JMB

Unusual Molecular Architecture of the *Yersinia pestis* Cytotoxin YopM: A Leucine-rich Repeat Protein with the Shortest Repeating Unit

Artem G. Evdokimov*, D. Eric Anderson, Karen M. Routzahn and David S. Waugh*

Protein Engineering Section Macromolecular Crystallography Laboratory National Cancer Institute at Frederick, P.O. Box B Frederick, MD 21702-1201, USA

Many Gram-negative bacterial pathogens employ a contact-dependent (type III) secretion system to deliver effector proteins into the cytosol of animal or plant cells. Collectively, these effectors enable the bacteria to evade the immune response of the infected organism by modulating host-cell functions. YopM, a member of the leucine-rich repeat protein superfamily, is an effector produced by the bubonic plague bacterium, *Yersinia pestis*, that is essential for virulence. Here, we report crystal structures of YopM at 2.4 and 2.1 Å resolution. Among all leucine-rich repeat family members whose atomic coordinates have been reported, the repeating unit of YopM has the least canonical secondary structure. In both crystals, four YopM monomers form a hollow cylinder with an inner diameter of 35 Å. The domain that targets YopM for translocation into eukaryotic cells adopts a well-ordered, α-helical conformation that packs tightly against the proximal leucine-rich repeat module. A similar α-helical domain can be identified in virulence-associated leucine-rich repeat proteins produced by *Salmonella typhimurium* and *Shigella flexneri*, and in the conceptual translation products of several open reading frames in *Y. pestis*.

Keywords: Yersinia pestis; YopM; leucine-rich repeat; contact-dependent secretion; type III secretion

*Corresponding authors

Introduction

Although invasive bacterial pathogens are a taxonomically diverse group of microbes, it is becoming clear that even distantly related organisms often employ similar virulence strategies.[1] One common theme, exemplified by pathogenic species of *Yersinia, Salmonella, Shigella* and *Xanthomonas*, among others, is the use of a contact-dependent (type III) secretion system to inject proteins directly into the cytosol of eukaryotic cells.[2–4] These proteins, termed effectors, enable the bacterium to evade the non-specific immune response of the infected organism by interfering with host-cell functions. Whereas the number and nature of the effectors delivered by each pathogen is quite variable, many structural components of the type III secretion apparatus (approximately 20 polypeptides in all), are conserved even among distantly related pathogens, and some are clearly related to proteins involved in the assembly of the bacterial flagellum, suggesting a common origin for these two systems.[2,5]

Yersinia pestis, the causative agent of plague, injects at least six effectors into the cytosol of mammalian cells: YopE, YopH, YopJ, YopM, YopT and YpkA. Most of these proteins are essential virulence factors.[6] YopH and YpkA contain functional eukaryotic-like protein tyrosine phosphatase and serine-threonine kinase domains, respectively.

YopH dephosphorylates focal adhesion kinase, paxillin, Fyb, p130[Cas], SKAP-HOM, and possibly other targets as well, thereby antagonizing bacterial phagocytosis and causing the disruption of focal adhesions at sites of cell attachment to the extracellular matrix.[7–9] The kinase activity of YpkA is stimulated by actin,[10] but the eukaryotic protein(s) it phosphorylates is still unknown. YpkA has been observed to associate with Rho-family GTPases.[11,12] However, YpkA does not phosphorylate these GTPases, and so the significance of this interaction is unclear. YopJ is a ubiquitin-like protein protease that inhibits mitogen-activated protein kinase (MAPK) and nuclear factor κb (NFκB) signaling in animal cells, thereby triggering apoptosis, and resulting in general suppression of the inflammatory response.[13] YopE and YopT elicit cytotoxic effects by modulating the activity of Rho family GTPases, which are pivotal regulators of the actin cytoskeleton.[14–16] The *Yersinia* effector about which the least is known is YopM. This highly acidic protein ($pK \sim 4.2$) is essential for virulence,[17] but its specific targets and anti-host functions have yet to be elucidated. YopM has been observed to accumulate in the cytosol, and in contrast to the other *Yersinia* effectors, also in the nucleus of mammalian cells.[18]

Y. pestis YopM has a modular architecture dominated by 15 tandem copies of a 20/22 residue leucine-rich repeat (LRR) motif. The repeating unit in YopM (Figure 1) is the shortest among all LRRs known to date.[19] The LRRs in *Y. pestis* YopM comprise the central 310 residues in the polypeptide sequence, and are bracketed by 73 and 24 residues on the N and C termini, respectively. The function of the C-terminal tail of YopM is unknown. However, as is thought to be true of all the *Yersinia* effectors, information contained in the N-terminal region of YopM targets the protein for secretion from the bacterium and translocation into mammalian cells.[20]

To investigate the molecular architecture of an LRR domain that is comprised of the smallest repeating unit, and to gain some insight into the function of *Y. pestis* YopM, we have crystallized the protein and determined its structure in two different crystal forms at 2.4 and 2.1 Å resolution.

Results and Discussion

Three crystal forms of YopM

Recombinant *Y. pestis* YopM with a C-terminal hexahistidine tag was overproduced in *Escherichia coli*, purified, and crystallized as described.[21] The first structure (space group $P4_222$), with one molecule in the asymmetric unit (AU), was solved by isomorphous replacement using platinum and mercury derivatives. Two other structures (space groups $I4_122$ and $C222_1$, with one and two molecules per AU, respectively) were solved by molecular replacement using the $P4_222$ structure as a model. Only the $P4_222$ and $I4_122$ structures are
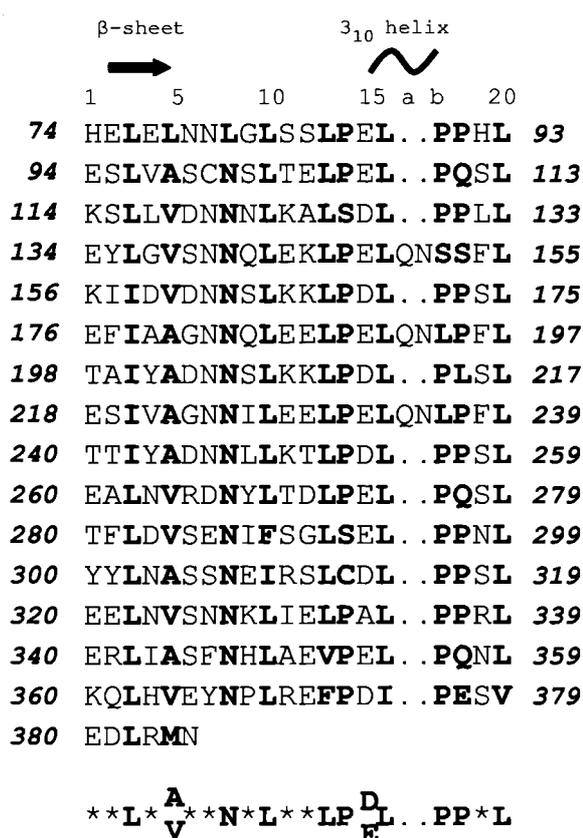

Figure 1. Sequence alignment of the individual repeats that constitute the LRR domain of YopM. Structurally important residues are shown in bold. The two extra residues that occur in LRRs 4, 6 and 8 are designated a and b. The positions of β-strands and 3_{10} helices are indicated above the alignment, and the consensus LRR sequence in YopM is shown below. Residue numbering according to YopM sequence is shown in bold italics within each row.

discussed here because the data for the $C222_1$ structure is of low quality, and the protein model is still undergoing refinement. Representative electron density from the initial experimental map and from the two fully refined structures is shown in Figure 2. The fold of the YopM monomer is illustrated in Figure 3.

The secondary structure of YopM

A distinguishing feature of YopM is its paucity of canonical secondary structure. As is typically the case in other LRR proteins, only three residues in each strand on the concave face of YopM engage in the characteristic backbone-backbone hydrogen bonding interactions that are the hallmark of β-sheet structure. However, whereas the strands that comprise the convex surface of LRR domains in other proteins almost invariably are helical (3_{10} helices in internalin B,[22] U2A′[23] and TAP,[24] α-helices in Rab geranylgeranyltransferase,[25]
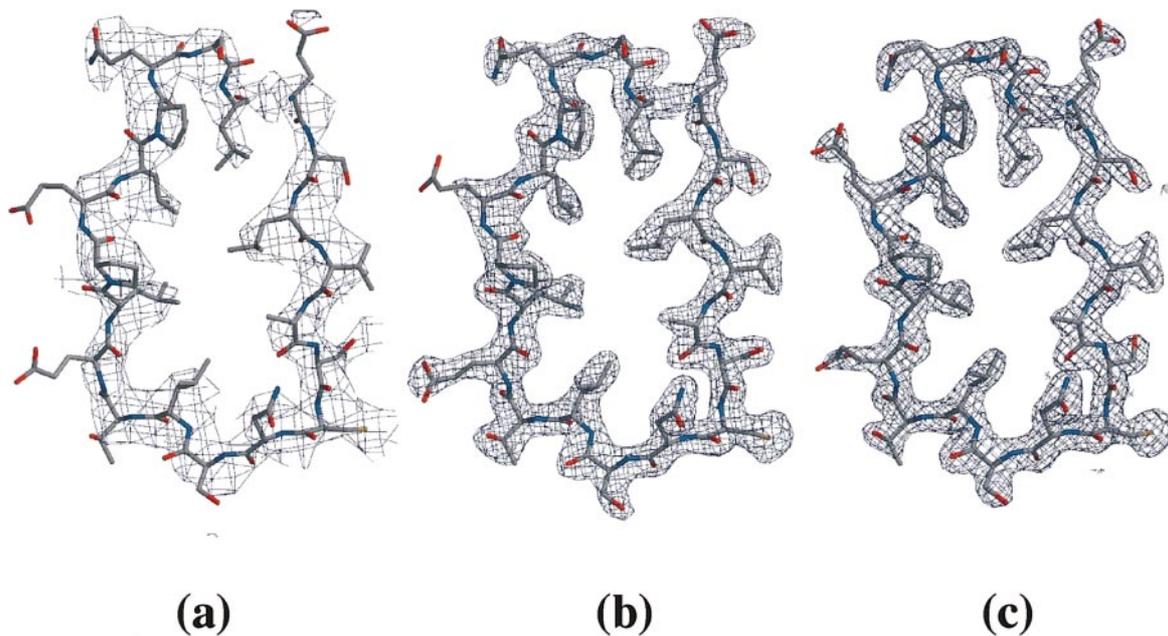
Figure 2. Representative electron density maps superimposed with the final models. (a) Experimental map contoured at the 1.2 σ level. (b) $|2F_o - F_c|$ simulated-annealing total-omit map contoured around the same portion of the protein in the $P4_222$ structure. (c) $|2F_o - F_c|$ simulated-annealing total-omit map contoured around the same portion of the protein in the $I4_122$ structure.

rna1p[26] and ribonuclease inhibitor[27]), in YopM this part of the structure adopts an extended conformation instead. This latest architectural twist underscores the versatility of the LRR motif as a building block for protein structures, and serves as a remarkable example of how even large proteins can adopt stable conformations in the absence of extensive α-helical and/or β-sheet structure. The region of the protein with the greatest density of canonical secondary structure consists of a tandem pair of α-helices, located just proximal to the first

LRR module (Figure 3). This part of the structure purportedly corresponds to the signal that targets YopM for translocation into eukaryotic cells.[20]

A theoretical model of the YopM LRR domain structure has been proposed by Kajava and co-workers.[28,29] This model is generally similar to the crystal structure of YopM, especially in the β-sheet portion of the LRR domain. However, it did not correctly predict the conformation of the polypeptide chain in the proline-rich region of each repeat, nor did it anticipate the overall helical twist of the protein.

The LRRs in YopM

YopM contains two slightly different types of LRR (Figure 4(a)). The longer variation (22 residues) includes a short helical segment at the apex of the loop; otherwise, the backbone conformations of these two repeat types are nearly identical. There are only three repeats of the 22-residue type in YopM, the remainder being 20 residues in length, and completely devoid of helical content. An individual LRR in YopM may be defined as 20 consecutive residues (or 22 residues in three cases), beginning at an arbitrary position. However, it is convenient to use the beginning of the first LRR in YopM (residue 74) as the starting point for our definition of the repeating unit. Least-squares superimposition of the two types of LRRs (Figure 4(a)) makes it clear that the differences between the 20 and 22-residue repeats are restricted to the upper-left corner of the LRR, so the two types may be analyzed in a uniform
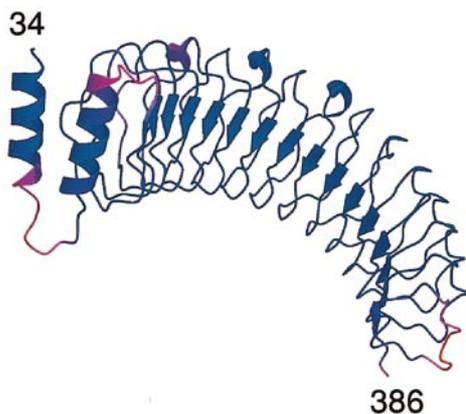

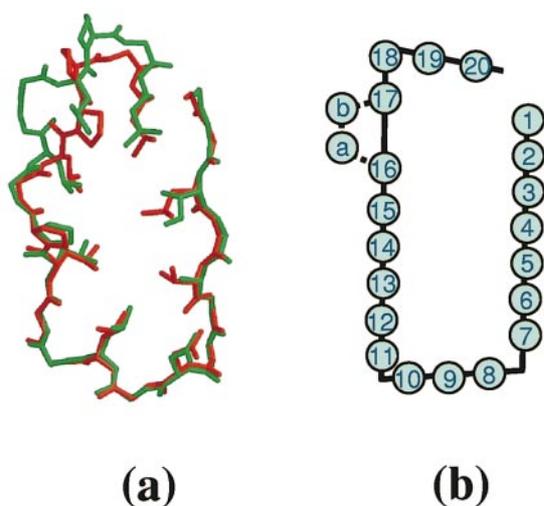**Figure 3.** Schematic representation of the YopM structure colored according to atomic displacement factors. Blue corresponds to low values, red to high values. The N and C-terminal residues that were observed in the crystal structure are labeled.

Figure 4. The two types of LRR in YopM. (a) Super-imposition of the 20-residue (LRR13, red) and the 22-residue (LRR8, green) repeats found in the structure of YopM. (b) Residue numbering scheme within the LRR.

fashion with the exception of the residues involved in the formation of the helix in the longer repeats. Thus, it is possible to represent both types of LRR in YopM as a four-sided box: a simplified two-dimensional projection of an open three-dimensional polygon (Figure 4(b)). Individual repeats are referred to by their order within the LRR domain; the N and C-terminal LRRs are designated LRR1 and LRR15, respectively.

Although the structures of the *P*4 and *I*4 crystal forms of YopM are not identical, individual repeats from the *P*4 structure superimpose with their analogues in the *I*4 structure with an average C^α rmsd of only 0.35 Å. Analysis of the least-squares alignment of all 15 LRRs in both crystal structures of YopM onto one another revealed that LRR13, having the least average C^α rmsd from all the other repeats, can be used as a standard of reference with which all the other LRRs can be compared, in both the *P*4 and the *I*4 structures. A plot of the C^α rmsd per residue demonstrates that apart from the first LRR, the last LRR, and the geometrically different parts of the three 22-residue LRRs, there are only a few residues whose C^α atoms deviate from the corresponding positions in LRR13 by more than 1 Å (Figure 5(a)). All of these differences are due to the substitutions of conserved proline residues with serine or cysteine residues.

Structurally important residues and inter-repeat interactions

Certain positions of the LRRs in YopM are occupied by highly conserved amino acids, whereas other positions can accommodate a wide variety of residues. Positions 3, 5, 10, 13, 16 and 20 contain the structural leucine, isoleucine, valine, and ala-
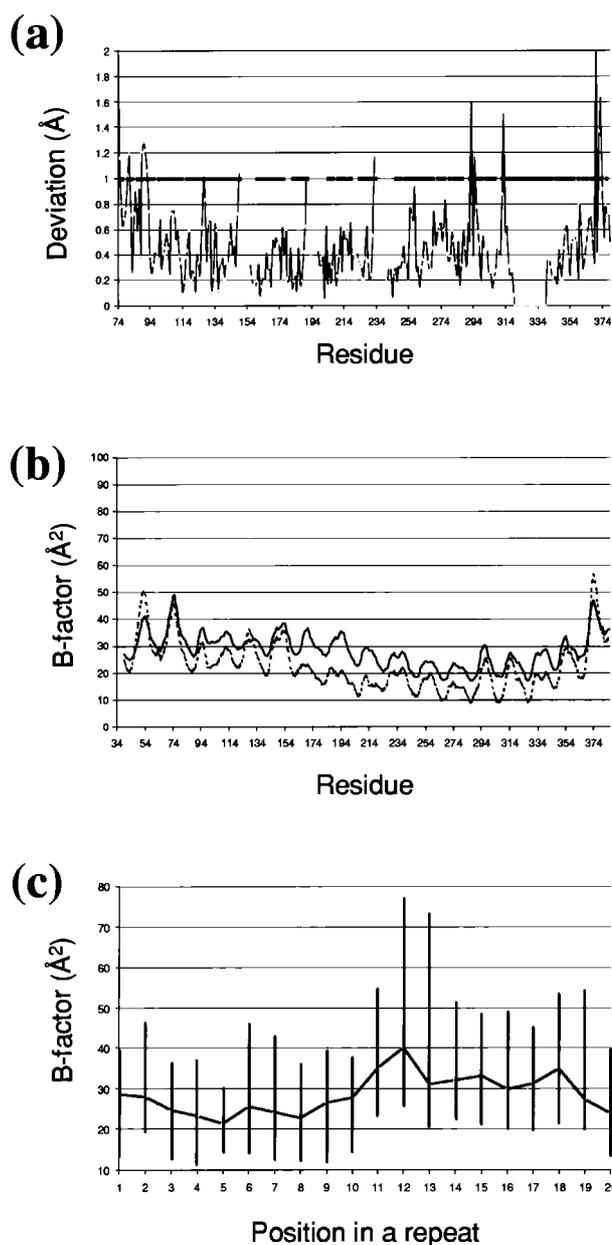
Figure 5. Geometry of the LRR domain in YopM (a) Deviation of individual C^α atoms in the LRR domain of YopM from the corresponding C^α atom positions of LRR13. Breaks in the line correspond to the helical parts of the 22-residue repeats. The thick line in the middle of the plot indicates residues that are identical with the corresponding residues in LRR13; breaks in the line represent sequence differences. (b) Running average (five-residue window) plot of the C^α atomic displacement parameters for the two YopM structures. The *B*-factors for the $P4_2$22 and $I4_1$22 structures are plotted as continuous and broken lines, respectively. (c) The average C^α displacement parameters are plotted as a function of the residue position within the repeat. The maximum and minimum values for each position correspond to the ends of the vertical bars.

nine residues. The van der Waals interactions involving these residues appear to be the main factor that stabilizes the overall fold of the LRR domain. Position 8 contains an invariant (except for the N-terminal repeat) asparagine residue, the side-chain of which is oriented towards the core of the LRR. This side-chain forms strong hydrogen bonds with the main-chain carbonyl group of the residue in position 5 of the same repeat, as well as with the main-chain carbonyl group in position 5 and main-chain amide group in position 8 of the previous repeat. Cooperatively, these asparagine residues form a ''ladder'' that serves as the spine of the LRR domain in YopM. Similar asparagine residue ladders are found in other LRR proteins as well (e.g. in the structure of U2A').[23] Position 19 contains a loosely conserved serine residue that makes a hydrogen bond with the main-chain carbonyl group of position 18 in the previous repeat. This interaction generates another cooperative stabilization pattern across the LRR domain of the protein. Positions 14, 17 and 18 contain highly conserved proline residues that are most likely important for the stabilization of the extended conformation of the polypeptide chain in this area. It can be readily observed that the C^α atomic displacement parameters (*B*-factors, ADPs) oscillate with a periodicity of 20 residues (Figure 5(b)), indicating that the distribution of *B*-factors within each repeat is non-random. Analysis of this distribution shows that the residues in positions 11-18 have systematically higher ADPs (Figure 5(c)). The conserved proline residues may serve to reinforce the most flexible region of the LRR backbone, where direct inter-repeat contacts do not occur. It is not clear why acidic residues are strongly preferred in position 15; their side-chains are not organized into any kind of regular array that would suggest they contribute to the stability of the fold.

Comparison of the YopM LRR with the repeating units of other LRR proteins

LRRs vary in their length and pattern of conserved residues.[30] In Figure 6 and Table 1, the YopM LRR is compared and contrasted with the repeats in internalin B,[22] Rab
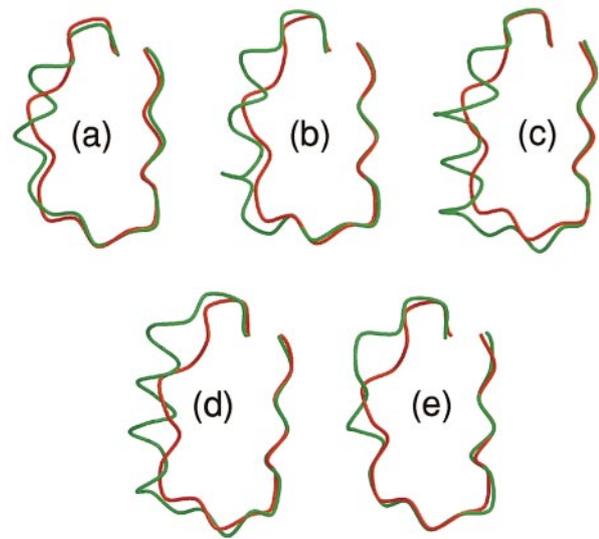

Figure 6. Least-squares superimposition of the 20-residue LRR from YopM (LRR13, red) with representative LRRs (green) from: (a) internalin B; (b) Rab geranylgeranyltransferase; (c) rna1p; (d) RNase inhibitor; and (e) U2A' spliceosomal protein.

geranylgeranyltransferase,[25] rna1p,[26] porcine RNase inhibitor,[27] and the U2A' spliceosomal protein.[23] For most of the LRR types, the right-hand side β-sheet portions, the lower-left corner turn, and the top of the backbone ''box'' overlap almost exactly with the corresponding parts of the YopM LRR. The left sides are quite variable and do not align well. The 3_{10} helix of the internalin B LRR occupies almost exactly the same space as the proline-rich, extended segment of the YopM LRR. This bacterial LRR is most similar to the repeat found in YopM.

The geometry of the YopM monomer

We used the backbone atoms of every repeat in YopM to construct 15 least-squares planes. Together with the centers of gravity of each LRR, these planes can be used to describe the geometry of the LRR domain in terms of its repetitive struc-

Table 1. Sequence alignment of the 20-residue LRR from YopM with other LRRs

Protein	Aligned sequences	rmsd (Å)
Internalin B	**ESLYLGNNKITDI**TVLSR**LTKL**	0.36
YopM	**EELNVSNNKLIEL**P..AL**PPRL**	
Rab geranylgeranyltransferase	L**QELLLCNNRL**QQSAAIQPLVSCP**RL**	0.45
YopM	.**EELNVSNNKL**IELPALP.....P**RL**	
rna1p	**HTVKMVQN**GIRPEGIEHLLLEGLAYC**QQL**	0.57
YopM	**EELNVSNN**....KLIELPALP.....**PRL**	
RNase inhibitor	**EKLQLEYC**RLTAASCEPLASVLRAT**RAL**	0.71
YopM	**EELNVSNN**.....KLIELPA.LP..**PRL**	
U2A' spliceosomal protein	**DAIDFSDNE**IRKLDGFPLL**RRL**	0.80
YopM	**EELNVSNNK**LIELPALP..**PRL**	

Residues with similar backbone geometries were used for alignment and are shown in bold type.

ture. For the *P*4 and *I*4 structures, the absolute average vertical/horizontal displacements of each LRR from its predecessor are 1.2/4.9 and 1.3/5.0 Å, respectively. The angle between two consecutive LRR least-squares planes is ~11.5° in the *P*4 structure and ~11.9° in the *I*4 structure.

Overall, the C^α positions in the YopM monomers from the *P*4 and *I*4 structures superimpose with an average rmsd of 1.4 Å. When the superposition is performed only on the bi-helical module of the two molecules (residues 34-74), the backbones diverge progressively toward their C termini (Figure 7(a) and (b)). This implies that the conformation of the YopM monomer can change depending on the environmental conditions. The *P*4 form of YopM is less curved and more extended than the *I*4 form. Analysis of the backbone torsion angles of the two conformations of YopM shows that there are small (~10° on average) systematic differences between the two structures in most of the ϕ and ψ angles, as well as a few significant differences in areas where the backbone radically changes conformation (Figure 7(c)). Interestingly, the areas of greatest conformational variability correspond to parts of the sequence where structurally conserved proline residues are substituted with serine or cysteine residues (e.g. Ser128 and Ser293), reinforcing the hypothesis that these conserved proline residues stabilize the otherwise flexible backbone sections that comprise the convex face of the protein. The crystallization conditions for both forms differ only by a few percent in ethylene glycol concentration, and so it is not clear what drives the protein into a particular conformation.

Water-mediated interactions

Tightly bound solvent molecules form a quasi-regular network that bridges the backbone on the convex side of the protein (Figure 8). A similar water network was observed in the high-resolution structure of internalin B.[22] Comparison of the solvent networks in the *P*4 and *I*4 structures reveals an overall pattern of conservation, but there are also differences, especially in regions of the LRR domain where the backbone adopts slightly different conformations in the two structures. The water network appears to compensate for the lack of direct inter-repeat interactions on the convex side of YopM, and may also help the protein to achieve a compromise between flexibility and structural stability; the inherent ''softness'' of the backbone-water-backbone solvent bridges enables the backbone to experience considerable variation in geometry without imposing a significant energy penalty.

Orthologs of YopM in *Yersinia enterocolitica*

Two YopM sequences from different strains of *Y. enterocolitica* have been published.[20,31] Knowledge of the formal geometry of the *Y. pestis* YopM ($YopM_p$) monomer allowed us to construct models
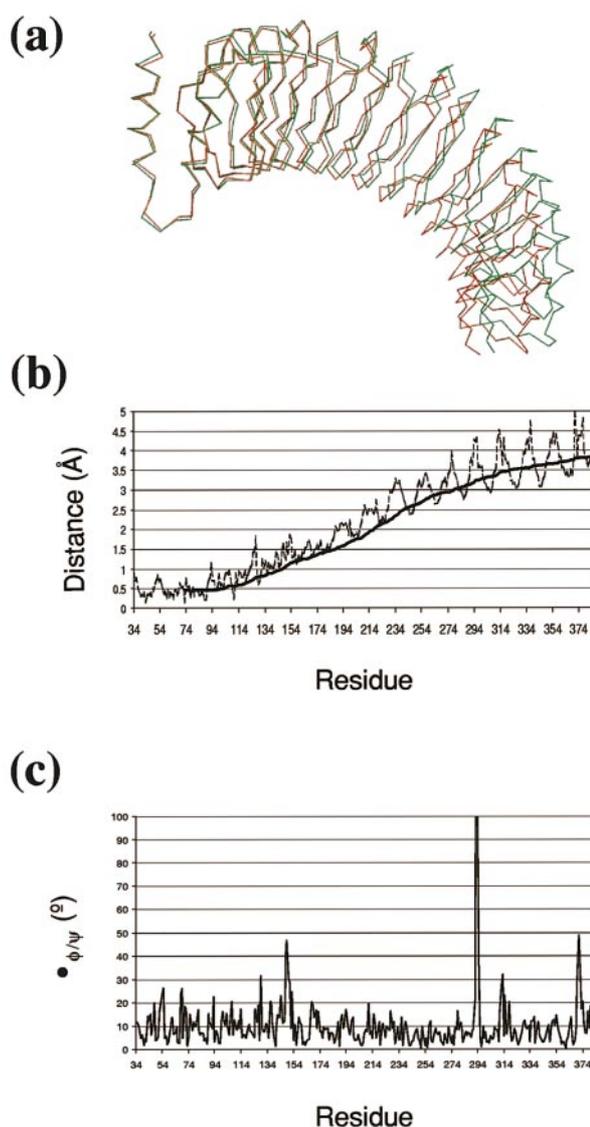
Figure 7. Differences between the $P4_2 22$ and $I4_1 22$ structures of YopM. (a) Superimposition of the $P4_2 22$ (red) and $I4_1 22$ (green) C^α-traces performed by least-squares alignment of residues 34-74. (b) Distance between corresponding C^α atoms of the $P4_2 22$ and $I4_1 22$ crystal structures, aligned as above. The broken line indicates the actual values, and the continuous line represents a running average with a 20-residue window. (c) Differences in backbone geometry, plotted as the distance between corresponding C^α atoms of the two structures in the ϕ-ψ space $\Delta_{\phi/\psi} = ((\phi - \phi')^2 + (\psi - \psi')^2)^{1/2}$ where ϕ and ψ are the values of backbone torsion angles from the $P4_2 22$ structure and ϕ' and ψ' are the values from the $I4_1 22$ structure.

of these YopM proteins, which contain 20 and 13 LRRs, and are termed $YopM_{e20}$ and $YopM_{e13}$, respectively. In addition to the two types of LRR that are present in $YopM_p$, $YopM_{e20}$ includes one 21-residue repeat that is very similar to the 20-residue LRR except that it contains an extra serine residue after position 6. The non-repetitive N and
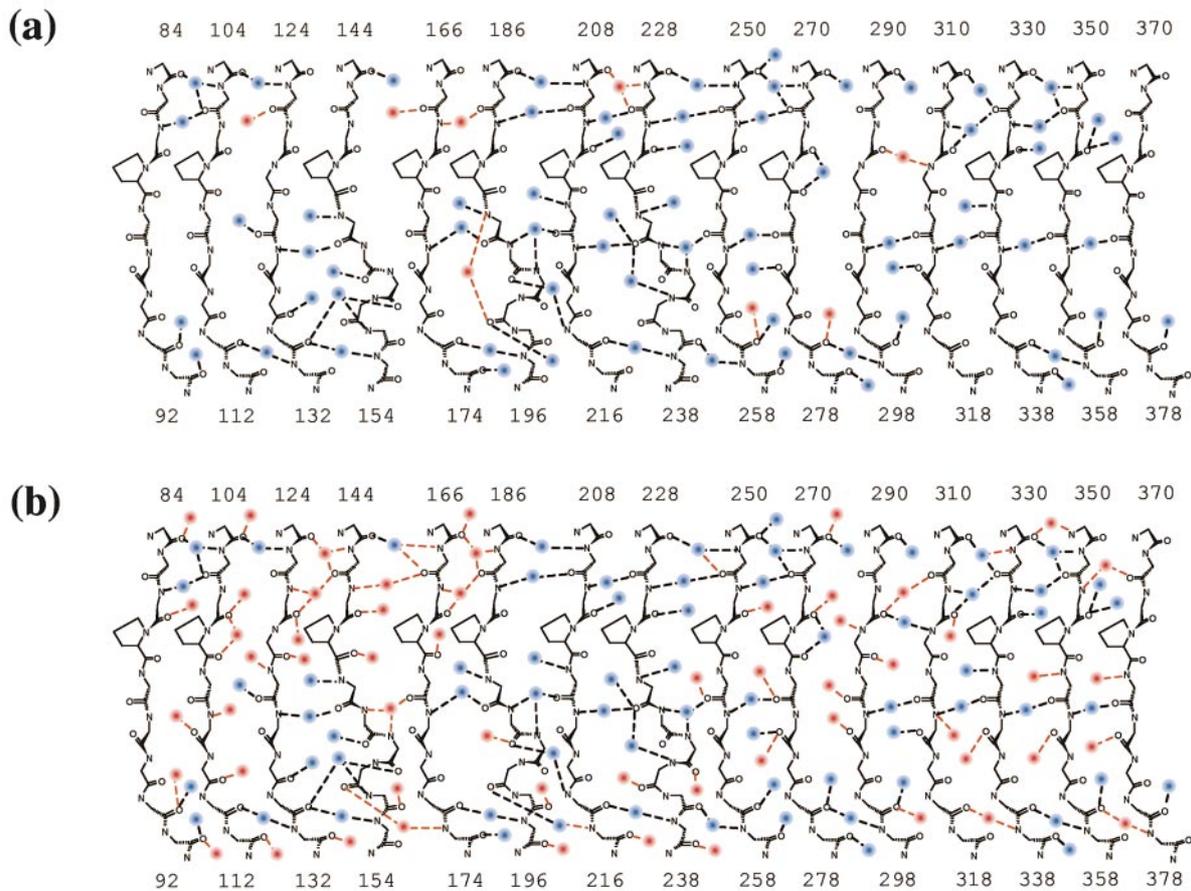
Figure 8. Water network at the convex surface of YopM in the (a) $P4_222$ and (b) $I4_122$ structures. Only those water molecules that participate in hydrogen bonds with the protein backbone are shown. Only the proline residues in position 14 are shown with side-chains. The water molecules that are conserved in both structures are represented by blue spheres, whereas the rest are shown as red spheres. Conserved and variable hydrogen bonds are represented by broken black and red lines, respectively.

C-terminal portions of the two *Y. enterocolitica* YopM proteins are virtually identical with the corresponding parts of $YopM_p$. Comparison of the LRR domain sequences from all three YopM proteins (Figure 9) makes it clear that the first five and the last four repeats can be aligned with confidence. The major differences are confined to the central part of the LRR domains, where duplications and/or deletions of LRR modules appear to have occurred. Only one 22-residue repeat is present in $YopM_{e20}$, which seems to lack LRRs 6-8 of $YopM_p$. The larger size of the LRR domain in $YopM_{e20}$ can be ascribed to the insertion of additional repeats that resemble LRRs 12-14 in $YopM_p$. $YopM_{e13}$ is very similar to $YopM_p$, except that the third 22-residue repeat and the LRR immediately following it are missing.

A model of the protein backbone of $YopM_{e20}$ was constructed by fusing together blocks of individual repeats, according to the rules of LRR organization derived from the structure of $YopM_p$. This backbone then was used as a scaffold onto which the sequence of the protein was overlaid either manually or automatically (e.g. using

SwissProt threading server[32]). A model of the $YopM_{e13}$ structure was constructed by deleting LRRs 8 and 9 from the structure of $YopM_p$, fusing the remaining portions in order to remove the gap, and then altering the amino acid sequence.

Surface features of the YopM monomer

A schematic diagram of the *Y. pestis* YopM LRR domain surface, colored in accordance with charge and hydrophobicity of the amino acid side-chains, is presented in Figure 10. It is readily apparent that the concave side of the protein exposes a variety of amino acid side-chains, whereas the convex side exposes a semi-regular array of mostly proline and acidic residues. The solvent-exposed proline residues create a thin hydrophobic strip on the otherwise negatively charged surface. It is clear from crystallographic studies that both U2A' and ribonuclease inhibitor utilize the concave surfaces of their LRR domains to bind their target proteins, and it has been inferred on the basis of sequence conservation and/or site-directed mutagenesis experiments that the same is true of internalin B,
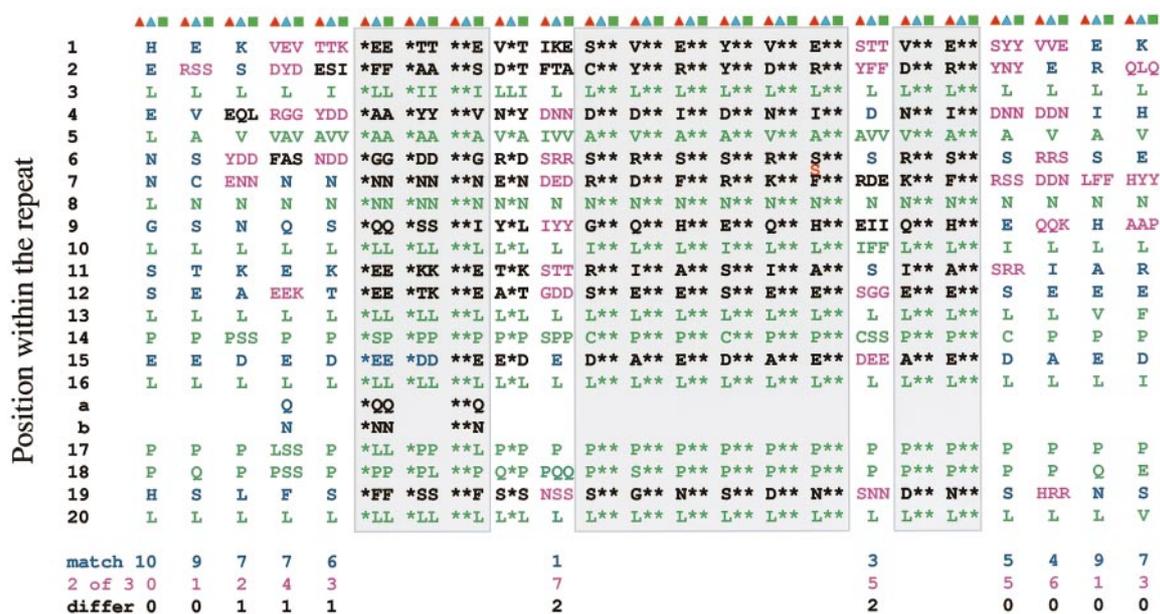
Position within the repeat

	▲▲■	▲▲■	■	▲▲■	▲▲■	▲▲	▲▲■	■	▲▲	▲▲■	▲▲■	▲▲■	▲▲■	▲▲■	▲▲	▲▲■	▲▲	▲▲■	▲▲	■	▲▲■
1	H	E	K	VEV	TTK	*EE	*TT	**E	V*T	IKE	S**	V**	E**	Y**	V**	E**	STT	V**	E**	SYY	VVE	E	K
2	E	RSS	S	DYD	ESI	*FF	*AA	**S	D*T	FTA	C**	Y**	R**	Y**	D**	R**	YFF	D**	R**	YNY	E	R	QLQ
3	L	L	L	L	I	*LL	*II	**I	LLI	L	L**	L**	L**	L**	L**	L**	L	L**	L**	L	L	L	L
4	E	V	EQL	RGG	YDD	*AA	*YY	**V	N*Y	DNN	D**	D**	I**	D**	N**	I**	D	N**	I**	DNN	DDN	I	H
5	L	A	V	VAV	AVV	*AA	*AA	**A	V*A	IVV	A**	V**	A**	A**	V**	A**	AVV	V**	A**	A	V	A	V
6	N	S	YDD	FAS	NDD	*GG	*DD	**G	R*D	SRR	S**	R**	S**	S**	R**	S**	S	R**	S**	S	RRS	S	E
7	N	C	ENN	N	N	*NN	*NN	**N	E*N	DED	R**	D**	F**	R**	K**	F**	RDE	K**	F**	RSS	DDN	LFF	HYY
8	L	N	N	N	N	*NN	*NN	**N	N*N	N	N**	N**	N**	N**	N**	N**	N	N**	N**	N	N	N	N
9	G	S	N	Q	S	*QQ	*SS	**I	Y*L	IYY	G**	Q**	H**	E**	Q**	H**	EII	Q**	H**	E	QQK	H	AAP
10	L	L	L	L	L	*LL	*LL	**L	L*L	L	I**	L**	L**	I**	L**	L**	IFF	L**	L**	I	L	L	L
11	S	T	K	E	K	*EE	*KK	**E	T*K	STT	R**	I**	A**	S**	I**	A**	S	I**	A**	SRR	I	A	R
12	S	E	A	EEK	T	*EE	*TK	**E	A*T	GDD	S**	E**	E**	S**	E**	E**	SGG	E**	E**	S	E	E	E
13	L	L	L	L	L	*LL	*LL	**L	L*L	L	L**	L**	L**	L**	L**	L**	L	L**	L**	L	L	V	F
14	P	P	PSS	P	P	*SP	*PP	**P	P*P	SPP	C**	P**	P**	C**	P**	P**	CSS	P**	P**	C	P	P	P
15	E	E	D	E	D	*EE	*DD	**E	E*D	E	D**	A**	E**	D**	A**	E**	DEE	A**	E**	D	A	E	D
16	L	L	L	L	L	*LL	*LL	**L	L*L	L	L**	L**	L**	L**	L**	L**	L	L**	L**	L	L	L	I
a				Q		*QQ		**Q															
b				N		*NN		**N															
17	P	P	P	LSS	P	*LL	*PP	**L	P*P	P	P**	P**	P**	P**	P**	P**	P	P**	P**	P	P	P	P
18	P	Q	P	PSS	P	*PP	*PL	**P	Q*P	PQQ	P**	S**	P**	P**	P**	P**	P	P**	P**	P	P	Q	E
19	H	S	L	F	S	*FF	*SS	**F	S*S	NSS	S**	G**	N**	S**	D**	N**	SNN	D**	N**	S	HRR	N	S
20	L	L	L	L	L	*LL	*LL	**L	L*L	L	L**	L**	L**	L**	L**	L**	L	L**	L**	L	L	L	V

match	10	9	7	7	6					1							3			5	4	9	7
2 of 3	0	1	2	4	3					7							5			5	6	1	3
differ	0	0	1	1	1					2							2			0	0	0	0

Figure 9. Repeat-based sequence alignment of YopM$_{e20}$ (columns marked by red triangles), YopM$_{e13}$ (blue triangles), and YopM$_p$ (green squares). Structurally important residues are colored green. Residues conserved in all three proteins are shown only once and are colored blue, while residues conserved in any two of the three proteins are colored cyan. A single 21-residue repeat in the sequence of YopM$_{e20}$ is shown with the extra serine residue in red. The areas where sequence insertions or deletions make alignment of all three proteins impossible are shaded.

rna1p, and TAP. It therefore seems likely that the concave surface of YopM will also be involved in target recognition.

We reasoned that additional insight into the potential binding site(s) on YopM might be gained by comparing the surfaces of YopMs from different species, as the functional surfaces are expected to be conserved. The models of YopM$_{e13}$ and YopM$_{e20}$ were used to derive schematic representations of their respective surfaces. Figure 10 illustrates the distribution of residue types on the surfaces of all three YopM proteins. Like YopM$_p$, the concave surfaces of the model monomers exhibit more variation in residue types than their convex surfaces. A two-dimensional alignment of the concave surfaces revealed a conserved patch of residues in all three proteins (Figure 10). In YopM$_p$ and YopM$_{e13}$, the center of this patch is located six repeats from the C termini of the proteins, whereas in YopM$_{e20}$ it is situated six repeats from the N terminus. The significance of these conserved residues could be evaluated by site-directed mutagenesis.

The YopM tetramer

A particularly striking aspect of the YopM structure is its quaternary architecture. In all three crystal forms, protein tetramers are found in the crystal lattice (Figure 11(a)). The tetramer is a superhelix that describes a hollow cylinder with an inner diameter of ~35 Å. Each strand of the helix is created by dimerization between the C-terminal LRR modules of two YopM monomers. This interaction buries a considerable amount of exposed hydro-

phobic surface on the distal LRR modules of the monomers (Figure 11(b)). The alignment of LRR modules between monomers gives rise to a nearly continuous coil, almost as if there were no interruption between repeats.

Two end-to-end dimers wrap around each other in a helical fashion to form a symmetrical tetramer. A considerable amount of surface area is occluded by this interaction as well (Figure 11(b)). In fact, most of the exposed hydrophobic surface on the YopM monomer is involved in protein-protein contacts in the tetramer. Two symmetry-related pairs of calcium ions are intimately associated with the interface between dimers of YopM (Figure 11(c)). They are held in place by interactions with Asn246, Asp266, Asn307, Glu308 and Asn326. Calcium site 1 coordinates six ligands into an almost perfect bipyramid, whereas site 2 coordinates seven ligands into a ''capped'' bipyramid. The geometric parameters of the calcium coordination spheres are presented in Table 2. Other residues involved in tetramerization include the tyrosine and isoleucine residues in position 9 of LRR9 and LRR10 that interact with their symmetry equivalents from the neighboring molecule to form a tight hydrophobic cluster.

The biological significance of the YopM tetramer is unclear at the present time. Although it occurs in all of the crystal forms of the protein and seems artfully designed, we cannot rule out the possibility that the tetramer is an artifact resulting from the high concentration of calcium in crystallization solutions. Experiments designed to demonstrate oligomerization under more physiological
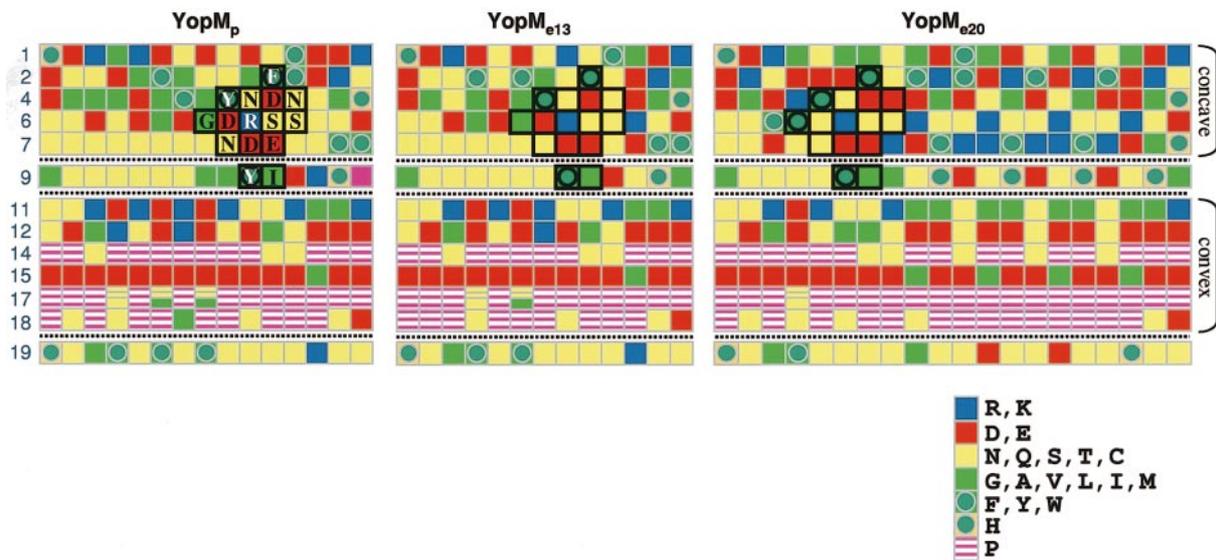
Figure 10. Schematic surface projections of YopM$_p$ (left) YopM$_{e13}$ (center) and YopM$_{e20}$ (right) colored in accordance with the chemical properties of the amino acid side-chains. The broken lines represent approximate ''fold lines'' that define the boundaries between the convex and concave surfaces. Buried structural residues have been omitted for clarity. A patch of residues conserved in all three structures is highlighted. In YopM$_p$ this patch corresponds to: Gly223 (residue 6 in LRR8); Tyr243, Asp245 and Asn246 (residues 4, 6 and 7 in LRR9); Asn263, Arg265, Asp266 and Tyr268 (residues 4, 6, 7 and 9 in LRR10); Phe281, Asp283, Ser285, Glu286 and Ile288 (residues 2, 4, 6, 7 and 9 in LRR11); and Asn303 and Ser305 (residues 4 and 6 in LRR12).

conditions produced contradictory results: size-exclusion chromatography coupled with refractive index/light-scattering measurements indicated that the protein is monomeric in both the absence and presence of 0.8 mM calcium ions, whereas glutaraldehyde cross-linking under similar conditions suggested that the protein undergoes oligomerization upon the addition of calcium (data not shown). It is possible that tetramers of *Y. pestis* YopM form concomitantly with the binding of its target(s) in eukaryotic cells but are not energetically favorable under physiological conditions in the absence of the specific cellular target(s). In this regard, it would be interesting to examine the

effect on virulence of mutants that are defective in tetramerization.

If the tetramer is biologically relevant, then one might anticipate that the potential to form oligomers would also exist in the two *Y. enterocolitica* YopMs. We attempted to assemble tetramers of YopM$_{e21}$ and YopM$_{e13}$ by aligning the last four LRRs in the corresponding monomers with the last four LRRs in the tetrameric YopM$_p$ structure. The model of the YopM$_{e13}$ tetramer constructed in this fashion is very similar to the tetramer of YopM$_p$. It can be speculated that the calcium-binding sites are formed in essentially the same fashion, although the equivalent of the key bidentate

Table 2. Calcium coordination in YopM structures

Calcium	Ligand	Distance (Å)	
		P4$_2$22	I4$_2$22
Ca$_1$	Asp266	2.21	2.05
	Asn246	2.25	2.05
	Asn326	2.36	2.21
	O$_1$	2.30	2.05
	O$_2$	2.28	2.24
	O$_3$	2.16	2.33
Ca$_2$	Asp266	2.49	2.25
	Asn307	2.28	2.30
	Glu308	2.34	2.25
	O$_4$	2.51	2.27
	O$_5$	2.53	2.38
	O$_6$	2.58	2.53
	O$_7$	2.60	2.78
Ca$_1$-Ca$_2$		5.94	5.59

Ligand numbering corresponds to Figure 11. Italicized residues belong to the symmetry-related protein monomer.
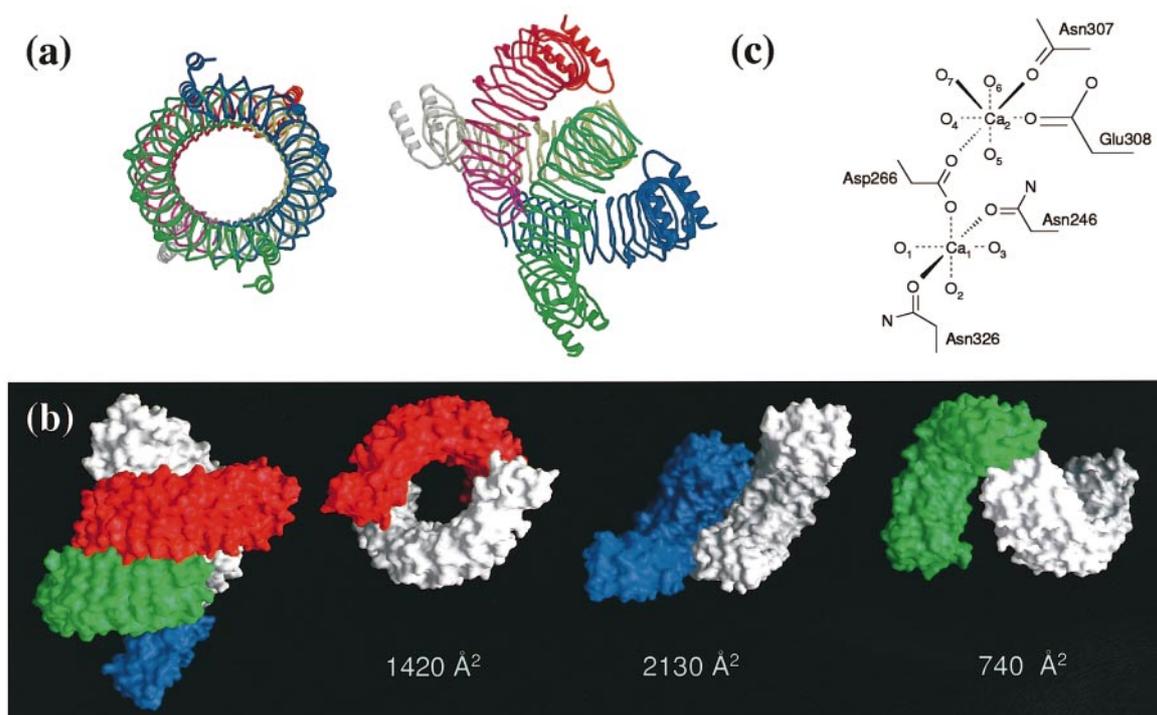
Figure 11. The *Y. pestis* YopM tetramer. (a) Two views of the YopM tetramer C^{α}-trace colored by monomer. (b) Molecular surface representation of the YopM tetramer and of the pairwise interactions of its constituent monomers. (c) Schematic representation of the calcium-binding sites.

calcium-binding aspartic acid 266 in YopM$_\mathrm{p}$ would be a glutamic acid residue in YopM$_\mathrm{e13}$. Conversely, many residues from the N-terminal region of each YopM$_\mathrm{e20}$ monomer come unacceptably close or even clash with the other monomers in the model of the YopM$_\mathrm{e20}$ tetramer. Consequently, analogous tetramers of YopM$_\mathrm{e20}$ could not form unless the geometry of the individual monomers differed significantly from what is observed in the YopM$_\mathrm{p}$ structure. We know that the YopM$_\mathrm{p}$ monomer is fairly flexible, because it adopts slightly different conformations in the *P*4 and *I*4 structures, but the limits of its flexibility are not clear. It is even remotely possible that the biologically relevant form of YopM is almost flat. Yet, the large hydrophobic patch on the surface of LRR15 that mediates the dimerization of YopM$_\mathrm{p}$ monomers to form one strand of the tetrameric superhelix is conserved in all three proteins, and so it seems likely that both *Y. enterocolitica* proteins could also form tail-to-tail dimers. We hope to be able to shed some light on the structure and oligomeric states of YopM$_\mathrm{e13}$ and YopM$_\mathrm{e20}$ in the near future.

Secretion and translocation signals in YopM

A major unresolved question pertaining to type III secretion systems concerns the precise nature of the signals that target the effectors for secretion from the bacterium and translocation into the mammalian cytosol. Secretion and translocation appear to be triggered by different signals, because some *Yersinia* proteins are secreted but not translocated.[6] By analyzing the secretion of partially deleted *Yersinia* effectors fused to various reporters, it was possible to conclude that the secretion signal is located at the very N termini of the coding sequences.[6] However, the N termini of the effectors bear no obvious similarity to each other, nor do they resemble the signal peptides involved in the general *sec*-dependent secretion pathway.[1] Furthermore, no residues are cleaved from the effectors in transit.

In YopM, the secretion signal was localized to within the first 40 codons.[20] However, no electron density can be observed for the first 33 residues of YopM, suggesting that this part of the protein may be disordered. This is in contrast to the secretion signal in the *Yersinia* effector YopH, which is intimately interwoven into the structure of the N-terminal domain of the protein.[33]

Deletion analyses also revealed that the translocation signals in the *Yersinia* effectors are located immediately adjacent to the secretion signals in the N-terminal region of the proteins.[6] However, there is no apparent sequence similarity between translocation signals in different effectors, even within the same organism, and so it is not clear what features are recognized by the type III translocation apparatus. Yet, the effectors from one organism are capable of being recognized by the translocation machinery of another, suggesting the existence of a universal signal in all proteins that transit type III pathways.[34] In YopM, the translocation signal

has been localized within the first 100 residues and extends farther than the first 41 residues.[20] Residues 34-75 adopt a well-ordered, α-helical conformation in the crystal structure of YopM (Figure 12(a)). The tandem pair of helices packs tightly against the edge of the proximal LRR module. It is not clear how the translocation signal in YopM could be universal, however, because we were unable to identify a similar motif in the amino acid sequences of the other *Yersinia* effectors, nor does one exist in the structure of the N-terminal domain of YopH.[33]

Conversely, nearly all of the bacterial LRR proteins that have been identified thus far appear to contain a similar pair of α-helices, including the *Salmonella typhimurium* proteins SlrP,[35] SspHI and SspH2,[36] the *Shigella flexneri* virulence factor IpaH,[37] and at least three other ORFs of unknown function in the *Y. pestis* genome (Figure 12(b)). The most highly conserved residues are buried at the interface with the proximal LRR module in YopM, suggesting that the tandem helices adopt a similar conformation in all of these LRR proteins. Because their consensus LRR repeats are very close to that of YopM, we anticipate that the tertiary structures of these other bacterial LRR proteins will also contain very little conventional secondary structure. It seems plausible that one role of the tandem helices

in these bacterial LRR proteins may be to facilitate the folding of the adjacent LRR domains by providing a nucleation site for the organization of the proximal LRR module, after which the folding of subsequent LRR modules could proceed in a stepwise fashion. Additionally, the helices form a hydrophilic cap on the hydrophobic core of the proximal LRR module, which would otherwise be exposed to solvent. It is also possible that the tandem helices provide the asymmetry that is necessary for the proper orientation of YopM polypeptides in the tetramer. In any case, we believe that the pattern of amino acid conservation seems more consistent with a structural role for these helices in YopM and other bacterial LRR proteins than with a role in the translocation process.

Materials and Methods

Protein expression, purification and crystallization

A recombinant form of *Y. pestis* YopM with a C-terminal polyhistidine affinity tag was overproduced in *E. coli*, purified to homogeneity and crystallized, as described.[21] A new crystal form (space group $I4_122$) of YopM was produced in the same general manner as described previously, using 25 % ethylene glycol, 0.15 M calcium acetate, and 12 % isopropanol as the reservoir solution in hanging drop experiments. An ethyl mercury
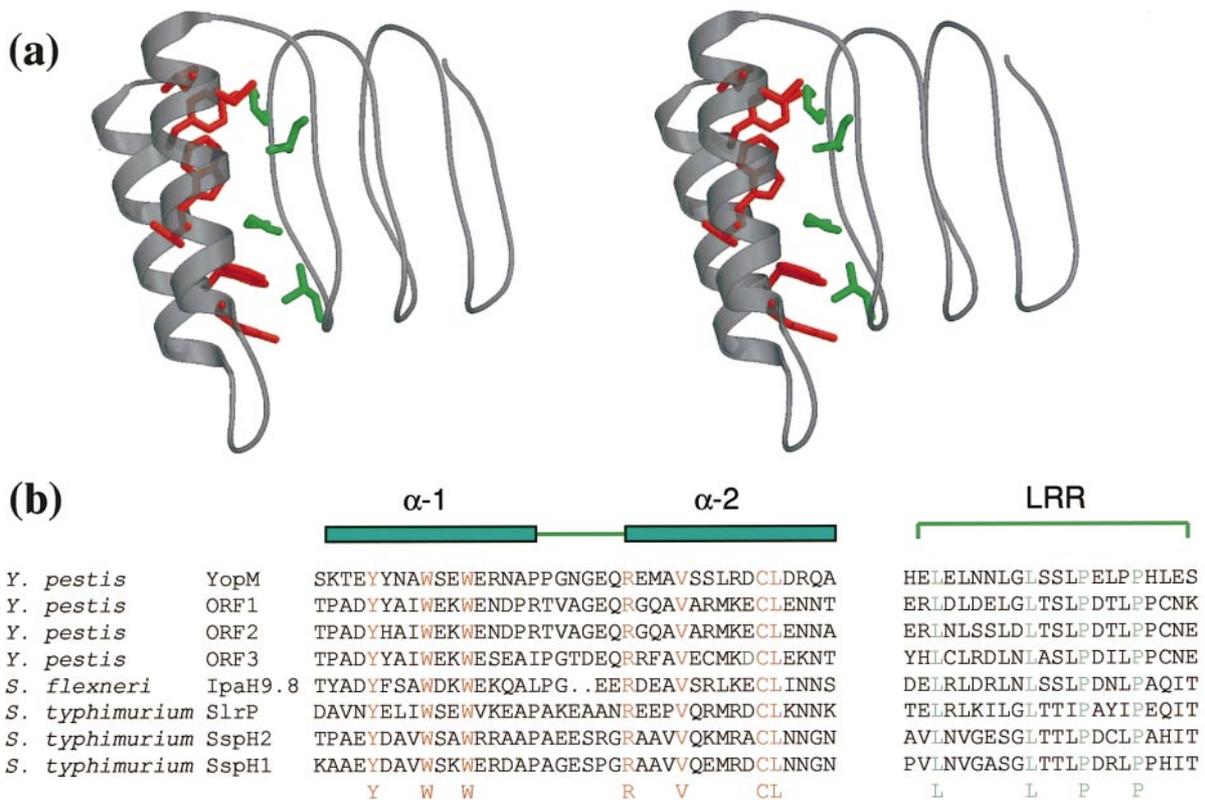

Figure 12. The putative translocation domain of YopM. (a) Stereo view of the residues involved in packing of the tandem α-helices (red residues) against the surface of the proximal LRR module (green residues) in YopM. (b) Alignment of amino acid sequences proximal to the first LRR module in YopM with other bacterial LRR proteins. Conserved amino acid residues that are involved in packing are colored according to the scheme in (a).

phosphate derivative was prepared by soaking the $P4_2 22$ crystals in a 6 mM solution of the reagent for two hours. Due to the lack of crystals of useful size, the second derivative was prepared by soaking the already mercury-derivatized crystals in 5 mM potassium tetrachloroplatinate for 2.5 hours.

Data collection and structure solution

Data collection was performed at beamline X9B of the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY). For data collection, the crystal of the $I4_1 22$ form (0.24 mm × 0.2 mm × 0.2 mm) was soaked in 45 % ethylene glycol for five minutes, mounted in a loop (Hampton Research) and flash-frozen in a cryogenic nitrogen stream at 100 K (Oxford Cryostream). X-ray diffraction was recorded using an ADSC Quantum 4 2 K × 2 K CCD detector positioned 130 mm from the crystal. Diffraction data were collected in two passes in order to obtain unbiased low-resolution values. During the first pass, the oscillation speed was eight minutes/deg. with an oscillation angle of 0.25°. For the second pass, the oscillation speed was increased to one minute/deg., using the same oscillation angle. A total of 55° of data were collected. Diffraction images were reduced and scaled using HKL[38] software. Essential data collection details for the other crystal forms were described.[21]

Data statistics and additional data collection details are summarized in Table 3.

Peaks corresponding to heavy atoms were clearly visible on isomorphous and anomalous Patterson maps. Structure solution for the $P4_2 22$ form was performed using the programs Solve[39] and CNS,[40] resulting in an overall figure of merit of 0.37 at 100-3 Å resolution. Density modification was performed with Resolve,[41] yielding a combined figure of merit of 0.71.

Model building and refinement

A total of 353 amino acid residues were fit to the electron density (Figure 2(a)) using the program O[42] and refined using SHELXL.[43] At this stage of refinement, two difference density peaks of ~15σ were found at the interface between crystallographically related monomers. These peaks were assigned as calcium ions on the basis of ligand composition and the coordination geometry of the metal. After several successive rounds of refinement and rebuilding, water molecules were assigned to >3σ peaks on the difference electron density map, which had reasonable hydrogen bonding geometry. A simulated annealing total omit map was calculated using CNS[40] to check for model inconsistencies. The 2.35 Å mercury derivative dataset was used to refine the final model. All

Table 3. Essential crystallographic data

	Native	Mercury derivative	Mercury + platinum derivative	Native
Space group	$P4_2 22$	$P4_2 22$	$P4_2 22$	$I4_1 22$
Solvent content (%)		61		79
Wavelength (Å)	1.00	1.0089	1.072	0.971
Mosaicity (deg.)	0.36	0.4	0.4	0.35
Unit cell dimensions				
a, b, c (Å)	109.36, 109.36, 101.50	109.52, 109.5, 101.63	109.48, 109.48, 101.71	149.43, 149.43, 191.18
($\alpha = \beta = \gamma = 90°$)				
Diffraction limit (Å)	2.48	2.35	3.0	2.1
Data completeness (%)	95.7 (93.6)[a]	96.8 (99.3)[a]	96.3 (89.1)[a]	97.3 (92.1)[a]
Unique reflections	23,645	24,924	12,752	58,456
$I/\sigma I$	12.4 (3.3)[a]	14.0 (3.5)[a]	12.1 (2.9)[a]	25.2 (3.5)[a]
Redundancy	2.9	3.4	2.8	4.9
R_{sym}	0.11(0.34)[a]	0.10(0.30)[a]	0.08(0.29)[a]	0.04(0.22)[a]
$\Delta_{isomorphous}$		0.15	0.17	
$\Delta_{anomalous}$		0.075	0.078	
Heavy atom sites		2 medium + 3 weak	3 medium[b]	
Overall figure of merit (100-3.0 Å)		0.36 (0.72 after density modification)		
Data:parameter ratio		3.5		5.1
R_{all} (%)		22		22
$R_{>4\sigma}$ (%)		19		18
R_{free} (%, for random 5 % of reflections)		26		24
$\langle B$-factors$\rangle$ (Å^2)				
All (Wilson plot)		18.3		19.4
All (structure)		24.2		33.2
Main-chain		22.4		28.4
C^α		22.6		28.5
Side-chain		26.3		32.3
Ramachandran plot				
Preferred (%)		85		78
Allowed (%)		15		22
RMSD				
Bond lengths (Å)		0.006		0.008
Bond angles (deg.)		2.4		1.9
Dihedral angles (deg.)		22.2		25.4

[a] Statistics for the highest-resolution shell are given in parentheses.
[b] Platinum sites only (platinum was soaked on top of mercury, and mercury sites are the same for both derivatives).

the mercury atoms found during structure solution were bound to either cysteine or histidine residues. Mercury atoms were refined with fixed partial occupancies that were estimated during the initial phase refinement, and with isotropic temperature displacement factors. To account for the anomalous effect of mercury, the Friedel equivalents in the dataset were not merged.

The structures of the other two crystal forms of YopM were solved by molecular replacement using the model obtained from the $P4_222$ crystal (AMoRe[44]). As expected, the AU of the $C222_1$ crystal form contained two protein monomers. Surprisingly, only one monomer was found in the AU of the $I4_122$ crystal form, resulting in a crystal with ~79% (v/v) solvent content. The initial refinement and rebuilding of the $I4_122$ crystal structure were performed, as described for the $P4_222$ structure. The high solvent content of the $I4$ crystal prompted us to use the more rigorous bulk solvent correction implemented in CNS. This correction resulted in ~4% drop of R_{free} (the same set of R_{free} reflections was used during all the stages of the refinement) and thus was deemed necessary. In order to maintain consistency in the refinement procedures, the final models of both $P4$ and $I4$ structures were subjected to the conjugated-gradient least-squares refinement procedure in CNS, followed by isotropic ADP refinement using individual B-factors for all atoms. The details of the crystallographic refinement are presented in Table 3.

Molecular graphics were generated with BobScript,[46] MolScript[47] and Raster3D.[48]

Structure validation, sequence alignment, modeling of orthologous YopMs and miscellaneous calculations

Experimental YopM models were verified with CNS,[40] PROCHECK,[49] and WHATIF.[50] Models of YopM$_{e13}$ and YopM$_{e20}$ from *Y. enterocolitica* were either constructed manually from individual LRRs, and then subjected to regularization in O and CNS, or submitted as sequence-backbone sets to the SWISS-MODEL homology threading server.[32] Regularized models were checked for abnormalities using WHATIF and corrected manually if necessary.

YopM sequences were aligned using the program Pile-Up (Wisconsin Package Version 10.2, Genetics Computer Group, Madison, WI) followed by considerable manual improvement. Two-dimensional surface map best-fit superposition was performed semi-manually using a custom PERL script.

Least-squares coordinate best-fit superimpositions, planes, interatomic distances, B-factor distributions, and backbone torsion angle calculations were calculated with LSQMAN,[51] MOLEMAN2,[52] CNS,[40] and several custom PERL scripts. Hydrogen bonds were calculated with HBPLUS,[53] and a custom PERL script was employed to sort and analyze them. All scripts are available from A.G.E. upon request.

Methods used to assess the oligomeric state of YopM

Cross-linking of YopM was performed by mixing solutions of divalent metal salts (Ca^{2+}, Mg^{2+}, Zn^{2+}) with 0.1 mg/ml^{-1} of protein in 30 mM NaPO$_4$ (pH 7.0), 50 mM NaCl. The solutions were incubated at room temperature for 20 minutes before glutaraldehyde was added to a final concentration of 0.1%. Cross-linking was performed for 15 minutes prior to gel electrophoresis.

Light-scattering analysis of YopM was performed using DAWN Eos (Wyatt Technology) 18-angle light scattering detector. YopM samples were diluted to the final concentration of 2 mg/ml in 20 mM Hepes (pH 7.4), 150 mM NaCl, in presence of either 1 mM EDTA or 0.8 mM CaCl$_2$. The samples (50 μl) were passed through a 1 cm × 30 cm Superdex-200 column that had been pre-equilibrated with one of these buffers at a flow rate of 0.3 ml/minute. UV/VIS, light-scattering, and refractive index (RI) were monitored during the run. Light-scattering from eight angles and the RI were used to estimate the molecular mass of the species in the peak fractions. Bovine serum albumin (Sigma) was used as a calibration standard.

Protein Data Bank accession numbers

The atomic coordinates and structure factors for the $P4_222$ and $I4_122$ crystal structures have been deposited with the Protein Data Bank[45] (accession codes 1G9U and 1JL5, respectively). The $C222_1$ crystal structure is still undergoing refinement and will be deposited at a later time.

References

1. Finlay, B. B. & Falkow, S. (1997). Common themes in microbial pathogenicity revisited. *Microbiol. Mol. Biol. Rev.* **61**, 136-169.
2. Hueck, C. J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62**, 379-433.
3. Galan, J. E. & Collmer, A. (1999). Type III secretion machines: bacterial devices for protein delivery into host cells. *Science,* **284**, 1322-1328.
4. Cheng, L. W. & Schneewind, O. (2000). Type III machines of Gram-negative bacteria: delivering the goods. *Trends Microbiol.* **8**, 214-220.
5. Galan, J. E. & Bliska, J. B. (1996). Cross-talk between bacterial pathogens and their host cells. *Annu. Rev. Cell Dev.* **12**, 221-255.
6. Cornelis, G. R. (2000). Molecular and cell biology aspects of plague. *Proc. Natl Acad. Sci. USA,* **97**, 8778-8783.
7. Persson, C., Carballeira, N., Wolf-Watz, H. & Fallman, M. (1997). PTPase YopH inhibits uptake of *Yersinia*, tyrosine phosphorylation of p130Cas and FAK, and the associated accumulation of these proteins in peripheral focal adhesions. *EMBO J.* **16**, 2307-2318.
8. Black, D. S., Marie-Cardine, A., Schraven, B. & Bliska, J. B. (2000). The *Yersinia* tyrosine phosphatase

YopH targets a novel adhesion-regulated signaling complex in macrophages. *Cell Microbiol.* **2**, 401-414.

9. Black, D. S., Montagna, L. G., Zitsmann, S. & Bliska, J. B. (1998). Identification of an amino-terminal substrate-binding domain in the *Yersinia* tyrosine phosphatase that is required for efficient recognition of focal adhesion targets. *Mol. Microbiol.* **29**, 1263-1274.

10. Juris, S. J., Rudolph, A. E., Huddler, D., Orth, K. & Dixon, J. E. (2000). A distinctive role for the *Yersinia* protein kinase: actin binding, kinase activation, and cytoskeleton disruption. *Proc. Natl Acad. Sci. USA,* **97**, 9431-9436.

11. Barz, C., Abahji, T. N., Trülzsch, K. & Heesemann, J. (2000). The *Yersinia* Ser/Thr protein kinase YpkA/YopO directly interacts with the small GTPases RhoA and Rac-1. *FEBS Letters,* **482**, 139-143.

12. Dukuzumuremyi, J.-M., Rosqvist, R., Hallberg, B., Akerström, B., Wolf-Watz, H. & Schesser, K. (2000). The *Yersinia* protein kinase A is a host factor inducible RhoA/Rac-binding virulence factor. *J. Biol. Chem.* **275**, 35281-35290.

13. Orth, K., Xu, X., Mudgett, M. B., Bao, Z. Q., Palmer, L. E. & Bliska, J. B., *et al.* (2000). Disruption of signaling by *Yersinia* effector YopJ, a ubiquitin-like protein protease. *Science,* **290**, 1594-1597.

14. von Pawel-Rammingen, U., Telepnev, M. V., Schmidt, G., Aktories, K., Wolf-Watz, H. & Rosqvist, R. (2000). GAP activity of the *Yersinia* YopE cytotoxin specifically targets the Rho pathway: a mechanism for disruption of actin microfilament structure. *Mol. Microbiol.* **36**, 737-748.

15. Black, D. S. & Bliska, J. B. (2000). The RhoGAP activity of the *Yersinia pseudotuberculosis* cytotoxin YopE is required for antiphagocytic function and virulence. *Mol. Microbiol.* **37**, 515-527.

16. Zumbihl, R., Aepfelbacher, M., Andor, A., Jacobi, C. A., Ruckdeschel, K., Rouot, B. *et al.* (1999). The cytotoxin YopT of *Yersinia enterocolitica* induces modification and cellular redistribution of the small GTP-binding protein RhoA. *J. Biol. Chem.* **274**, 29289-29293.

17. Leung, K. Y., Reisner, B. S. & Straley, S. C. (1990). YopM inhibits platelet aggregation and is necessary for virulence of *Yersinia pestis* in mice. *Infect. Immun.* **58**, 3262-3271.

18. Skrzypek, E., Cowan, C. & Straley, S. C. (1998). Targeting of the *Yersinia pestis* YopM protein into HeLa cells and intracellular trafficking to the nucleus. *Mol. Microbiol.* **30**, 1051-1065.

19. Kobe, B. & Deisenhofer, J. (1995). Proteins with leucine-rich repeats. *Curr. Opin. Struct. Biol.* **5**, 409-416.

20. Boland, A., Sory, M.-P., Iriarte, M., Kerbourch, C., Wattiau, P. & Cornelis, G. R. (1996). Status of YopM And YopN in the Yersinia Yop virulon: YopM of *Y. enterocolitica* is internalized inside the cytosol of PU5. 8 macrophages by the YopB,D,N delivery apparatus. *EMBO J.* **15**, 5191-5201.

21. Evdokimov, A. G., Anderson, D. E., Routzahn, K. M. & Waugh, D. S. (2000). Overproduction, purification, crystallization and preliminary X-ray diffraction analysis of YopM, an essential virulence factor extruded by the plague bacterium *Yersinia pestis*. *Acta Crystallog. sect. D,* **56**, 1676-1679.

22. Marino, M., Braun, L., Cossart, P. & Ghosh, P. (1999). Structure of the InlB leucine rich repeats, a domain that triggers host cell invasion by the bacterial pathogen *L. monocytogenes*. *Mol. Cell,* **4**, 1063-1072.

23. Price, S. R., Evans, P. R. & Nagai, K. (1998). U2B″-U2A′ protein complex bound to a fragment of U2 small nuclear RNA. *Nature,* **394**, 645-650.

24. Liker, E., Fernandez, E., Izaurralde, E. & Conti, E. (2000). The structure of the mRNA export factor TAP reveals a *cis* arrangement of a non-canonical RNP domain and an LRR domain. *EMBO J.* **19**, 5587-5598.

25. Zhang, H., Seabra, M. C. & Deisenhofer, J. (2000). Crystal structure of Rab geranylgeranyltransferase at 2.0 Å resolution. *Structure,* **8**, 241-245.

26. Hillig, R. C., Renault, L., Vetter, I. R., Drell, T., Wittinghofer, A. & Becker, J. (1999). The crystal structure of Rna1p: a new fold for a GTPase-activating protein. *Mol. Cell,* **3**, 781-791.

27. Kobe, B. & Deisenhofer, J. (1993). Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature,* **366**, 751-756.

28. Kajava, A. V. (1998). Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.* **277**, 519-527.

29. Hines, J., Skrzypek, E., Kajava, A. V. & Straley, S. C. (2001). Structure-function analysis of *Yersinia pestis* YopM's interaction with α-thrombin to rule on its significance in systemic plague and to model YopM's mechanism of binding host proteins. *Microb. Pathog.* **30**, 193-209.

30. Kobe, B. & Deisenhofer, J. (1994). The leucine-rich repeat, a versatile binding motif. *Trends Biochem. Sci.* **19**, 415-421.

31. Boland, A., Havauxs, S. & Cornelis, G. R. (1998). Heterogeneity of the *Yersinia* YopM protein. *Microb. Pathog.* **25**, 343-348.

32. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis,* **18**, 2714-2723.

33. Evdokimov, A. G., Tropea, J., Routzahn, K. M. & Waugh, D. S. (2001). Crystal structure of the N-terminal domain of *Yersinia pestis* YopH at 2.0 Å resolution. *Acta Crystallog. sect. D,* **57**, 793-799.

34. Rosqvist, R., Hakansson, S., Forsberg, A. & Wolf-Watz, H. (1995). Functional conservation of the secretion and translocation machinery for virulence proteins of yersiniae, salmonellae, and shigellae. *EMBO J.* **14**, 4187-4195.

35. Tsolis, R. M., Townsend, S. M., Miao, E. A., Miller, S. I., Ficht, T. A., Adams, L. G. *et al.* (1999). Identification of a putative *Salmonella enterica* serotype typhimurium host range factor with homology to IpaH and YopM by signature-tagged mutagenesis. *Infect. Immun.* **67**, 6385-6393.

36. Miao, E. A., Scherer, C. A., Tsolis, R. M., Kingsley, R. A., Adams, L. G., Baumler, A. J. *et al.* (1999). *Salmonella typhimurium* leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. *Mol. Microbiol.* **34**, 850-864.

37. Demers, B., Sansonetti, P. J. & Parsot, C. (1998). Induction of type III secretion in *Shigella flexneri* is associated with differential control of transcription of genes encoding secreted proteins. *EMBO J.* **17**, 2894-2903.

38. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276A**, 307-326.

39. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D,* **55**, 849-861.

40. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998).

Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

41. Terwilliger, T. C. (2000). Reciprocal-space solvent flattening. *Acta Crystallog. sect. D,* **55**, 1863-1871.

42. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect, A,* **47**, 110-119.

43. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319-343.

44. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. D,* **50**, 157-163.

45. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rogers, J. R. *et al.* (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.

46. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 132-134.

47. Kraulis, P. J. (1991). Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

48. Merrit, E. A. & Murphy, M. E. P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallog. sect. D,* **50**, 869-873.

49. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

50. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52-56.

51. Kleywegt, G. J. & Jones, T. A. (1994). A super position. *CCP4/ESF-EACBM Newsletter Prot. Crystallog.* **30**, 9-14.

52. Kleywegt, G. J. (1995). Dictionaries for heteros. *CCP4/ESF-EACBM Newsletter Prot. Crystallog.* **31**, 45-50.

53. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777-793.