How pathogens use linear motifs to perturb host-cell networks

Allegra Via[1], Bora Uyar[2], Christine Brun[3,4,5], Andreas Zanzoni[3,4,*]

[1]Department of Physics, Sapienza University of Rome, 00185 Rome, Italy, [2]Structural and Computational Biology, European Molecular Biology Laboratory, 69117 Heidelberg, Germany, [3]Inserm, UMR1090 TAGC, Marseille, F-13288, France, [4]Aix-Marseille Université, UMR1090 TAGC, Marseille, F-13288, France, [5]CNRS, Marseille, F-13402, France.

* Corresponding author: Zanzoni, A. (zanzoni@tagc.univ-mrs.fr).

Keywords: molecular mimicry, short linear motifs, SLiMs, host-pathogen interactions, virulence

Abstract

Molecular mimicry is one of the powerful stratagems pathogens employ to colonise their hosts and take advantage of host cell functions to guarantee their replication and dissemination. In particular, several viruses evolved the ability to interact with host cell components through protein short linear motifs (SLiMs) that mimic host SLiMs, thus facilitating their internalisation and the manipulation of a wide range of cellular networks. Here, we present convincing evidence from the literature that motif mimicry also represents an effective widespread hijacking strategy in prokaryotic and eukaryotic parasites. Further insights into host motif mimicry would be of great help for the elucidation of the molecular mechanisms behind host-cell invasion and the development of anti-infective therapeutic strategies.

Pathogens, molecular mimicry and linear motifs

Increasing amounts of evidence indicate that phylogenetically distant pathogens show remarkably similar *modi operandi* in host cell entry and subversion [1,2]. Commonalities are expected to occur particularly at the pathogen-host interface; that is, in pathogen macromolecules involved in cytoadherence and/or cell penetration, two mechanisms for which pathogens belonging to distant clades are confronted with similar host components and pathways.

Transient protein-protein interactions are often mediated by short stretches of contiguous amino acids – known as short linear motifs (SLiMs, see Box 1) [3] – that embody functions independently of a larger sequence and structure context. In many cases, the presence of a SLiM is sufficient to promote ligand

binding, targeting, control of protein stability, and, more generally, to regulate several pathways, provided the motif is adequately exposed on the protein surface. SLiMs are mostly located in natively disordered protein regions and, if within folded domains, they tend to reside in accessible loops [3], which are evolutionary variable segments where motifs may appear or disappear as a result of single point mutations [4]. This plasticity of SLiMs makes them ideal elements to tune functionality in eukaryotic regulatory proteins. This adaptive process is facilitated by convergent evolution events, in which similar SLiMs may arise *de novo* in unrelated protein sequences. Including post-transcriptional modification sites, it has been estimated that more than 1 million linear motif functional sites may exist in the human proteome [5], a testament to the complexity of cell regulatory systems. Pathogenic organisms may, however, take advantage of host systems if their secreted proteins also contain convergently evolved SLiMs. These are the so-called "mimicry motifs"; (Box 2); short stretches of amino acids that are similar - if not identical - to host SLiMs in both composition and function. Thus, motif plasticity represents a double-edged sword or, as nicely expressed by Davey et al. [6], creates an Achilles' heel, permitting pathogens to exploit the host cell machinery itself to interact with host pathways.

The occurrence of mimicry motifs has been extensively reported in viral proteomes [6–8]. As highlighted previously [6], these organisms, probably thanks to their quickly evolving genomes, have evolved numerous host-resembling SLiMs that facilitate their internalisation into the host cell and the manipulation of a wide range of host cellular pathways involved in immune response, cell cycle and transcription regulation [6]. It is worth noting that

78 experimentally validated examples of viral mimicry could be shown for one third
79 of the motif classes in the Eukaryotic Linear Motif database (ELM) [9] and
80 recent computational studies identified a large number of potentially functional
81 eukaryotic motifs in an exhaustive set of viral genomes [7,8].
82 The extensive use of viral motif mimicry suggests that this tactic could be a
83 general practice not only in viruses but also among pathogens belonging to
84 other taxonomic domains such as bacterial and eukaryotic parasites, especially
85 if they invade host cells. For instance, Src homology 3 (SH3) interacting proline-
86 rich motifs are found in proteins from different pathogen phyla, from Bacteria to
87 Apicomplexa, where they are generally used to interact with SH3 domains
88 within the infected cell and modulate host-signalling pathways [10].
89
90 Here, we wish to highlight the growing number of motif mimicry examples being
91 observed in pathogens belonging to the cellular domains of life. While the
92 available data is much less than for viruses, it does raise the question as to
93 what extent motif mimicry represents the expression of a general strategy used
94 by pathogens to "sneak" into host cell regulatory networks.
95
96 Exploratory literature searches revealed that motif mimicry events occur both
97 in prokaryotic and eukaryotic parasites (Table 1) where they represent the
98 etiological agents of distinct diseases in both plants and animals (Table 2).
99 Mimicry motifs facilitate the adhesion to host cells and the release of pathogenic
100 effectors that can ultimately reach specific intracellular target sites and perturb
101 interaction networks. Interestingly, among the representative examples of
102 pathogen mimicry of host SLiMs reported in Table 1, there are some – such as

Arg-Glu-Asp (RGD), mitogen-activated protein kinase (MAPK) docking and SH3 domains or 14-3-3 binding motifs – that occur in proteins of pathogens spanning different phyla (Figure 1 and 2) and are also observed in viruses [6] (see Box 3).

We discuss examples from the growing literature on the subject, revealing how even evolutionary distant pathogens exploit molecular mimicry of host-like motifs to assist in host-pathogen interactions.

Mimicry motifs in prokaryotic pathogens

RGD

The extracellular tri-peptide motif RGD plays a crucial role in cell adhesion by mediating the interaction of several extracellular glycoproteins, such as thrombospondins and cell adhesion receptors with members of the integrin superfamily [11].

Several pathogenic bacteria are known to possess RGD motifs in surface proteins that allow them to establish intimate contacts with host tissues, an essential preliminary step of infection, favouring either the pathogen internalisation or the injection of specific protein effectors into the host cells [12]. For instance, the *Helicobacter pylori* CagL protein, a component of the type IV secretion system, has an RGD sequence that mediates the interaction with the host α5β3 and α5β1 integrins and is required for pathogen adhesion to gastric epithelial cells [13,14]. Other examples of RGD mimicry by pathogenic bacteria were identified in the calcium-dependent pilus biogenesis factor PiY1 [15] of *Pseudomonas aeruginosa*, an opportunistic pathogen of immuno-compromised

individuals, and in the Lipoprotein T of *Mycoplasma conjunctivae*, responsible for the infectious keratoconjunctivitis of the domestic sheep [16].

Motif mimicry is not a prerogative of bacterial proteins involved in host cell recognition. Indeed, several pathogenic effectors interfere with key intracellular cell processes involved in immune response, cytoskeleton dynamics, cell survival and transcriptional regulation by mimicking ligand-binding motifs that can be recognised by host signalling proteins [17] such as MAP kinases, 14-3-3, and Src homology 2 (SH2)/SH3 domain-containing proteins.

14-3-3 binding motifs

The 14-3-3s are phospho-serine/threonine binding proteins that play an important role in the regulation of signalling cascades [18]. Recent lines of evidence suggest an involvement of 14-3-3 proteins in mediating plant defence response at various levels [19,20]. The *Pseudomonas syringae* effector HopQ1 interacts, *via* a conserved RxxS$_p$xP phospho-motif, both with tomato and tobacco 14-3-3 proteins in a phosphorylation-dependent manner. These interactions alter the HopQ1 localization, modulate its activity and promote virulence [21,22]. Similarly, the *Xanthomonas campestris* XopQ protein binds to several members of the 14-3-3 protein family [23]. For instance, upon binding to 14-3-3 isoforms, XopQ suppresses immunity-related cell death in *Capsicum annuum* (pepper) by interacting with the TFT4 protein [23].

Mimicry motifs involved in the interaction with host 14-3-3 proteins are not limited to plant pathogens. Indeed, incG, one of the inclusion proteins of *Chlamydia trachomatis*, an obligate human intracellular pathogen, binds specifically to the 14-3-3 beta isoform, thus probably influencing the interplay between the chlamydial inclusion and host vesicular trafficking [24]. The

interaction with 14-3-3 proteins may also occur in a phosphorylation-independent fashion. This is the case for *Pseudomonas aeruginosa* exoenzyme S (ExoS), a bifunctional toxin, which interacts with human 14-3-3-zeta through unphosphorylated DALDL motifs (Figure 2A). This interaction mostly relies on hydrophobic contacts and is required for the bacterial ADP-ribosyltransferase activity, which induces host cell death in a mouse model of pneumonia [25].

MAPK docking motifs

Mitogen-activated protein kinases (MAPKs) regulate many types of cellular processes including innate immune response [26]. For this reason, several different stratagems have been adopted by intracellular pathogens to rewire immune-related signalling pathways that are controlled by MAPKs [27]. For instance, the presence of a MAPK docking motif (also known as D motif) in the phosphothreonine lyases belonging to the OspF effector family of enteric pathogens (OspF in *Shigella* and SvpC in *Salmonella*), is required for the inhibition of the host dual-phosphorylated ERK2 and p38 kinases, leading to down-regulation of the immune response [28]. Interestingly, a recent report shows that the *Campylobacter jejuni* effector CiaD (Campylobacter invasion antigen D), which also contains a MAPK docking motif, activates host ERK1/2 and p38 kinases resulting in interleukin-8 secretion [29]. In this case, the activation of MAPK signalling is necessary to promote host cell invasion.

EPIYA-related motifs

Recent studies have revealed a class of apparently unrelated bacterial effectors characterised by the presence of one or more EPIYA-related motifs (Glu-Pro-Ile-Tyr-Ala) [30]. The tyrosine of the EPIYA motif is phosphorylated by Src family

177 kinases in the host cell, thus enabling the interaction between the bacterial
178 effectors and several host SH2 domain-containing proteins [31].
179 For instance, the EPIYA-mediated interactions of the *Helicobacter pylori* type
180 IV secretion system cagA effector with the host proteins Grb2, Shp2 and Csk
181 perturb multiple cellular pathways leading to cell transformation [32].
182 Early studies of enteropathogenic *Escherichia coli* (EPEC) showed that the
183 phosphorylated EPIYA motif in membrane-inserted Tir protein is able to recruit
184 host Nck adaptor proteins which, in turn, activate N-WASP to promote host actin
185 polymerisation and actin pedestal formation [33]. More recently, researchers
186 found that members of another host adaptor family, Crk, can bind Tir
187 phosphorylated tyrosine through their SH2 domains thus competing with Nck
188 proteins [34]. Since Crk proteins do not activate N-WASP, they cannot
189 substitute for Nck in promoting actin polymerisation. Nck and Crk adaptors
190 competition for Tir binding occurs in the early stage of infection. Indeed, during
191 later stages, Abl phosphorylates both Tir and Crk, whose SH2 domain can
192 interact intramolecularly with a phosphorylated tyrosine thus reducing the
193 competition with Nck [35].
194 Notably, in enteroheamorrhagic *Escherichia coli* (EHEC), a close EPEC
195 relative, the formation of pedestals occurs independently of Tir tyrosine
196 phosphorylation [36]. Indeed, EHEC secretes an additional effector, EspF(U),
197 which interacts with Tir through the host protein IRTKS (insulin receptor tyrosine
198 kinase substrate) [37] and recruits N-WASP to activate actin assembly [38]. In
199 this case, phosphorylation is not needed as the interaction between pathogen
200 EspF(U) and host IRTKS is mediated by a C-terminal PxxP motif of EspF(U)
201 that binds to the SH3 domain of IRTKS [39,40].

PxxP motifs

Short polyproline motifs were also detected in proteins from *Mycobacterium tuberculosis* [10]. In particular, some PxxP motifs unique to this organism were observed [10]. Although experimental verification of the biological necessity of proline-rich motifs in *M. tuberculosis* is not yet available, it can be speculated that the motifs could mediate interactions with SH3-containing proteins from the host to modulate its cell signalling pathways. In this regard PPE34, a cell wall-associated/secretory Rv1917c antigen of *M. tuberculosis* that is characterised by the presence of SH3-interacting proline-rich regions, induces functional maturation of human dendritic cells, thus executing subversion of the host immune response against the pathogen.

The reported examples of bacterial SLiM mimics indicate that molecular mimicry of host proteins is a common hijacking strategy in prokaryotic pathogens. Interestingly, it was suggested that these may only represent a small subset of the approximately 100 potential mimicry relationships identified in a comparative analysis of the proteomes of 62 pathogenic and 66 non-pathogenic bacterial species [41].

Mimicry motifs in eukaryotic pathogens

Interesting examples of motif mimicry are reported in both Apicomplexa and Nematoda, as well as in pathogenic Ascomycota (a phylum of the Fungi kingdom).

Plasmodium falciparum proteins

As in bacteria, the adherence of eukaryotic pathogens to host cells is a key event during infection. In particular, the extracellular RGD cell recognition tripeptide motif is also displayed by multiple *Plasmodium falciparum* proteins. These include the erythrocyte membrane protein 1 (PfEMP1) [42] and PfTRAP (thrombospondin related anonymous protein) [43]. PfEMP1 is a parasitised erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1 [42].

PfTRAP is a transmembrane parasite protein recognising endothelial and hepatocyte receptors and playing a key role in liver cell infection of sporozoites [44,45]. More in detail, PfTRAP is stored in specialised sporozoite organelles (micronemes) and released onto the parasite cell surface at the anterior tip upon contact with a host cell and translocated to the posterior pole of the sporozoite during host-cell penetration, where the protein is cleaved releasing its extracellular portion [46,47].

Remarkably, PfTRAP also shares the cell adhesive sequence CSVTCG with the parasite circumsporozoite (CS) protein, human thrombospondin (TSP), host properdin and terminal complement proteins, and, among others, the neural adhesion molecules F-spondin and Unc-5 [48,49]. This extracellular motif, which appears in three copies in region I of human TSP and six copies in properdin, has been demonstrated to confer sulfated glycoconjugate binding properties to the hosting proteins [50,51]. In TSP, the two cysteine residues in the CSVTCG sequence are involved in two separate disulphide bonds and the threonine is O-glycosylated [52,53]. In PfTRAP, the motif mediates the attachment of the protein to the surface of hepatocyte-derived cell lines (HepG2) [50]. Incidentally, the CSVTCG sequence in malarial CS protein is not

sufficient for cell adhesion; downstream basic residues are also required (basically a Rx[RK] motif) [54].

But these are not about the only linear motif in PfTRAP that is involved in host-cell interactions. I ndeed, PfTRAP also contains intracellular SH3-domain binding PxxP motifs in its C terminus [45]. Intriguingly, PfTRAP can interact with, and be phosphorylated by, host SH3-domain containing tyrosine kinases belonging to the Src-family, suggesting a role for PfTRAP not only in binding to hepatocytes via its extracellular portion but also in human cell signalling during sporozoite invasion and homing inside the liver cells. More generally, PxxP motifs, which are commonly present in proteins from evolutionary distant species, were also found in the *P. falciparum* proteome [8]. Noticeably, some of these motifs occur in the highly conserved cytoplasmic domains of *P. falciparum* erythrocyte membrane proteins (PfEMPs) [10] and in Pfg27 [55], an abundantly expressed sexual stage–specific protein, essential for gametocytogenesis in the parasite. Both PfEMP and Pfg27 are expressed by erythrocytic forms of the parasite, though Pfg27 does not seem to localise in the red blood cells' cytoplasm [56].

PfEMPs interact with the host erythrocyte cytoskeletal protein spectrin – which contains an SH3 module – and actin, thereby assisting anchorage of PfEMP in the erythrocyte cytoskeleton. The crystal structure of Pfg27 [55] suggests that the protein can accommodate two SH3 modules simultaneously, and pull-down binding assays indicate that it is able to interact with SH3 domains from human proteins (i.e. Src, Grb2, Hck, and Fyn) [55]. However, it remains to be verified whether these interactions are specifically mediated by the polyproline motif. As genome-wide annotation studies failed to identify SH3-like domains in the

P. falciparum predicted proteome [57], the presence of PxxP motifs in Pfg27 raises questions about the natural binding partners of this protein, suggesting a possible role of Pfg27 in host-pathogen interactions.

Toxoplasma gondii

Another case of mimicry in a different species from the Apicomplexa phylum is the MAPK-docking motif detected in proteins delivered into the host intra-cellular space by the parasite *Toxoplasma gondii*. *T. gondii* is an obligate intracellular parasite causing toxoplasmosis. During its life cycle, the parasite invades host cells and multiplies within a parasitophorous vacuole (PV). The maintenance of this structure is guaranteed by a set of parasite proteins secreted from *Toxoplasma* organelles called dense granules [58]. Recent studies showed that several dense granule proteins traffic to the PV membrane and are delivered into the host intra-cellular space [59–61]. For instance, the dense granule protein GRA24 localises in the host cell nucleus and activates the host p38-alpha MAP kinase (*MAPK14*) through a stable interaction that is mediated by a MAPK-docking motif [61]. The GRA24-p38-alpha interaction triggers the expression of macrophage pro-inflammatory genes that are required for controlling early parasite replication [62]. Another dense granule protein, GRA16, also reaches the host cell nucleus and modulates the p53 tumour suppressor pathway by interacting with the HAUSP protein, a deubiquitinase coded by the gene USP7 [60]. The region responsible for the interaction with HAUSP has been identified in the N-terminal of GRA16, which indeed contains, according to the ELM database, an instance of the USP7 docking motif.

Plant parasites

Further examples of mimicry come from plant parasites. Cyst nematodes are unique phytopathogens that synthesise proteins with functional similarity to plant CLAVATA3/ESR (CLE)-like proteins [63]. Plant CLEs function as small-secreted peptide ligands that bind to extracellular receptors and activate signalling networks regulating aspects of plant growth and development. By secreting peptide mimics of plant CLEs, the nematode can developmentally reprogram root cells for the formation of unique feeding sites within host roots for its own benefit. One such example is the plant-parasitic nematode *Heterodera schachtii* CLE2 protein (HsCLE2), which shares an identical 12-amino acid CLE motif RVSPGGPDPQHH with CLE5 and CLE6 proteins from *Arabidopsis thaliana* [64]. Though nematode-secreted peptides function as molecular mimics of endogenous plant peptides to promote parasitism, for the sake of clarity, it is worth noting that CLE motifs are examples of peptide hormone [63] rather than 'classical' SLiM mimicry.

Pathogenic Fungi

Finally, instances of mimicry motifs in pathogenic Fungi have also been reported. For instance, *Candida albicans* secreted aspartic proteases 4-6 bind host cell integrins through RGD motifs thus favouring their internalisation in the host cell [65] (Figure 2B). The RGD motif is also employed by another fungal species, *Pyrenophora tritici-repentis*, which is the etiological agent of tan spot disease. In this case, RGD mediates the interaction of the pathogen protein toxin A (ToxA) with the membrane of sensitive wheat mesophyll host cells, which is a prerequisite for toxin internalisation and eventual cell death [66].

Phomopsis amygdali, the causative agent of peach and almond canker, provides an intriguing example of how some pathogens can combine toxin and

host motif usage in one to hijack the host cell. Indeed, this fungal pathogen secretes fusicoccin, a phytotoxic terpenoid mimicking in plants some of the effects of the phytohormone auxin. Fusicoccin binds to and stabilises the endogenous H^+-ATPase/14-3-3 complex, thus irreversibly activating the plasma membrane H^+-ATPase and inducing uncontrolled stomata opening [67,68]. Interestingly, the H^+-ATPase binds 14-3-3 through a classical 14-3-3 binding motif (Figure 2A). The usage of toxin mimicry was previously observed also in bacteria [69,70].

The survey of SLiM mimics in prokaryotic and eukaryotic pathogens highlights a clear unbalance between intracellular and extracellular binding motifs (Table 1). This unbalance can be also observed in the ELM database (http://elm.eu.org). However, the paucity of extracellular binding motifs should not be a surprise. Indeed, there could be multiple reasons for not favouring motif-mediated interactions in the extracellular environment. In particular, motifs tend to appear in disordered regions, which might be easily cleaved by proteases; low affinity partners could not be able to establish stable interactions; extracellular cooperative systems are simpler than the intracellular ones; last but not least, pathogens could exploit them.

Motifs indirectly required for virulence

It is important to note that the SLiM-related virulence of pathogens is not always restricted to mimicry of the host motifs. Pathogen virulence may depend on motifs that are not necessarily present in the host proteome. For instance, the transport of virulence factors from pathogens to host cells is a critical step that

depends on motif functionality. A well-studied example is the Plasmodium Export Element (PEXEL) motif (also known as 'host targeting motif' and defined as 'RxLxE/D/Q') [71,72]. The PEXEL motif can be found in Apicomplexan parasite such as *Plasmodium* falciparum [73] and Toxoplasma gondii [74]. In malaria parasites (e.g. *Plasmodium falciparum*) the PEXEL motif is required for the export of virulence proteins in malaria parasites from the parasitophorous vacuole (PV) into red blood cells [73]. In *Toxoplasma gondii*, instead, the motif is not required for the transport of such proteins. Indeed, it serves as cleavage site in certain dense granule proteins and is required for the association with the PV membrane [74]. Motifs similar to the Apicomplexan PEXEL motif are also found in different pathogens notably *Oomycete* and fungal plant pathogens. In the case of *Oomycetes,* a slightly different version of the PEXEL motif (an RxLR motif) is used to translocate virulence factors from parasitic invasive structures (called haustoria) into the plant cells. Similarly, filamentous fungi employ RxLR-like motifs to export their virulence factors into plant cells with a similar strategy [75]. Thus, convergently evolved instances of these motifs in exported proteins may be indirectly required for the exertion of the functions of virulence factors from pathogens of different phyla.

Concluding remarks

A recent survey showed that mimicry of host protein motifs is a common phenomenon in viral proteomes [6] and its pervasiveness was recently confirmed by a comprehensive analysis of more than 2000 viral genomes [8]. Cellular pathogens can use strategies that are not available to viruses, such as toxin production, and their larger genomes offer opportunities for the mimicry of

a wider range of host genes, namely a greater variety of specific strategies such as the highly selective enzymatic inactivation of cellular signalling molecules [17]. While less is known about motif mimicry in cellular pathogens than in viruses, our understanding is now growing rapidly. The evidence is increasingly convincing that motif mimicry may indeed represent a common and effective subversion stratagem for prokaryotic and eukaryotic pathogens. We anticipate that only a small portion of the pathogen mimicry motifs has been discovered.

The viral literature is larger because viral–host protein interactions are the essence of viral cell-hijacking and so researchers began to study them as soon as methods became available. Most of the work on prokaryotic and eukaryotic pathogens started later. But in addition, the detection of such motifs in more complex pathogens, as well as the identification of putative pathogenic effector proteins, is a more difficult task. Indeed, prokaryotic and eukaryotic parasites not only have much larger genomes than viruses, but they are also characterised by more sophisticated and diverse invasion mechanisms. Indeed, these pathogens may damage the host through mechanisms in addition to those as described herein. For instance, some of them can also produce enzymes capable of post-translationally modifying proteins of host cell signaling pathways [17]. Moreover, some non-invasive bacteria and fungi injure the host through the secretion and delivery into the host cells of toxins such as small molecules (such as fusicoccin), short peptides or even proteins [76]. Finally, several eukaryotic parasites, such as Apicomplexa, develop different forms throughout their life cycles and invade of a variety of cell types and even different host species [58]. This complexity offers a very large number of parasite-host interfaces, each implying numerous host-parasite transient

protein-protein interactions, which in turn provides occasions for the evolution of mimicry motifs. Clearly, this poses difficulties in the identification and experimental validation of such motifs. Incidentally, in many cases the molecular mechanisms of host-cell invasion are still obscure.

Motif mimicry is a mechanism utilised by pathogens to rewire host-signalling pathways by co-opting SLiM-mediated protein interactions to the disadvantage of the host. Perhaps, though, these mechanisms will also offer an intervention strategy that can be used for the development of drugs that target such interactions. Small molecule inhibitors and peptide mimetics have already shown promise in manipulation of SLiM-mediated interactions. Notably cilengitide, an RGD cyclic peptide, is being explored in cancer therapy trials [77] and might be adaptable to interactions mediated by RGD motifs [78]. Small molecule inhibitors have been developed for other motif interactions discussed in this manuscript, such as 14-3-3 binding motifs [79], MAPK docking motifs [80], and SH3-domain binding motifs [81]. These promising developments suggest that SLiM-mediated interactions at host-pathogen interfaces can be targeted using small molecule compounds or peptide mimetics as part of anti-infective treatment strategies.

The implementation of computational tools [82–84] and the systematic application of state-of-the-art proteomics techniques [85–87], may help discover novel mimicry events and motif occurrences in known and putative pathogenic effectors thus revealing new aspects of host invasion and subversion mechanisms.

Figure captions

Figure 1. **Representative examples of mimicry motifs found in parasites belonging to different Phyla**. Prokaryotes are depicted in dark orange and Eukaryotes in light orange). Mimicry motifs (text in grey box) bind their targets (pink boxes) in a wide range of hosts (Animals: dark blue; Plants: light blue). (A) The RGD motif is widely used by distinct pathogens to adhere to target cells either in Animals and Plants. (B) KR-rich sequences are exploited both by Proteobacteria and Apicomplexa effectors to localise in the host-cell nucleus. (C) The mimicry of MAPK docking motifs allows pathogens to modulate the host immune response. (D) PxxP and RxxK motifs are used by different pathogens to subvert several signalling pathways. See main text and Table 1 for more details.

Figure 2. **Examples of motifs binding to their targets and of identical motifs on different protein folds**. The leftmost panels show human motifs, whereas the pathogenic motifs are represented in the middle and rightmost panels. (A) Solved structures of three different 14-3-3 binding motifs in complex with their target 14-3-3 domains. Motifs are shown as blue sticks. Left: human 14-3-3 sigma in complex with the RAF proto-oncogene serine/threonine-protein kinase (RAF1) peptide hosting the $RSTS_pT$ motif (ELM: LIG_14-3-3_1); PDB: 3IQU; middle: human 14-3-3 beta/alpha in complex with the *Pseudomonas aeoruginosa* ExoS peptide containing the DALDL motif (see Table 1); PDB: 2C23; right: *Nicotiana tabacum* 14-3-3-like protein C in complex with the plasma membrane H^+ATPase peptide from *Nicotiana plumbaginifolia* (hosting

Host target	Host organism	Pathogen	Pathogen protein	Uniprot KB	PDB	Instance	Motif[a]	ELM	Ref.
14-3-3	*S. lycopersicum, N. benthamiana*	*Pseudomonas syringae*	hopQ1	Q888Y7	-	RSKSAP	Rx[^P]([ST])[^P]P	LIG_14-3-3_1	[21,22]
	C. annuum, S. lycopersicum	*Xanthomonas campestris*	XopQ	Q3BM44	-	RRAQSLP	Rxx[^P]([ST])[IVLM]x	LIG_14-3-3_2	[23]
	Homo sapiens	*Chlamydia trachomatis*	IncG	B0B9M6	-	RSRSF	[RHK][STALV]x([ST])x[PESRDIFTQ]	LIG_14-3-3_3	[24]
	Homo sapiens	*Pseudomonas aeruginosa*	ExoS[b]	Q51451	2C23	$_{424}$DALDL$_{428}$			[25]
CASP4	*Homo sapiens*	*Shigella flexneri*	OspC3	-	-	$_{333}$LSTDN$_{337}$	xYxDx		[88]
CLE-receptors (signalling receptors in plant developmental pathways)	*Solanum tuberosum*	*Globodera rostochiensis*	CLE-1	D1FNJ7	-	KRVTPGGPDPLHN KRVTPGVPDRQHR	KRxxPxGPDPxHh		[89]
			CLE-4A	D1FNJ9	-	KRVAGAGPDPIHH KRAVPAGPDPKHH			
			CLE-4B	D1FNK2	-	KRVAGAGPDPIHH KRAVPAGPDPKHH KRGAPAGPDPIHH			
			CLE-4C	D1FNK4	-	KRVAGAGPDPIHH			
			CLE-4D	D1FNK5	-	KRGAPAGPDPIHH KRVAGAGPDPIHH			
	Arabidopsis thaliana	*Heterodera glycines*	CLE1	Q9BN21	-	RLSPSGPDPHHH			[90]
			CLE2	Q86RQ1	-	RLSPSGPDPHHH			
	Arabidopsis thaliana	*Heterodera schachtii*	CLE2	F2WA39	-	RVSPGGPDPQHH			[64]
EPB41 RBC cytoskeleton	*Homo sapiens*	*Plasmodium falciparum*	MESA/PfEMP2	Q8I492	-	$_{99}$NYIECLRNAPYID$_{111}$	NYx[EK]C[LI][KIR][NT]APYID		[91]
			PF10_0378	Q8IJ23	-	$_{342}$NFCECLISAPYID$_{354}$			
			PF11_0034	Q8IIX8	-	$_{108}$KYFHCIKTAKFID$_{120}$			
			PF14_0018	Q8IM73	-	$_{349}$QYNECMKTSDYID$_{361}$			
			PFA0675w	B9ZSJ2	-	$_{113}$NYLKCVKSAPYID$_{325}$			
			PFB0925w	O96277	-	$_{57}$NYYKSIQNAPYID$_{69}$			
			PFD0095c	Q8I205	-	$_{309}$NYYKSIQNAPYID$_{322}$			
			PFF0075c	C6KSL7	-	$_{438}$NYKKCLIEAPYID$_{450}$			
			PFF1510w	C6KTE6	-	$_{438}$NYKKCLIEAPYID$_{450}$			
			PFI0130c	Q8I3C1	-	$_{321}$NYEAFLETVPYID$_{333}$			
			PFI1790w	C0H593	-	$_{342}$NYEKCLINAPYID$_{354}$			
			PFL0055c	Q8I634	-	$_{388}$NYLECCRTARHID$_{401}$			
			PFL2540w	Q8I4Q8	-	$_{356}$NYTECLKMAEYVD$_{368}$			
ERD2	*Homo sapiens*	*Vibrio cholerae*	ctxA	P01555	1XTC	KDEL	KDEL	TRG_ER_KDEL_1	[92]
Importins	*Arabidopsis thaliana*	*Ralstonia solanacearum*	PopB	Q9RBS1	1V12[c]	KRKRDDETDPNAETEGGKKKKKR	KR-rich	TRG_NLS_Bipartite_1	[93]
	Homo sapiens	*Toxoplasma gondii*	GRA10	Q5UAF5	-	RKKRRRSGKKKRGKR			[94]
Integrins	*Homo sapiens*	*Candida albicans*	SAP4	Q5A8N2	-	RGD	RGD[d]	LIG_RGD	[65]
			SAP5	P43094	2QZX	RGD			
			SAP6	Q5AC08	-	RGD			
		Helicobacter pylori	CagL	D5LL10	-	RGD			[13,14]
		Toxoplasma gondii	GRA10	Q5UAF5	-	RGD			[95]
		Plasmodium falciparum	TRAP	P16893	-	RGD			[43]
		Pseudomonas aeruginosa	PilY1	A8WE64	-	RGD			[15]
	Ovis aries	*Mycoplasma conjunctivae*	LppT	Q8KMK0	-	RGD			[16]
	Triticum spp	*Pyrenophora tritici-repentis*	ToxA	P78737	1ZLD, 1ZLE	RGD			[66]
MAPK	*Homo sapiens*	*Shigella flexneri*	OspF	Q8VSP9	-	KKPCLKLNL	[RK]xxxxΦxΦ	DOC_MAPK_1	[28]
		Toxoplasma gondii	GRA24	S8F317	-	RRGVSELPPLYI RRGVSELPPLRI			[61]
N-WASP	*Homo sapiens*	*Escherichia coli (EHEC)*	EspF(U)	C6UYI3	-	LPDVAQRLMQHLAEHGI	x{3}Φx{3}Φx{3}Φx{5}		[38,40]
PDZ domain	*Homo sapiens*	*Escherichia coli (EPEC)*	EspI[e]	B7UR60	-	IQETRV	[ST]x[VIL]$	LIG_PDZ_Class_1	[96,97]
			Map[f]	B7UMA0	-	VQDTRL			[97]
			NleH[f]	B7UR62	-	VVLSKI			[97]
RB1	*Homo sapiens*	*Shigella flexneri*	OspF	Q8VSP9	3I0U	IMCLE	[LI]xCx[DE]	LIG_Rb_LxCxE_1	[98]
SH2 domain	*Homo sapiens*	*Escherichia coli (EPEC)*	Tir[g]	B7UM99	2CI9, 2CIA	EHIYD	EPIYA-related[i]		[99]
		Helicobacter pylori	CagA[h]	Q9RF15	-	EPIYA			[100,101]
		Chlamydia trachomatis	Tarp	O84462	-	ENIYE			[30]
		Bartonella henselae	bepD	Q5QT02	-	EPLYA, NPLYE, EHLYA			[30]
			bepE	Q5QT01	-	Several EPIYA-related			[30]
			bepF	Q5QSZ9	-	TPLYA, EPLYA			[30]
		Haemophilus ducreyi	LspA1	Q7VLE8	-	EPIYG, EPVYA			[30]
			LspA2	G1UB82	-	EPIYG			[30]
		Anaplasma phagocytophilum	AnkA	A0S0L6	-	ESIYE, EDLY, other EPIYA-related			[30]
SH3 domain	*Homo sapiens*	*Escherichia coli (EHEC)*	EspF(U)[j]	P0DJ89	2KXC, 2LNH	IPPAPNWPAPTPP	xx[PV]xxP	LIG_SH3_3	[39]
		Mycobacterium tuberculosis	PPE family protein	several		multi PxxP[k]			[10]
		Plasmodium falciparum	PfTRAP	Q94662	-	multi PxxP[k]			[45]
			PfEMP1	several		multi PxxP[k]			[10]
			Pfg27	Q27336	1N81	KPIPALP APLSPR	[RKY]xxPxxP PxxPx[KR]	LIG_SH3_1	[55]

								LIG_SH3_2	
		Listeria monocytogenes	InlC[l]	P71451	-	RNNK	RxxK		[102]
Thrombospondin	*Homo sapiens*	*Plasmodium berghei*	CS protein	P06915	-	WSQCNVTCG	W[ST]xCsVTCG[d]		[103]
		Plasmodium falciparum	CS protein	P02893	2AKB[c]	WSPCSVTCG			[43]
			PfTRAP	Q94662	-	WSPCSVTCG			[44]

the QSYT$_p$V motif) and fusicoccin (in yellow), a toxin from *Pseudomonas amygdali* stabilizing the 14-3-3 /H$^+$ATPase complex; PDB: 1O9F. (B) Solved structures of three different RGD-containing proteins. RGD motifs are shown as orange sticks. Left: Human Del-1 EGF2 domain. The RGD motif forms a type II' beta turn at the tip of a long loop, dubbed the RGD finger, which is critical for integrin binding; PDB: 4D90; middle: Secreted aspartic proteinase SAP5 from *Candida albicans* (see Table 1); PDB: 2QZX; right: RGD-containing host-selective toxin Ptr ToxA from *Pyrenophora tritici-repentis* (see Table 1); PDB: 1ZLD. The occurrence of identical motifs on completely unrelated scaffolds supports the hypothesis they represent convergent evolution events.

Tables

Table 1. Representative list of mimicry motifs in Animal and Plant parasites.

[a] A 'motif' is a general consensus (expressed through the syntax of regular expressions where appropriate). The actual sequence ('instance') may slightly violate the general consensus. Regular expression syntax: letters denote a specific amino acid; x denotes any amino acid; square brackets denote a subset of allowed amino acids; round brackets indicate a position targeted for PTM after motif recognition; [^P] indicates any residues but P; Φ denotes hydrophobic residues; lower case letters denote conserved but not strictly conserved residues; $ indicates the carboxy terminus of the protein.

[b] The interaction between ExoS and 14-3-3 is phosphorylation-independent.

[c] Theoretical model

[d] Extracellular binding motif

[e] EspI interacts with *LIN7C, PDZD11, SNTA1, SNX27*.

[f] Map, NleH interact with *NHERF2*.

[g] Tir interacts with *NCK1*.

[h] cagA interacts with *GRB2, PTPN11*.

[i] The tyrosine residue serves as a phosphorylation site

[j] EspF(U) interacts with *IRTKS.*

[k] It remains to be experimentally verified if the motif directly mediates the interaction.

[l] Non canonical binding motif to *DNMBP*.

Table 2. Taxonomy and disease implication of the pathogens discussed in this review.

Pathogen	Domain	Phylum	Host	Disease
Anaplasma phagocytophilum	Bacteria	Proteobacteria	Animals (including Human)	Anaplasmosis
Bartonella henselae	Bacteria	Proteobacteria	Cat, Human	Cat-scratch disease
Chlamydia trachomatis	Bacteria	Chlamydiae	Human	Genital diseases
Escherichia coli (EHEC, EPEC)	Bacteria	Proteobacteria	Mammals (including Human)	Gastroenteritis, kidney failure
Haemophilus ducreyi	Bacteria	Proteobacteria	Human	Chancroid
Helicobacter pylori	Bacteria	Proteobacteria	Human	Gastric ulcers, stomach cancer
Listeria monocytogenes	Bacteria	Firmicutes	Animals	Listeriosis
Mycobacterium tuberculosis	Bacteria	Actinobacteria	Mammals (including Human)	Tubercolosis
Mycoplasma conjunctivae	Bacteria	Tenericutes	Domestic and wild Caprinae species	Keratoconjunctivitis
Pseudomonas aeruginosa	Bacteria	Proteobacteria	Animals (including Human), Plants	Nosocomial infections (Human), soft rot (Plants)
Pseudomonas syringae	Bacteria	Proteobacteria	Plants	Bacterial speck
Ralstonia solanacearum	Bacteria	Proteobacteria	Plants	Wilting
Shigella flexneri	Bacteria	Proteobacteria	Animals	Disentery
Vibrio cholerae	Bacteria	Proteobacteria	Human	Cholera
Xanthomonas campestris	Bacteria	Proteobacteria	Plants	Leaf spot
Candida albicans	Eukaryota	Ascomycota	Human	Candidiasis
Globodera rostochiensis	Eukaryota	Nematoda	Plants	Chlorisis, wilting
Heterodera glycines	Eukaryota	Nematoda	Plants	Chlorosis

Heterodera schachtii	Eukaryota	Nematoda	Plants	Wilting
Plasmodium berghei	Eukaryota	Apicomplexa	Mammals (not Human)	Malaria
Plasmodium falciparum	Eukaryota	Apicomplexa	Human	Malaria
Pyrenophora tritici-repentis	Eukaryota	Ascomycota	Plants	Tan spot
Toxoplasma gondii	Eukaryota	Apicomplexa	Animals (warm-blooded)	Toxoplasmosis

Boxes

Box 1. Why linear motifs are easy to imitate?

Short linear motifs (SLiMs, LMs or minimotifs) are sequences of a few (~3-10) amino acid residues [104] encoding particular molecular functions. In a cell, SLiMs mediate protein–protein interactions, cell compartment targeting, regulation by post-translational modifications including phosphorylation, acetylation, glycosylation and others [9].

Due to their limited intermolecular contacts with their partners, SLiMs tend to bind with low affinity (in the low-micromolar range) and, therefore, mediate transient interactions. They often reside in natively unstructured, or disordered, protein regions, which may become ordered upon binding. Motifs may easily arise, disappear or become dysfunctional as a result of point mutations [4]. Some are highly conserved whereas others only retain a certain pattern of charge or hydrophobicity [105].

Several resources have been developed for the collection, annotation and/or detection of SLiMs. Examples are ELM (http://elm.eu.org), Prosite (http://prosite.expasy.org), Scansite (http://scansite.mit.edu), Minimotif miner (http://mnm.engr.uconn.edu), and SLiMFinder (http://bioware.ucd.ie/slimfinder.html).

The ELM resource was established in 2003 with the goal to collect, annotate, classify and detect SLiMs [106]. In this context, SLiMs have been grouped in six classes according to the functional site: 1) proteolytic cleavage sites (CLV), for example Caspase-3 cleavage sites; 2) post-translational modification sites (MOD), for example phosphorylation sites; 3) sub-cellular targeting sites (TRG), for example nuclear localisation signals; 4) general ligand binding sites (LIG), for example SH3 binding sites; 5) docking sites (DOC), that is, motifs recruiting a modifying enzyme using a site that is distinct from the active site, such as MAPK docking sites, and 6) degradation sites (DEG), that is, specific regions of a protein sequence that direct protein polyubiquitylation and target the protein to the proteasome for degradation [9].

A classical way to represent SLiMs is using regular expressions, that is, sequences made up of characters and metacharacters expressing a consensus [107].

The short length and the high levels of evolutionary plasticity make SLiMs an ideal place for the occurrence of convergent evolution events. In particular, the accidental appearance of a given combination of amino acids that is advantageous to the interaction of pathogen with host proteins will be more frequent if such functions are enclosed in short motifs with sequence specificity but little or no structural constraints.

Box 2. Molecular mimicry.

The term "molecular mimicry" originally referred to the sharing of antigens between microorganisms and their hosts [108]. Antigenic mimicry, which allows

535 pathogens to evade or even exploit the host immune system, is a well
536 recognised mechanism and numerous examples have been reported [109].
537 Here, we refer to molecular mimicry in the broader sense adopted by Ludin et
538 al. [110]: the display of any structure by the pathogen that (i) resembles a
539 structure of the host at the molecular level and (ii) confers a benefit to the
540 pathogen because of this resemblance. Such mimicry structures include
541 peptides and hormones [63], toxins [70] and protein domains [111]; they may
542 arise by horizontal gene transfer and divergence or by convergent evolution.
543 Work over the last decades has been instrumental in identifying and describing
544 the molecular mechanisms underlying pathogen mimicry [112–115].
545 A particular case of molecular mimicry is represented by mimicry motifs; that is,
546 SLiMs occurring in pathogenic proteins that are similar or identical to host
547 SLiMs in amino acid composition and function. When such SLiMs mediate
548 interactions with one or more host proteins and when such interactions are
549 necessary or beneficial to the pathogen, they can be considered as mimicry
550 motifs. Examples of such motifs have been observed in proteins of different
551 classes of pathogens and, in particular, in proteins involved in cytoadherence
552 and host cell penetration – typical cellular processes at the host-pathogen
553 interface – where it is particularly convenient for the pathogen to evolve
554 molecular structures imitating host modules. SLiMs represent an easily
555 accessible form of mimicry thanks to their intrinsic characteristics.
556 One of the best-characterised examples of molecular mimicry in viruses (which
557 is also observed in prokaryotic and eukaryotic pathogens (see Box 3)) is the
558 polyproline P-x-x-P-x-R motif found in the Human immunodeficiency virus type
559 1 (HIV-1) Nef protein as well as in the hepatitis C virus non-structural 5A protein

(NS5A). This motif is able to interact with the SH3 domains of a variety of host cellular proteins [116,117].

Box 3. Mimicry motifs conserved in distantly related pathogens

Pathogen SLiMs imitating host motifs and mediating host-pathogen transient interactions may have independently evolved by very distant infectious agents. Some SLiMs, with tiny variations, are conserved not only across bacteria and eukaryotic pathogen species but also in viruses.

One example is the motif RSxSxP (LIG_14-3-3_1), which mediates interactions with host 14-3-3 proteins and found in a number of bacterial proteins belonging to different phyla (Table 1) and the virus AAV protein Rep68 [6].

The KDEL motif, allowing interactions with ER (endoplasmic reticulum) lumen protein-retaining receptor **(**ERD2), is conserved in the ctxA protein from *Vibrio cholerae* and in its homologous protein in phage CTX. The RGD consensus, which is used instead to interact with host integrins, is very well conserved in both viruses, prokaryotic and eukaryotic pathogens. Other examples of motifs evolved by distantly related pathogens are TxVS (LIG_PDZ), YxxV (LIG_SH2), and PxxP (LIG_SH3), which bind to PDZ, SH2, and SH3 modular domains, respectively, and are found in several proteins governing host signalling pathways (Table 1 and [6]).

Glossary

14-3-3 proteins: a family of evolutionary conserved proteins that play a key role in multiple biological processes by interacting with a plethora of client

proteins. They bind to phosphorylated serine or threonine residues.

Cytotoxicity associated immunodominant antigen (CagA): a phosphotyrosine-containing protein that is secreted by the *Helicobacter pylori* type IV secretion system. It induces morphological changes to the infected cell by interacting with proteins of host's signalling pathways including Grb2, Shp2 and Csk.

CagL protein: a component of the type IV secretion system (T4SS) of the gastric pathogen *Helicobacter pylori*. It is a specialised adhesin of the T4SS pilus interacting via an RGD motif with the host α5β3 and α5β1 integrins. It is required for pathogen adhesion to gastric epithelial cells.

Cluster of differentiation 36 (CD36): an integral membrane protein found on the surface of many cell types in vertebrate animals.

***Campylobacter* invasion antigens (Cia):** proteins exported via a type III secretion system and delivered to the host cell to promote maximal cell invasion.

CLAVATA3/Embryo Surrounding Region-related (CLAVATA3/ESR-related or CLE) gene family: a gene family composed of numerous genes that contain conserved CLE domains in various plant species and plant-parasitic nematodes. Plant CLE genes encode small proteins with an N-terminal secretion signal peptide and a conserved 14-amino-acid domain called the CLE motif at the carboxyl terminus. CLE proteins have roles in shoot, floral, and root meristem maintenance, organ size regulation, apical dominance, and vascular development.

CT10 regulator of kinase (Crk) adaptor family: a family of important adaptor molecules that participates in diverse signalling pathways and that localizes to

610 the EPEC pedestals.

611 **C-Src kinase (Csk)**: an enzyme that phosphorylates tyrosine residues located
612 in the C-terminal end of Src-family kinases.

613 **Dense granule protein 24 (GRA24)**: a *Toxoplasma gondii* protein secreted
614 from the parasitophorous vacuole to the host cell nucleus where it activates
615 host kinases using two high-affinity MAPK docking motifs.

616 **Dense granule protein 16 (GRA16):** a dense granule protein that is exported
617 through the Toxoplasma gondii vacuole membrane and reaches the host cell
618 nucleus, where it positively modulates genes involved in cell cycle progression
619 and the p53 tumor suppressor pathway.

620 **ERK2:** a serine/threonine kinase that plays a critical role in the regulation of cell
621 growth and differentiation.

622 **EspF(U) (or TccP) effector protein**: enteroheamorrhagic *Escherichia coli*
623 (EHEC) protein injected through a type III secretion system into host cells where
624 it stimulates actin polymerization by activating host WASP proteins.

625 **Exoenzyme S (ExoS)**: a *Pseudomonas aeruginosa* type III secretion effector
626 targeting multiple substrates in the host. It exerts complex effects on eukaryotic
627 cell function, including inhibition of DNA synthesis, alterations in cell
628 morphology, microvillus effacement, and loss of cellular adherence.

629 **Fusicoccin**: a phytotoxic terpenoid secreted by the fungus *Phomopsis*
630 *amygdali*. The terpenoid binds and stabilizes the host H^+-ATPase/14-3-3
631 complex, thus irreversibly activating the plasma membrane H^+-ATPase and
632 inducing uncontrolled stomata opening.

Growth factor receptor-bound protein 2 (Grb2): an adaptor protein involved in signal transduction. It binds several membrane receptors and contains one SH2 domain and two SH3 domains.

The Hrp outer protein Q (HopQ1): a type III effector secreted by *Pseudomonas syringae* effector protein. It enhances bacterial virulence and associates with host 14-3-3 proteins in a phosphorylation dependent manner.

Inclusion membrane protein G (incG): one of the transmembrane proteins of the chlamydial inclusion, a vacuole in which *Chlamydia trachomatis* developmental cycle takes place.

Integrins: a family of heterodimeric receptors that link the surface of cells to different extracellular mebrane matrix components. They mediate the transduction of cell-extracellular membrane matrix signalling.

Non-catalytic region of tyrosine kinase adaptor protein (Nck): an adaptor protein involved in transducing signals from receptor tyrosine kinases to downstream signal proteins. It contains SH2 and SH3 domains and interacts with the WASP/Arp2/3 complex to coordinate actin cytoskeletal remodelling.

Neuronal Wiskott–Aldrich syndrome protein (N-WASP): family of proteins are involved in transduction of signals from receptors on the cell surface to the actin cytoskeleton.

OspF effector family: type III secretion system effectors that down-regulate the host innate immune response.

p38 MAP kinases: a class of mitogen-activated protein kinases that are responsive to different stress stimuli such as cytokines, lipopolysaccharides, ultraviolet light, heat and osmotic shock. They are also involved in cell differentiation and apoptosis.

Pilus-biogenesis factor (PiY1): an essential, calcium-dependent regulator of *Pseudomonas aeruginosa* twitching/surface motility.

Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1): a protein encoded by the var genes. It interacts with adhesion molecules such as ICAM-1, CD36, or thrombospondin via different domains.

Thrombospondin related anonymous protein PfTRAP: a *Plasmodium falciparum* type 1 membrane protein that possesses multiple adhesive domains in its extracellular region. It is essential for sporozoite motility and for liver cell invasion.

Pfg27: *Plasmodium falciparum* sexual stage-specific protein involved in maintaining cell integrity in the uniquely long gametocytogenesis of the parasite.

Protein toxin A (ToxA):

Shp2: a tyrosine-protein phosphatase encoded by the gene PTPN11 and involved in several intracellular signalling pathways. It contains two SH2 domains.

Src homology 3 (SH3) domain: a conserved docking modules recognising and interacting with poly-proline motifs. They are commonly found in several intracellular signalling proteins.

Src homology 2 (SH2) domain: a conserved docking modules recognising and interacting with phosphorylated tyrosine residues. They are commonly found in several intracellular signalling proteins.

Translocated intimin receptor (Tir): One of the effectors delivered into the host cells by the enteropathogenic *Escherichia coli* (EPEC) type III secretion

682 system. It drives the major pathway responsible for regulating actin
683 polymerisation in the host cell.
684 **XopQ protein:** a type-III effector protein found in pytho-pathogens of the genre
685 Xanthomonas.
686
687
688 **Acknowledgements**
689 The authors are grateful to David G. Biron (CNRS, France), Toby J. Gibson
690 (EMBL, Germany), and Vincenzo Petrarca (Sapienza University, Italy) for
691 critically reading the manuscript and providing fruitful suggestions. AV
692 acknowledges the King Abdullah University of Science and Technology
693 (KAUST), Award No. KUK-I1-012-43, for funding support.
694
695 **References**

696 1 Bliska, J.B. and Casadevall, A. (2009) Intracellular pathogenic bacteria
697 and fungi–a case of convergent evolution? *Nat. Rev. Microbiol.* 7, 165–71
698 2 Romano, J.D. and Coppens, I. (2013) Host Organelle Hijackers: a similar
699 modus operandi for Toxoplasma gondii and Chlamydia trachomatis: co-
700 infection model as a tool to investigate pathogenesis. *Pathog. Dis.* DOI:
701 10.1111/2049-632X.12057
702 3 Van Roey, K. *et al.* (2014) Short linear motifs: ubiquitous and functionally
703 diverse protein interaction modules directing cell regulation. *Chem. Rev.*
704 114, 6733–78
705 4 Neduva, V. and Russell, R.B. (2005) Linear motifs: evolutionary interaction
706 switches. *FEBS Lett.* 579, 3342–5
707 5 Tompa, P. *et al.* (2014) A Million Peptide Motifs for the Molecular Biologist.
708 *Mol. Cell* 55, 161–169

6 Davey, N.E. *et al.* (2011) How viruses hijack cell regulation. *Trends Biochem. Sci.* 36, 159–69
7 Garamszegi, S. *et al.* (2013) Signatures of pleiotropy, economy and convergent evolution in a domain-resolved map of human-virus protein-protein interaction networks. *PLoS Pathog.* 9, e1003778
8 Hagai, T. *et al.* (2014) Use of host-like peptide motifs in viral proteins is a prevalent strategy in host-virus interactions. *Cell Rep.* 7, 1729–39
9 Dinkel, H. *et al.* (2014) The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res.* 42, D259–66
10 Ravi Chandra, B. *et al.* (2004) Distribution of proline-rich (PxxP) motifs in distinct proteomes: functional and therapeutic implications for malaria and tuberculosis. *Protein Eng. Des. Sel. PEDS* 17, 175–82
11 D'Souza, S.E. *et al.* (1991) Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends Biochem. Sci.* 16, 246–50
12 Hauck, C.R. *et al.* (2012) Exploitation of integrin function by pathogenic microbes. *Curr. Opin. Cell Biol.* 24, 637–44
13 Conradi, J. *et al.* (2012) Cyclic RGD peptides interfere with binding of the Helicobacter pylori protein CagL to integrins αVβ3 and α5β1. *Amino Acids* 43, 219–32
14 Barden, S. *et al.* (2013) A helical RGD motif promoting cell adhesion: crystal structures of the Helicobacter pylori type IV secretion system pilus protein CagL. *Struct. Lond. Engl. 1993* 21, 1931–41
15 Johnson, M.D.L. *et al.* (2011) Pseudomonas aeruginosa PilY1 binds integrin in an RGD- and calcium-dependent manner. *PloS One* 6, e29629
16 Zimmermann, L. *et al.* (2010) RGD motif of lipoprotein T, involved in adhesion of Mycoplasma conjunctivae to lamb synovial tissue cells. *J. Bacteriol.* 192, 3773–9
17 Alto, N.M. and Orth, K. (2012) Subversion of cell signaling by pathogens. *Cold Spring Harb. Perspect. Biol.* 4, a006114
18 Aitken, A. (2006) 14-3-3 proteins: a historic overview. *Semin. Cancer Biol.* 16, 162–72
19 Yang, X. *et al.* (2009) Arabidopsis 14-3-3 lambda is a positive regulator of RPW8-mediated disease resistance. *Plant J. Cell Mol. Biol.* 60, 539–50

20 Oh, C.-S. *et al.* (2010) Tomato 14-3-3 protein 7 positively regulates immunity-associated programmed cell death by enhancing protein abundance and signaling ability of MAPKKK {alpha}. *Plant Cell* 22, 260–72
21 Giska, F. *et al.* (2013) Phosphorylation of HopQ1, a type III effector from Pseudomonas syringae, creates a binding site for host 14-3-3 proteins. *Plant Physiol.* 161, 2049–61
22 Li, W. *et al.* (2013) The Pseudomonas syringae effector HopQ1 promotes bacterial virulence and interacts with tomato 14-3-3 proteins in a phosphorylation-dependent manner. *Plant Physiol.* 161, 2062–74
23 Teper, D. *et al.* (2014) Xanthomonas euvesicatoria type III effector XopQ interacts with tomato and pepper 14-3-3 isoforms to suppress effector-triggered immunity. *Plant J. Cell Mol. Biol.* 77, 297–309
24 Scidmore, M.A. and Hackstadt, T. (2001) Mammalian 14-3-3beta associates with the Chlamydia trachomatis inclusion membrane via its interaction with IncG. *Mol. Microbiol.* 39, 1638–50
25 Ottmann, C. *et al.* (2007) Phosphorylation-independent interaction between 14-3-3 and exoenzyme S: from structure to pathogenesis. *EMBO J.* 26, 902–13
26 Cuadrado, A. and Nebreda, A.R. (2010) Mechanisms and functions of p38 MAPK signalling. *Biochem. J.* 429, 403–17
27 Krachler, A.M. *et al.* (2011) Manipulation of kinase signaling by bacterial pathogens. *J. Cell Biol.* 195, 1083–92
28 Zhu, Y. *et al.* (2007) Structural Insights into the Enzymatic Mechanism of the Pathogenic MAPK Phosphothreonine Lyase. *Mol. Cell* 28, 899–913
29 Samuelson, D.R. *et al.* (2013) The Campylobacter jejuni CiaD effector protein activates MAP kinase signaling pathways and is required for the development of disease. *Cell Commun. Signal. CCS* 11, 79
30 Hayashi, T. *et al.* (2013) Bacterial EPIYA effectors–where do they come from? What are they? Where are they going? *Cell. Microbiol.* 15, 377–85
31 Selbach, M. *et al.* (2009) Host Cell Interactome of Tyrosine-Phosphorylated Bacterial Proteins. *Cell Host Microbe* 5, 397–403

32 Hatakeyama, M. (2008) Linking epithelial polarity and carcinogenesis by multitasking Helicobacter pylori virulence factor CagA. *Oncogene* 27, 7047–54

33 Goosney, D.L. *et al.* (2001) Recruitment of cytoskeletal and signaling proteins to enteropathogenic and enterohemorrhagic Escherichia coli pedestals. *Infect. Immun.* 69, 3315–22

34 Nieto-Pelegrin, E. *et al.* (2014) Crk adaptors negatively regulate actin polymerization in pedestals formed by enteropathogenic Escherichia coli (EPEC) by binding to Tir effector. *PLoS Pathog.* 10, e1004022

35 Swimm, A. *et al.* (2004) Enteropathogenic Escherichia coli use redundant tyrosine kinases to form actin pedestals. *Mol. Biol. Cell* 15, 3520–3529

36 Navarro-Garcia, F. *et al.* (2013) Actin cytoskeleton manipulation by effector proteins secreted by diarrheagenic Escherichia coli pathotypes. *BioMed Res. Int.* 2013, 374395

37 Weiss, S.M. *et al.* (2009) IRSp53 links the enterohemorrhagic E. coli effectors Tir and EspFU for actin pedestal formation. *Cell Host Microbe* 5, 244–58

38 Sallee, N.A. *et al.* (2008) The pathogen protein EspF(U) hijacks actin polymerization using mimicry and multivalency. *Nature* 454, 1005–8

39 Aitio, O. *et al.* (2010) Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly. *Proc. Natl. Acad. Sci. U. S. A.* 107, 21743–8

40 Aitio, O. *et al.* (2012) Enterohaemorrhagic Escherichia Coli Exploits a Tryptophan Switch to Hijack Host F-Actin Assembly. *Structure* 20, 1692–1703

41 Doxey, A.C. and McConkey, B.J. (2013) Prediction of molecular mimicry candidates in human pathogenic bacteria. *Virulence* 4, 453–466

42 Baruch, D.I. *et al.* (1996) Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proc. Natl. Acad. Sci. U. S. A.* 93, 3497–502

804 43 Robson, K.J. *et al.* (1988) A highly conserved amino-acid sequence in
805 thrombospondin, properdin and in proteins from sporozoites and blood
806 stages of a human malaria parasite. *Nature* 335, 79–82
807 44 Müller, H.M. *et al.* (1993) Thrombospondin related anonymous protein
808 (TRAP) of Plasmodium falciparum in parasite-host cell interactions.
809 *Parassitologia* 35 Suppl, 69–72
810 45 Akhouri, R.R. *et al.* (2008) Role of Plasmodium falciparum
811 thrombospondin-related anonymous protein in host-cell interactions.
812 *Malar. J.* 7, 63
813 46 Yuda, M. and Ishino, T. (2004) Liver invasion by malarial parasites--how
814 do malarial parasites break through the host barrier? *Cell. Microbiol.* 6,
815 1119–1125
816 47 Morahan, B.J. *et al.* (2009) No TRAP, no invasion. *Trends Parasitol.* 25,
817 77–84
818 48 Lawler, J. and Hynes, R.O. (1986) The structure of human
819 thrombospondin, an adhesive glycoprotein with multiple calcium-binding
820 sites and homologies with several different proteins. *J. Cell Biol.* 103,
821 1635–48
822 49 Tuszynski, G.P. *et al.* (1992) Biological activities of peptides and peptide
823 analogues derived from common sequences present in thrombospondin,
824 properdin, and malarial proteins. *J. Cell Biol.* 116, 209–17
825 50 Müller, H.M. *et al.* (1993) Thrombospondin related anonymous protein
826 (TRAP) of Plasmodium falciparum binds specifically to sulfated
827 glycoconjugates and to HepG2 hepatoma cells suggesting a role for this
828 molecule in sporozoite invasion of hepatocytes. *EMBO J.* 12, 2881–9
829 51 Pancake, S.J. *et al.* (1992) Malaria sporozoites and circumsporozoite
830 proteins bind specifically to sulfated glycoconjugates. *J. Cell Biol.* 117,
831 1351–7
832 52 Hofsteenge, J. *et al.* (2001) C-mannosylation and O-fucosylation of the
833 thrombospondin type 1 module. *J. Biol. Chem.* 276, 6485–98
834 53 Tan, K. *et al.* (2002) Crystal structure of the TSP-1 type 1 repeats: a novel
835 layered fold and its biological implication. *J. Cell Biol.* 159, 373–82

54 Gantt, S.M. *et al.* (1997) Cell adhesion to a motif shared by the malaria circumsporozoite protein and thrombospondin is mediated by its glycosaminoglycan-binding region and not by CSVTCG. *J. Biol. Chem.* 272, 19205–13

55 Sharma, A. *et al.* (2003) Structure of a gametocyte protein essential for sexual development in Plasmodium falciparum. *Nat. Struct. Biol.* 10, 197–203

56 Camarda, G. *et al.* (2010) Regulated oligomerisation and molecular interactions of the early gametocyte protein Pfg27 in Plasmodium falciparum sexual differentiation. *Int. J. Parasitol.* 40, 663–673

57 Aravind, L. *et al.* (2003) Plasmodium biology: genomic gleanings. *Cell* 115, 771–785

58 Plattner, F. and Soldati-Favre, D. (2008) Hijacking of host cellular functions by the Apicomplexa. *Annu. Rev. Microbiol.* 62, 471–87

59 Rosowski, E.E. *et al.* (2011) Strain-specific activation of the NF-kappaB pathway by GRA15, a novel Toxoplasma gondii dense granule protein. *J. Exp. Med.* 208, 195–212

60 Bougdour, A. *et al.* (2013) Host cell subversion by Toxoplasma GRA16, an exported dense granule protein that targets the host cell nucleus and alters gene expression. *Cell Host Microbe* 13, 489–500

61 Braun, L. *et al.* (2013) A Toxoplasma dense granule protein, GRA24, modulates the early immune response to infection by promoting a direct and sustained host p38 MAPK activation. *J. Exp. Med.* 210, 2071–86

62 Mashayekhi, M. *et al.* (2011) CD8α(+) dendritic cells are the critical source of interleukin-12 that controls acute infection by Toxoplasma gondii tachyzoites. *Immunity* 35, 249–59

63 Mitchum, M.G. *et al.* (2012) Role of nematode peptides and other small molecules in plant parasitism. *Annu. Rev. Phytopathol.* 50, 175–195

64 Wang, J. *et al.* (2011) Identification of potential host plant mimics of CLAVATA3/ESR (CLE)-like peptides from the plant-parasitic nematode Heterodera schachtii. *Mol. Plant Pathol.* 12, 177–86

65 Wu, H. *et al.* (2013) Candida albicans secreted aspartic proteases 4-6 induce apoptosis of epithelial cells by a novel Trojan horse mechanism. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 27, 2132–44
66 Manning, V.A. *et al.* (2008) The Arg-Gly-Asp-containing, solvent-exposed loop of Ptr ToxA is required for internalization. *Mol. Plant-Microbe Interact. MPMI* 21, 315–25
67 Würtele, M. *et al.* (2003) Structural view of a fungal toxin acting on a 14-3-3 regulatory complex. *EMBO J.* 22, 987–94
68 Camoni, L. *et al.* (2013) The phytotoxin fusicoccin, a selective stabilizer of 14-3-3 interactions? *IUBMB Life* 65, 513–7
69 Jørgensen, R. *et al.* (2005) Exotoxin A-eEF2 complex structure indicates ADP ribosylation by ribosome mimicry. *Nature* 436, 979–984
70 Yates, S.P. *et al.* (2006) Stealth and mimicry by deadly bacterial toxins. *Trends Biochem. Sci.* 31, 123–133
71 Marti, M. *et al.* (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* 306, 1930–3
72 Hiller, N.L. *et al.* (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 306, 1934–7
73 Bonnefoy, S. and Ménard, R. (2008) Deconstructing export of malaria proteins. *Cell* 134, 20–2
74 Hsiao, C.-H.C. *et al.* (2013) A HT/PEXEL motif in Toxoplasma dense granule proteins is a signal for protein cleavage but not export into the host cell. *Traffic Cph. Den.* 14, 519–531
75 Petre, B. and Kamoun, S. (2014) How do filamentous pathogens deliver effector proteins into plant cells? *PLoS Biol.* 12, e1001801
76 Henkel, J.S. *et al.* (2010) Toxins from bacteria. *EXS* 100, 1–29
77 Scaringi, C. *et al.* (2012) Integrin inhibitor cilengitide for the treatment of glioblastoma: a brief overview of current clinical results. *Anticancer Res.* 32, 4213–23
78 Mas-Moruno, C. *et al.* (2010) Cilengitide: the first anti-angiogenic small molecule drug candidate design, synthesis and clinical evaluation. *Anticancer Agents Med. Chem.* 10, 753–68

79 Milroy, L.-G. *et al.* (2013) Stabilization and inhibition of protein-protein interactions: the 14-3-3 case study. *ACS Chem. Biol.* 8, 27–35
80 Willemen, H.L.D.M. *et al.* (2014) A novel p38 MAPK docking-groove-targeted compound is a potent inhibitor of inflammatory hyperalgesia. *Biochem. J.* 459, 427–39
81 Cussac, D. *et al.* (1999) A Sos-derived peptidimer blocks the Ras signaling pathway by binding both Grb2 SH3 domains and displays antiproliferative activity. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 13, 31–8
82 Arnold, R. *et al.* (2009) Sequence-based prediction of type III secreted proteins. *PLoS Pathog.* 5, e1000376
83 Wang, Y. *et al.* (2014) Prediction of bacterial type IV secreted effectors by C-terminal features. *BMC Genomics* 15, 50
84 Ruhanen, H. *et al.* (2014) Potential of known and short prokaryotic protein motifs as a basis for novel peptide-based antibacterial therapeutics: a computational survey. *Front. Microbiol.* 5, 4
85 Dyer, M.D. *et al.* (2010) The human-bacterial pathogen protein interaction networks of Bacillus anthracis, Francisella tularensis, and Yersinia pestis. *PloS One* 5, e12089
86 Jäger, S. *et al.* (2012) Global landscape of HIV-human protein complexes. *Nature* 481, 365–70
87 Panek, J. *et al.* (2014) Hijacking of Host Cellular Functions by an Intracellular Parasite, the Microsporidian Anncaliia algerae. *PloS One* 9, e100791
88 Kobayashi, T. *et al.* (2013) The Shigella OspC3 effector inhibits caspase-4, antagonizes inflammatory cell death, and promotes epithelial infection. *Cell Host Microbe* 13, 570–83
89 Lu, S.-W. *et al.* (2009) Structural and functional diversity of CLAVATA3/ESR (CLE)-like genes from the potato cyst nematode Globodera rostochiensis. *Mol. Plant-Microbe Interact. MPMI* 22, 1128–42
90 Wang, J. *et al.* (2010) Dual roles for the variable domain in protein trafficking and host-specific recognition of Heterodera glycines CLE effector proteins. *New Phytol.* 187, 1003–17

91 Kilili, G.K. and LaCount, D.J. (2011) An erythrocyte cytoskeleton-binding motif in exported Plasmodium falciparum proteins. *Eukaryot. Cell* 10, 1439–47

92 Majoul, I. *et al.* (1998) KDEL receptor (Erd2p)-mediated retrograde transport of the cholera toxin A subunit from the Golgi involves COPI, p23, and the COOH terminus of Erd2p. *J. Cell Biol.* 143, 601–12

93 Guéneron, M. *et al.* (2000) Two novel proteins, PopB, which has functional nuclear localization signals, and PopC, which has a large leucine-rich repeat domain, are secreted through the hrp-secretion apparatus of Ralstonia solanacearum. *Mol. Microbiol.* 36, 261–77

94 Ahn, H.-J. *et al.* (2007) Nucleolar translocalization of GRA10 of Toxoplasma gondii transfectionally expressed in HeLa cells. *Korean J. Parasitol.* 45, 165–74

95 Ahn, H.-J. *et al.* (2005) Host cell binding of GRA10, a novel, constitutively secreted dense granular protein from Toxoplasma gondii. *Biochem. Biophys. Res. Commun.* 331, 614–20

96 Lee, S.F. *et al.* (2008) A C-terminal class I PDZ binding motif of EspI/NleA modulates the virulence of attaching and effacing Escherichia coli and Citrobacter rodentium. *Cell. Microbiol.* 10, 499–513

97 Martinez, E. *et al.* (2010) Binding to Na(+) /H(+) exchanger regulatory factor 2 (NHERF2) affects trafficking and function of the enteropathogenic Escherichia coli type III secretion system effectors Map, EspI and NleH. *Cell. Microbiol.* 12, 1718–31

98 Zurawski, D.V. *et al.* (2009) Shigella flexneri type III secretion system effectors OspB and OspF target the nucleus to downregulate the host inflammatory response via interactions with retinoblastoma protein. *Mol. Microbiol.* 71, 350–68

99 Gruenheid, S. *et al.* (2001) Enteropathogenic E. coli Tir binds Nck to initiate actin pedestal formation in host cells. *Nat. Cell Biol.* 3, 856–9

100 Mimuro, H. *et al.* (2002) Grb2 is a key mediator of helicobacter pylori CagA protein activities. *Mol. Cell* 10, 745–55

101 Jones, K.R. *et al.* (2009) Polymorphism in the CagA EPIYA motif impacts development of gastric cancer. *J. Clin. Microbiol.* 47, 959–68

965 102 Rajabian, T. *et al.* (2009) The bacterial virulence factor InlC perturbs
966 apical cell junctions and promotes cell-to-cell spread of Listeria. *Nat. Cell*
967 *Biol.* 11, 1212–8
968 103 Chatterjee, S. *et al.* (1995) A conserved peptide sequence of the
969 Plasmodium falciparum circumsporozoite protein and antipeptide
970 antibodies inhibit Plasmodium berghei sporozoite invasion of Hep-G2 cells
971 and protect immunized mice against P. berghei sporozoite challenge.
972 *Infect. Immun.* 63, 4375–4381
973 104 Davey, N.E. *et al.* (2012) Attributes of short linear motifs. *Mol. Biosyst.*
974 8, 268–81
975 105 Gould, C.M. *et al.* (2010) ELM: the status of the 2010 eukaryotic linear
976 motif resource. *Nucleic Acids Res.* 38, D167–80
977 106 Puntervoll, P. I *et al.* (2003) ELM server: A new resource for
978 investigating short functional sites in modular eukaryotic proteins. *Nucleic*
979 *Acids Res.* 31, 3625–30
980 107 Aasland, R. *et al.* (2002) Normalization of nomenclature for peptide
981 motifs as ligands of modular protein domains. *FEBS Lett.* 513, 141–4
982 108 Damian, R.T. (1964) Molecular Mimicry: Antigen Sharing by Parasite
983 and Host and Its Consequences. *Am. Nat.* 98, 129–149
984 109 Kohm, A.P. *et al.* (2003) Mimicking the way to autoimmunity: an
985 evolving theory of sequence and structural homology. *Trends Microbiol.*
986 11, 101–5
987 110 Ludin, P. *et al.* (2011) Genome-wide identification of molecular mimicry
988 candidates in parasites. *PloS One* 6, e17546
989 111 Dean, P. (2011) Functional domains and motifs of bacterial type III
990 effector proteins and their roles in infection. *FEMS Microbiol. Rev.* 35,
991 1100–25
992 112 Stebbins, C.E. and Galán, J.E. (2001) Structural mimicry in bacterial
993 virulence. *Nature* 412, 701–705
994 113 Sikora, S. *et al.* (2005) Convergent evolution as a mechanism for
995 pathogenic adaptation. *Trends Microbiol.* 13, 522–527
996 114 Elde, N.C. and Malik, H.S. (2009) The evolutionary conundrum of
997 pathogen mimicry. *Nat. Rev. Microbiol.* 7, 787–797

115 Gowthaman, U. and Eswarakumar, V.P. (2013) Molecular mimicry: good artists copy, great artists steal. *Virulence* 4, 433–434

116 Lee, C.H. *et al.* (1995) A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. *EMBO J.* 14, 5006–15

117 Shelton, H. and Harris, M. (2008) Hepatitis C virus NS5A protein binds the SH3 domain of the Fyn tyrosine kinase with high affinity: mutagenic analysis of residues within the SH3 domain that contribute to the interaction. *Virol. J.* 5, 24

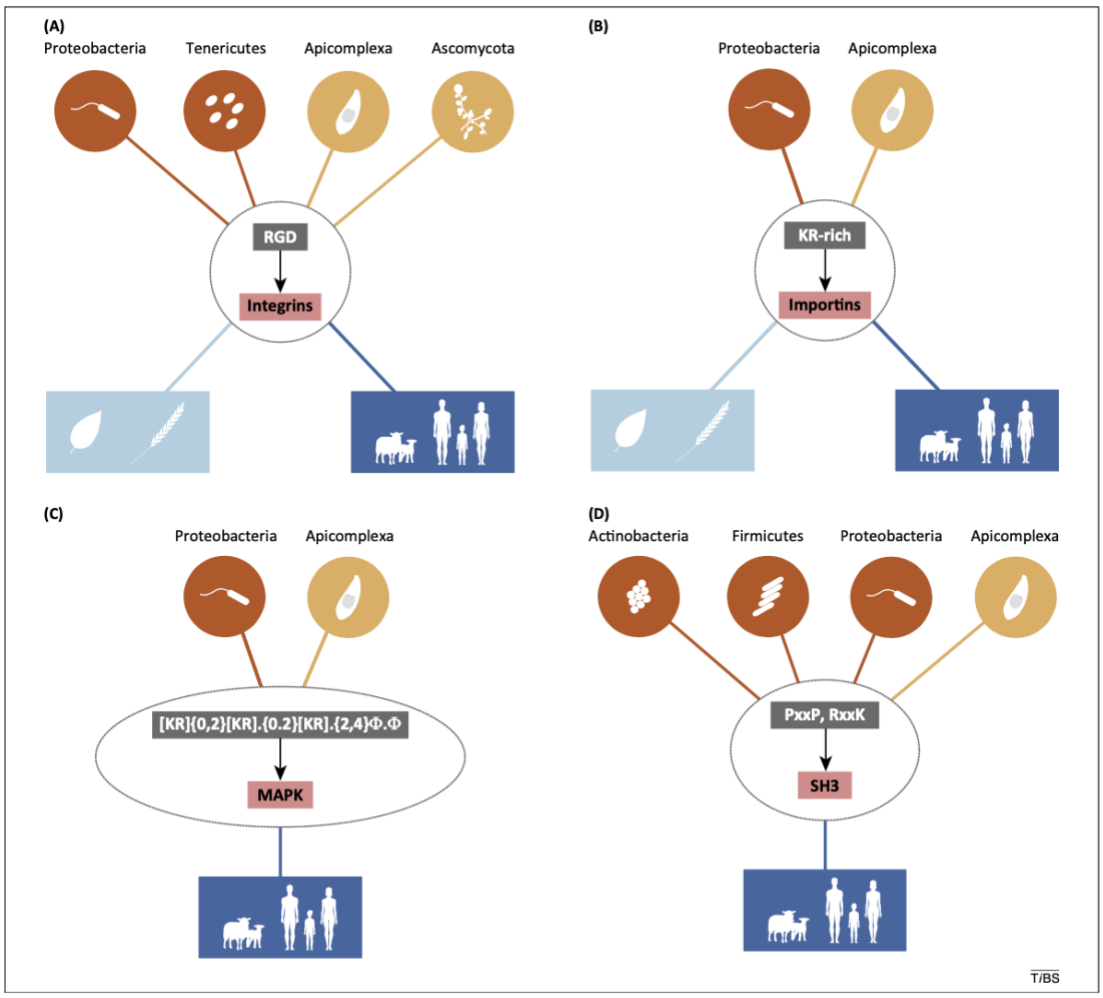

Figure 1. Representative examples of mimicry motifs found in parasites belonging to various phyla. Prokaryotes are depicted in dark orange and eukaryotes in light orange. Mimicry motifs (text in grey box) bind their targets (pink boxes) in a wide range of hosts (animals, dark blue; plants, light blue). **(A)** The Arg–Glu–Asp (RGD) motif is widely used by distinct pathogens to adhere to target cells in both animals and plants. **(B)** Lys and Arg (KR)-rich sequences are exploited by both Proteobacteria and Apicomplexa effectors for localisation in the host cell nucleus. **(C)** Mimicry of mitogen-activated protein kinase (MAPK)-docking motifs allows pathogens to modulate the host immune response. **(D)** PxxP and RxxK motifs are used by various pathogens to subvert several signalling pathways. See text and Table 1 for more details.

1015

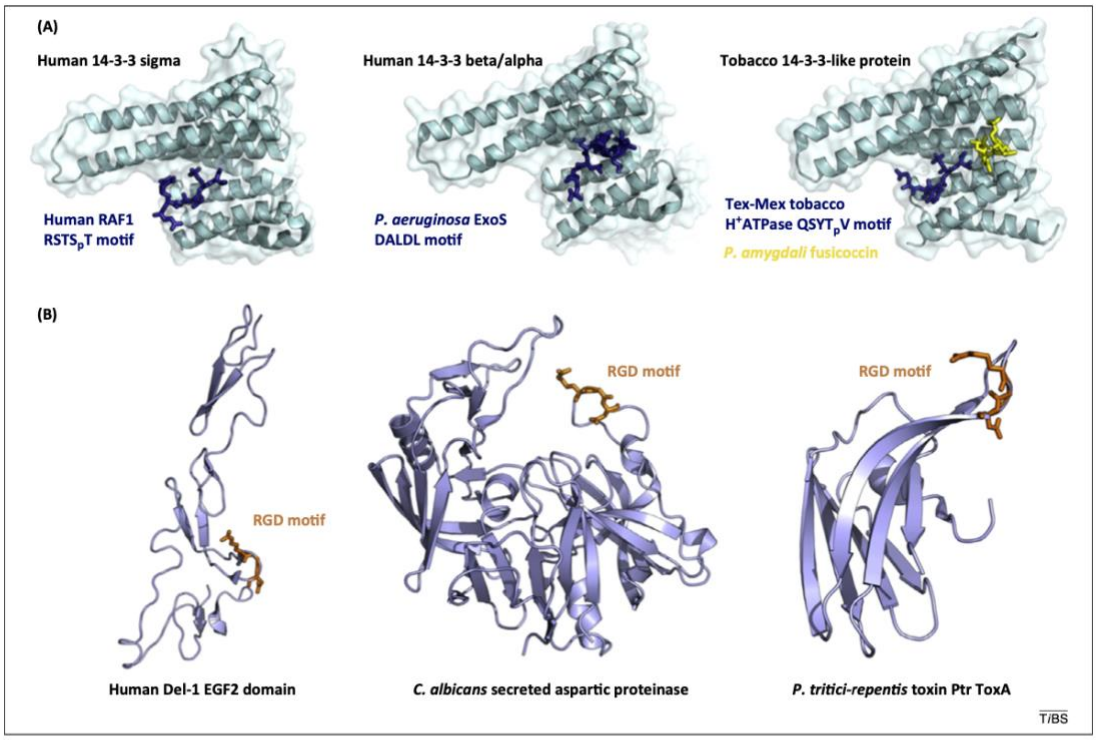

Figure 2. Examples of motifs binding to their targets and of identical motifs on different protein folds. The left panels show human motifs, whereas the pathogenic motifs are represented in the middle and right panels. **(A)** Solved structures of three different 14-3-3-binding motifs in complex with their target 14-3-3 domains. Motifs are shown as blue sticks. Left: Human 14-3-3 sigma in complex with the RAF proto-oncogene serine/threonine protein kinase (RAF1) peptide hosting the RSTS$_p$T motif [ELM: LIG_14-3-3_1; Protein Data Bank (PDB): 3IQU]. Middle: Human 14-3-3 beta/alpha in complex with the *Pseudomonas aeruginosa* ExoS peptide containing the DALDL motif (Table 1) (PDB: 2C23). Right: *Nicotiana tabacum* 14-3-3-like protein C in complex with the plasma membrane H$^+$-ATPase peptide from *Nicotiana plumbaginifolia* (hosting the QSYT$_p$V motif) and fusicoccin (in yellow), a toxin from *Pseudomonas amygdali* stabilising the 14-3-3/H$^+$-ATPase complex (PDB: 1O9F). **(B)** Solved structures of three different Arg–Glu–Asp (RGD)-containing proteins. RGD motifs are shown as orange sticks. Left: Human Del-1 EGF2 domain. The RGD motif forms a type II′ beta turn at the tip of a long loop, dubbed the RGD finger, that is critical for integrin binding (PDB: 4D90). Middle: Secreted aspartic proteinase 5 (SAP5) from *Candida albicans* (Table 1) (PDB: 2QZX). Right: RGD-containing host-selective toxin Ptr toxin A (ToxA) from *Pyrenophora tritici-repentis* (Table 1) (PDB: 1ZLD). The occurrence of identical motifs on completely unrelated scaffolds supports the hypothesis they represent convergent evolution events.

1016