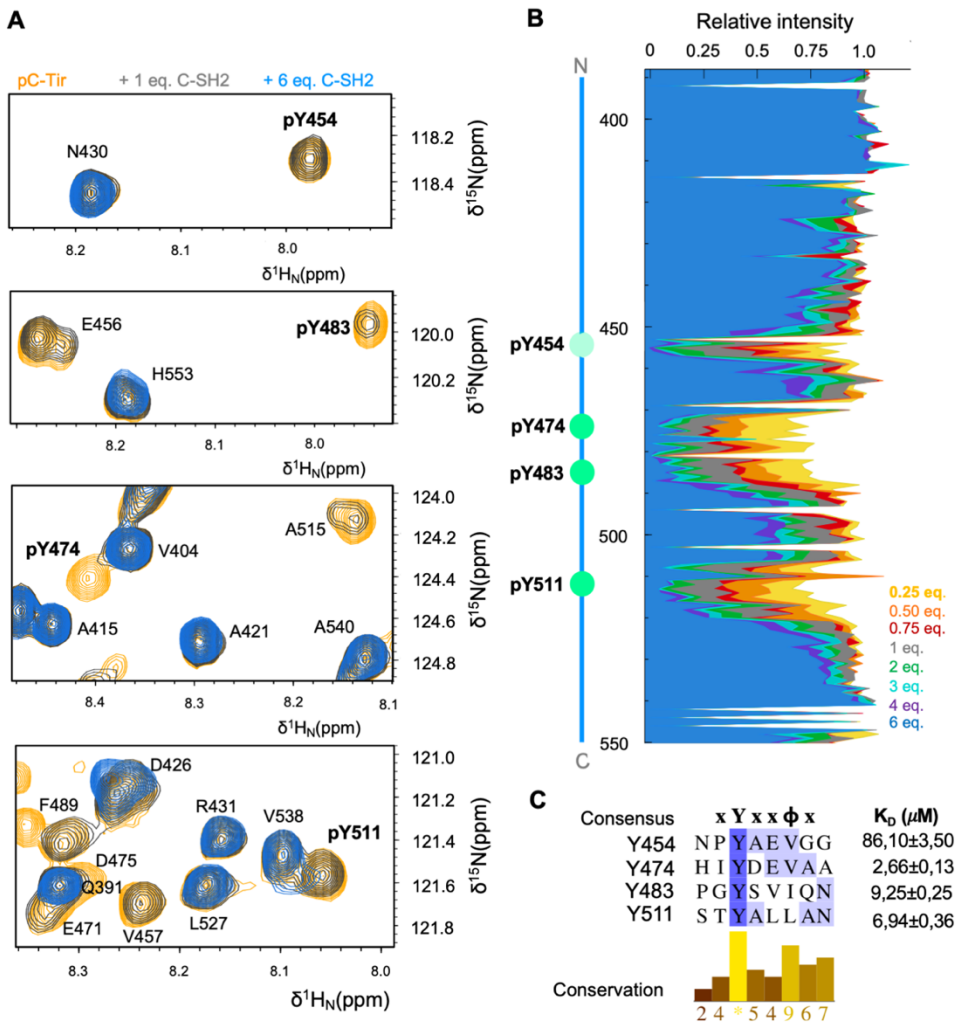


Figure 6. C-Tir contains multiple phosphorylated docking sites. (*A*) NMR-detected binding of C-SH2 to phosphotyrosine pY454, pY474, pY483, pY511 before (orange) and after adding 1.0 and 6.0 equivalents of C-SH2 (blue and grey, respectively). (*B*) Site-specific NMR signal attenuation on pC-Tir due to 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 intensities 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 bar-plot below the alignment. Green circles denote phosphorylation tyrosine-sites with color intensity reflecting the apparent relative affinity.

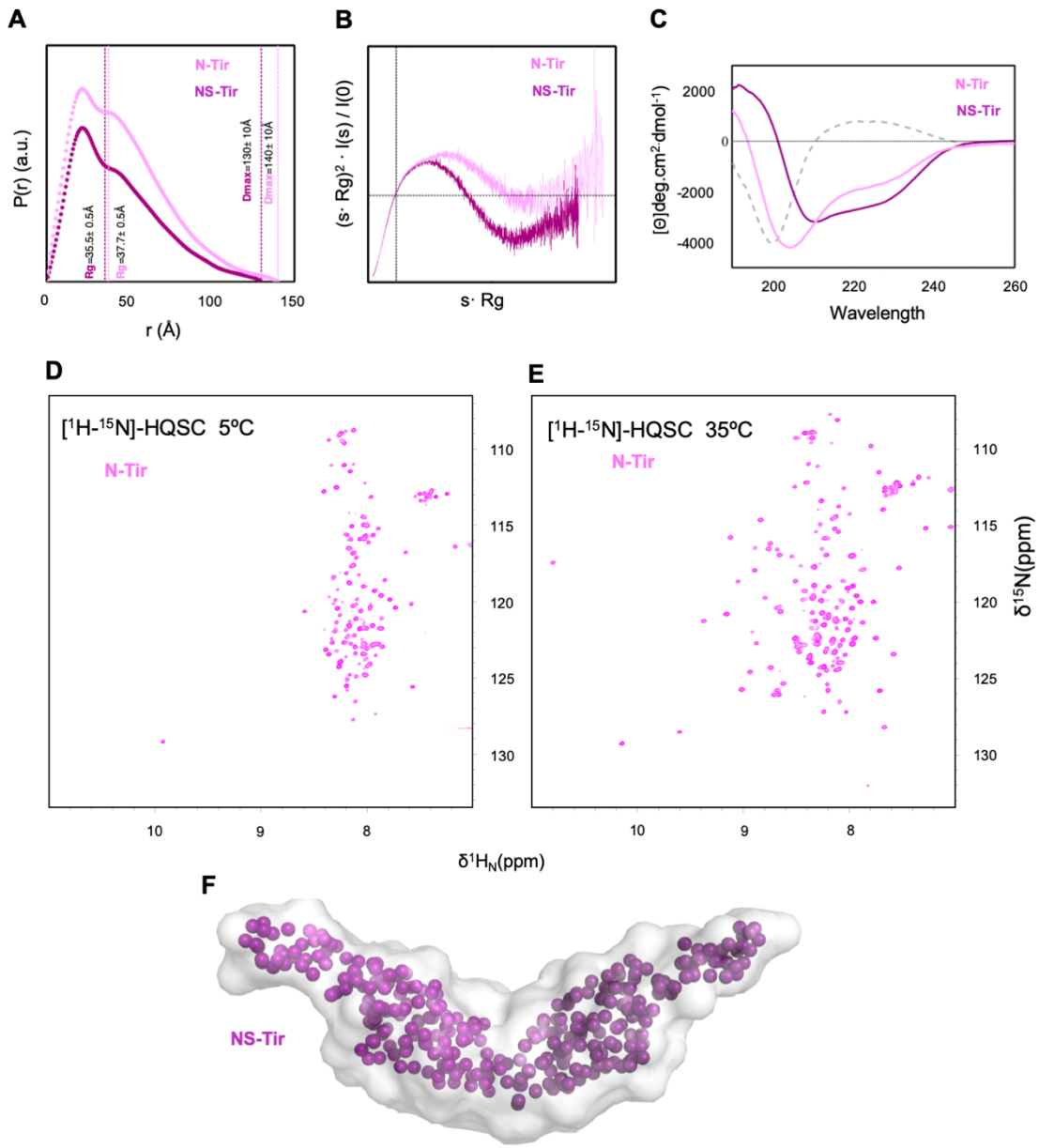


Figure 7. N-Tir is a partially disordered dimer. (*A*) Normalized pairwise distance distribution, P(r), computed from experimental SAXS curves of N-Tir (pink) and NS-Tir (purple) consistent with an elongated dimer. Dashed lines indicate the derived Rg and Dmax values. (*B*) Kratky plots of N-Tir (pink) 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.

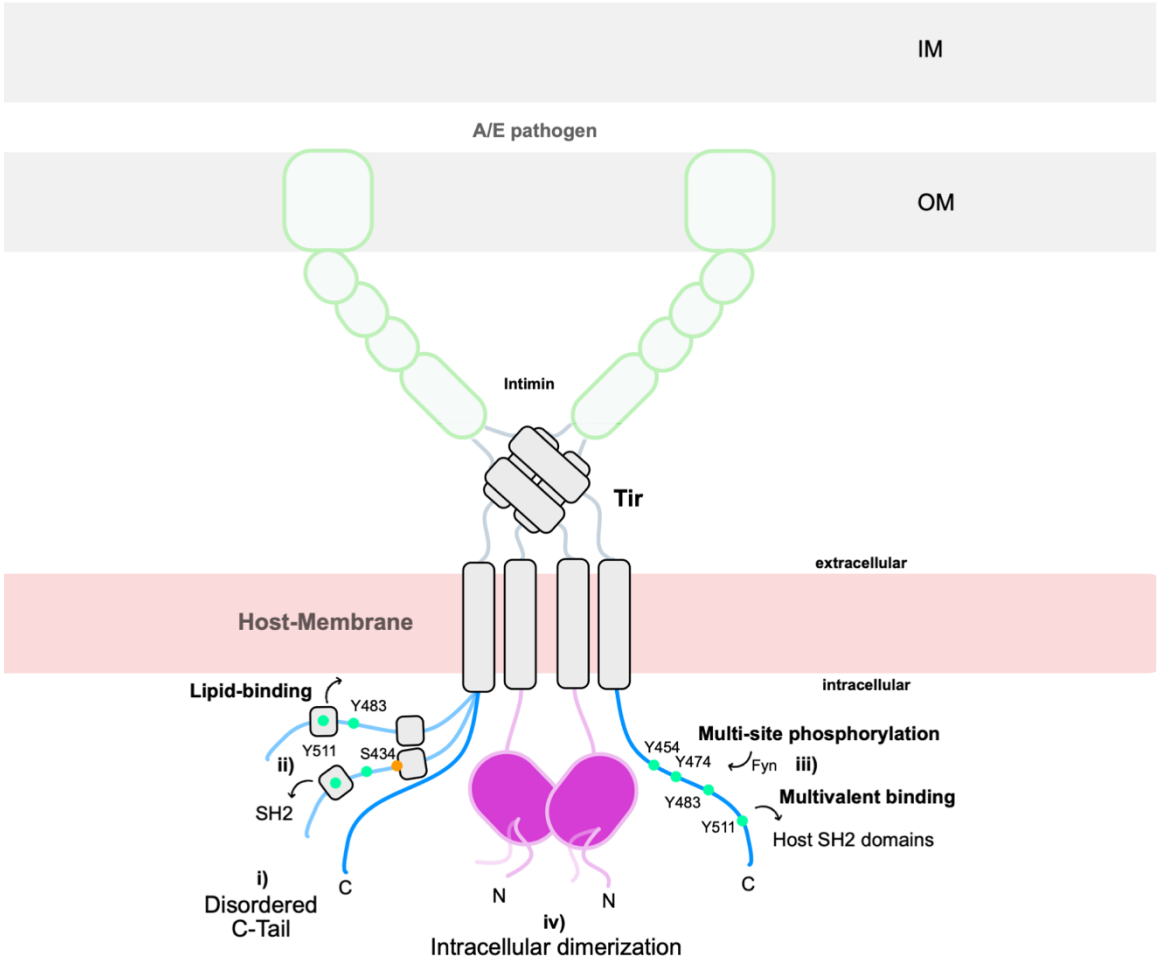


Figure 8. A proposed model of Tir structural organization. Schematics of Tir in the host plasma membrane in a hairpin-topology while bound to intimin at the bacterial outer membrane. Our data support that (i) the C-terminal intracellular region of Tir (C-Tir) membrane receptor is disordered; (ii) with transient helical structural elements involved in 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.