Crystallographic Studies of Invasin, a Bacterial Adhesion Molecule from *Yersinia pseudotuberculosis*

Thesis by

Zsuzsa Andrea Hamburger

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2002

(Submitted May 9, 2002)

Acknowledgments

First and foremost I would like to thank my advisor, Pamela Bjorkman, for all of her support during my graduate career at Caltech. Pamela has been a great source of inspiration and it has been an honor having her as a mentor and a role model. Most of all I would like to thank Pamela for her patience, encouragement, color coordination and for allowing me to pursue my interest in bacterial pathogenesis and membrane protein crystallography.

I would also like to thank my collaborator, Ralph Isberg at Tufts University, for all of his enthusiasm and help with the invasin project. In addition, many thanks go to my thesis committee at Caltech including Marianne Bronner-Fraser, Steve Mayo, Doug Rees, and Ellen Rothenberg, for their invaluable input on my research.

I am indebted to Art Chirino, Luis Sanchez, Melanie Bennett, Dan Vaughn, Xiao-Dong Su, and Andy Yeh for teaching me crystallography. I would especially like to thank Melanie Bennett for teaching me how to grow and handle crystals, for helping me with my Ramachandran plots and for sharing with me her incredible notes.

Kaspar Locher has been integral in helping to express, refold and crystallize the transmembrane region of invasin. I would like to thank him for being a very good and patient teacher, for always taking time to talk to me about science as well as life and for letting me borrow tips, pipettes, trays, pumps, columns, buffers, cells, etc.

I would also like to thank all past and present members of the Bjorkman and Rees labs for their invaluable help and advice.

I would like to thank Ben Willcox, Chris O'Callaghan, Lynda Llamas, Renny Feldman, Joel Pomerantz, Amy Greenwood, Mariela Zirlinger, Sarah Lord and Krisztina

de la Torre for their support and friendship. I would also like to thank Jose Lebron and Marta Murphy for helping me master the art of Photoshop. How else could we have gotten Pamela and George W. Bush on the same picture? And of course many thanks go to Tara Chapman who provided friendship and much entertainment by always being up for crazy ideas, including tattooing, hair dying and piercings.

Life at Caltech would not have been the same without the surfing trips. Thank you Adam Politzer, Randal Bass, Astrid Heikema, Jeroen Corver, Lance Martin, Jim Pierce, and Pavel Strop, for letting me plan surfing trips and getting up early in the morning to catch low tide at Bolsa Chica.

Mark Weinswig, thank you for your love and encouragement during the completion of this thesis. Without the turkeys, chickens, kolbasz, barbecues, salads, cakes and pies I'd still be eating tuna out of a can. Most of all, thank you for listening to me and for teaching me to see the humor in even the worst situations. P.S. There is only my side of the story.

My special thanks go to Rejon Sarrazine who is a hero for choosing to live through the graduate work of not only one but two Hamburger girls. Thank you Rejon for listening to us laugh, cry and complain, for fixing furniture, cars, computer programs, for all the dinners you prepared and brought to us in lab, for all the movies we made you watch and for your A.V. expertise.

Finally, and most importantly, thank you Agi Hamburger for being the best sister, friend, baymate anybody can ever have, for always being enthusiastic and helpful with my projects, oh, and the oligos. Samy and Alex, thank you for the whiskers and the cuddles. Mama, Papa and Nagyi, thank you for always being on my side, for letting me

call at all hours of the night with good or bad news, for your unconditional love and support and for learning about DNA, proteins and crystallography.

Abstract

Bacterial pathogens, such as *Yersinia pseudotuberculosis*, must bind and enter normally non-phagocytic cells to establish infection. The protein responsible for mediating uptake of the bacterium is a 986-residue outer membrane protein called invasin. Invasin binds to several members of the β_1 integrin family, presumably activating a reorganization of the host cytoskeleton to form pseudopods that envelop the bacterium. Integrin binding has been localized to the extracellular region of invasin (Inv497) comprised by the COOH-terminal 497 residues. In order to gain insight into host cell entry by *Yersinia pseudotuberculosis*, we solved the 2.3 Å crystal structure of Inv497. The structure reveals five domains that form a 180 Å rod with structural similarities to tandem fibronectin-III domains. The integrin-binding surfaces of invasin and fibronectin include similarly located key residues, but in the context of different folds and surface shapes. The structures of invasin and fibronectin provide an example of convergent evolution, in which invasin presents an optimized surface for integrin binding compared with host substrates. We have also initiated structural analyses of the NH_2-terminal ~500 residues of invasin, which are required for outer membrane localization and for presentation of the integrin-binding region of invasin. We expressed this region of invasin as inclusion bodies in *E. coli*, and refolded the protein in the presence of detergents. We have also obtained microcrystals of this membrane protein. Circular dichroism studies indicate that this region of invasin is composed of mainly β-structure. As the transmembrane regions of outer membrane proteins of known structure are β-barrels, this region of invasin is also presumed to fold into such a structure. Proteolysis experiments suggest that the N-terminal 70 amino acids of invasin may form a separate

domain from the invasin transmembrane region, analogous to that found in another outer membrane protein, the sucrose-specific porin ScrY. Equilibrium sedimentation analytical ultracentrifugation studies indicate the protein is monomeric in solution. Black bilayer experiments suggest that this region of the protein does not contain a pore and thus plays the role of an outer membrane anchor for the presentation of the integrin-binding domain on the cell surface.

Table of Contents

Chapter 1:

Introduction

Introduction

Yersinia Infection

The ability of bacterial pathogens to bind to host cells is an essential step in the establishment of disease. Adherence to host cell surfaces localizes pathogens to the appropriate target tissues and results in either extracellular colonization or penetration of the bacteria into host cells. Entry into host cells provides bacteria with an environment permissive for growth, a protective niche to avoid host defense mechanisms or a means to gain access to deeper tissues. The latter strategy is used by *Yersinia pseudotuberculosis* and *Yersinia enterocolitica*, two enteropathogenic Gram-negative bacteria that cause gastroenteritis. *Yersinia* infection occurs when contaminated food is ingested and the bacteria translocate via M cells across the intestinal epithelium at Peyer's patches. Translocated bacteria enter the lymphatic system and colonize the liver and spleen where they divide extracellularly [1-3].

Yersinia Virulence Factors

Multiple *Yersinia* factors have been identified as being important for interaction with mammalian cells. These include the pYV virulence plasmid-encoded YadA protein [4] and the chromosomally encoded ail [5] and invasin proteins [6].

YadA YadA is encoded by the 70-kb virulence plasmid (pYV) that is common to pathogenic *Yersiniae* [7]. YadA expression is temperature regulated; it is expressed at highest levels at 37 °C and not below 30 °C [8]. The protein is a homotrimeric outer membrane protein that has been shown to mediate low rate invasion to tissue culture cells [9, 10]. Expression of YadA is important for virulence by *Y. enterocolitica* but is

dispensable for *Y. pseudotuberculosis* and is not expressed by the closely related *Y. pestis*, the causative agent of the bubonic plague [4, 11]. The exact mechanism by which YadA contributes to virulence is not known, however, a large number of functions have been assigned to this protein. YadA has been shown to bind to a number of host factors including collagen, laminin, fibronectin, epithelial cells and mucus (reviewed in [12]. Collagen plays an important structural role in joints and high affinity collagen binding by YadA is thought to be necessary for *Y. enterocolitica*-induced reactive arthritis [13]. Bacteria expressing YadA autoagglutinate, which may protect them against host defenses [14]. YadA is also thought to render *Yersiniae* serum resistant by binding to serum resistance factor H, a negative regulatory component of complement activation [15].

Ail Ail, from *Y. enterocolitica* was also identified to play a role in *Yersinia* virulence by its ability to confer an invasive phenotype to *E. coli* cells [5]. Ail is an outer membrane protein that is predicted to fold into an 8-stranded β-barrel protein based on its homology (45% sequence identity) to OmpX [16, 17]. In addition to promoting cell attachment in *Y. enterocolitica*, expression of ail has also been shown to promote serum-resistance in both *Y. enterocolitica* and *Y. pseudotuberculosis* [18-20].

Invasin Although both YadA and ail have been shown to promote low level adhesion to and invasion of bacteria into host cells, the most efficient and best characterized pathway for *Yersinia* uptake involves the outer membrane protein invasin. Mutations in invasin result in the inability of bacteria to enter M cells and inefficient translocation into Peyer's patches [2, 21].

Invasin is a 986 amino acid outer membrane protein encoded by the chromosomal *inv* gene and was first identified by its ability to convert *E. coli* K-12 cells into

pathogenic bacteria capable of entering normally non-phagocytic cells [6]. Invasin is maximally produced at temperatures below 28 °C and the protein is thought to play a role during the early stages of infection [22, 23]. At later stages of infection, after colonization of deeper tissues, invasin expression is not required for virulence [21]. In fact, the bacteria actively resist uptake by phagocytic cells due to the expression of virulence plasmid encoded factors including the Yops that are expressed at 37 °C [24]. These proteins are translocated into host cells by the type III secretion system where they interrupt phagocytic signaling pathways by interacting with cytoskeletal and signaling proteins including focal adhesion kinase (FAK), CAS, Paxillin and RAC1 [25-29].

Invasin is expressed by both *Yersinia pseudotuberculosis* and *Yersinia enterocolitica* but not *Yersinia pestis*, due to the presence of an insertion that silences the gene [30]. It is thought that invasin expression is not required for this pathogen because *Yersinia pestis* does not need to cross the intestinal epithelial layer, being instead delivered directly to the blood stream by a flea bite.

The invasin protein contains multiple domains with separable functions (Figure 1). Invasin-mediated cell binding has been localized to the C-terminal 497 amino acids [31]. This region of the protein binds to a subset of β_1 integrins and mediates uptake of the bacteria into host cells [32]. The N-terminal ~500 amino acids of invasin contain a Sec-dependent signal sequence and are critical for outer membrane localization. This region of the protein is predicted to fold into an amphipathic channel and to function as an autotransporter for expression of the C-terminal domain on the cell surface [31].

Autotransporters, a family of large, multidomain outer membrane proteins, self-direct their transport across the outer membrane of Gram-negative bacteria (Figure 2).

These proteins contain three functional domains: 1) an N-terminal Sec-dependent signal sequence that allows transport of the proteins across the inner membrane, 2) a C-terminal, ~500 residue autotransporter domain predicted to fold into an amphipathic β-barrel that promotes translocation of 3) the N-terminal passenger domain across the outer membrane [33-35]. The autotransporter domains of these proteins share high sequence similarity. Because invasin contains the three functional units found in proteins of the autotransporter family, it has been proposed to function as an autotransporter despite a lack of significant sequence similarity with known autotransporters [31].

Invasin-mediated cell binding is cation-dependent and is localized to the C-terminal region of invasin [32]. In fact, the C-terminal 497 residues of invasin can be expressed as a soluble protein that is capable of binding and penetrating mammalian cells when attached to bacteria or beads [36]. The smallest region of invasin capable of cell binding and entry is the C-terminal 192 amino acids [31, 37]. Monoclonal antibodies specific for this region of the protein block attachment of invasin to mammalian cells and prevent bacterial penetration [38]. Invasin residues required for integrin binding map to this region and include residues within the 903-913 region (Figure 1) [32, 36]. The presence of a disulfide bond formed between Cys^{906} and Cys^{982} is also required for integrin binding, presumably because it is necessary for correct folding [39]. In addition, residues added to the C-terminus of the protein have been shown to interfere with cell attachment activity suggesting that the C-terminal portion of invasin is at the integrin binding site [40].

The Invasin Family

Invasin shares sequence homology with intimins, a family of enteropathogenic bacterial proteins expressed on the surface of enteropathogenic and enterohaemorrhagic *E. coli* [41]. The two proteins have different functions as invasin mediates uptake of *Yersinia* into host cells while intimin mediates attachment to host cell surfaces that leads to the formation of actin-rich pedestals beneath the bacteria [42, 43]. Unlike most bacterial receptors, which are host proteins residing upon the host cell surface, the receptor for intimin is a bacterially secreted protein, Tir. Tir is secreted into the host cell using the bacterially-encoded type III secretion system, and becomes inserted into the host membrane upon translocation [44]. Tir then mediates the intimate attachment of bacteria to intestinal epithelial cells by binding intimin.

The highest degree of sequence similarity between invasin and intimin lies at the N-terminal ~500 residue region of the proteins. This region is functionally exchangeable; replacing the N-terminal 489 amino acids of invasin for the corresponding N-terminal 539 amino acids of intimin results in proper membrane localization and presentation of invasin epitopes on the bacterial cell surface [45]. The highest degree of divergence lies at the C-terminal, receptor-binding region of the proteins, thus reflecting the different receptor specificities of invasin and intimin.

Intimin Structure The three-dimensional structure of a fragment of the extracellular region of intimin (residues 658-939 of enteropathogenic *E. coli* intimin) including the C-terminal Tir-binding domain has been solved both by NMR and x-ray crystallography (Figure 3A) [46, 47]. The structures reveal that intimin is composed of tandem domains including two N-terminal immunoglobulin superfamily (IgSF)-like

domains and a C-terminal C-type lectin-like domain. IgSF domains consists of two antiparallel β-sheets stacked to form a β-sandwich and are found in many cell surface adhesion molecules where they occur in tandem and project away from the cell surface to mediate a variety of cell recognition events (reviewed in [48]. C-type lectins are cell surface proteins that mediate Ca^{2+}-dependent carbohydrate recognition via an α/β fold with a stretch of extended structure that contains residues directly involved in Ca^{2+} coordination (reviewed in [49]. The presence of a lectin-like domain in intimin has been suggested to function in carbohydrate recognition on host cells [46]. However, intimin lacks the surface-exposed loop and conserved residues required for carbohydrate recognition (Figure 3B) [46, 47]. Thus if intimin binds carbohydrates, it does in a significantly different manner than C-type lectins [47].

Structural Basis of Tir Recognition by Intimin Mutagenesis studies indicated that residues in the C-type lectin like domain of intimin are required for Tir binding (J.M. Leong, personal communication). These findings were confirmed when the crystal structure of the intimin-binding domain of Tir (Tir IBD) was solved in complex with the intimin fragment described above (Figure 3) [47]. Tir is predicted to cross the host cell membrane twice with the N- and C-termini of Tir located in the cytoplasm of the host cell (Figure 3C). The Tir IBD is found between the two transmembrane regions and is located extracellularly [50]. The structure of Tir IBD is hairpin shaped with two long α-helices and a β-hairpin. Two molecules of Tir IBD dimerize to form an antiparallel four-helix bundle mediated by the two α-helices of each monomer [47]. Intimin and Tir binding is mediated by the C-type lectin-like domain of intimin and the β-hairpin and the N-terminus of the second helix of Tir. The rod-like structure of intimin, as seen in the

crystal structures, suggests that during bacterial attachment intimin and Tir are horizontally oriented in the intermembrane space (Figure 3C) [47].

Invasin Receptors

Integrins Although invasin and intimin share structural and sequence features, invasin binds to host-derived receptors that are members of the β_1 integrin family [32]. Integrins are heterodimeric integral membrane proteins that mediate communication between the extracellular environment and the cytoskeleton by binding to extracellular matrix proteins or cell surface receptors and cytoskeletal components [51]. By binding to cytoskeletal components, integrins trigger an intracellular signal, "outside-in signaling," that includes phosphorylation and rearrangement of the actin cytoskeleton. Integrins can also send signals in the opposite direction across the cell membrane, "inside-out signaling," since the adhesiveness of the integrin extracellular regions can be modulated by signals within the cells. The ability to link extracellular ligand binding to alterations in cellular architecture makes integrins a key player in many fundamental cell processes including cell-cell interaction, cell adhesion, cell migration and cell differentiation [51-53].

An integrin heterodimer is formed by the noncovalent interaction of one α subunit (150-180 kDa) and one β subunit (~90 kDa). Both the α and the β subunits are type I membrane proteins with large extracellular domains that bind extracellular ligands and short cytoplasmic tails that bind cytoskeletal components. In mammals, eighteen α and eight β subunits have been identified that combine to form over 20 different receptors with distinct ligand specificities. The α subunits are classified into two groups depending

on the presence of a ~200 amino acid insertion (I domain) [54]. These I domains are required for ligand binding for those integrins that contain them and isolated I domains retain many of the ligand binding properties of the parent integrins, including divalent cation dependency [55, 56]. For those integrins that lack an I domain ligand recognition requires an I-like domain (βA) present in all integrin β subunits [57, 58]. β subunits are numbered one through eight (β_1-β_8) and integrin families are classified based on the specific β subunit they contain.

Integrin Structures Until recently, the only three-dimensional information about integrins came from crystal structures of isolated I domains in "liganded" or "unliganded" conformations [57, 59-62]. I domains fold into classic Rossmann (nucleotide-binding) folds composed of a central 6-stranded β-sheet surrounded by seven α-helices. The structures revealed that ligand binding is achieved by acidic residues present in the integrin ligands that complete coordination of the metals at the metal ion-dependent adhesion sites (MIDAS) present in all I domains (Figure 4B) [57, 62].

Further information about integrin structure, including information about the conformation of the entire extracellular region, was recently obtained when the three-dimensional crystal structure of the extracellular region of the $\alpha_v\beta_3$ integrin was determined [58]. The integrin extracellular region consists of 12 domains. The α_v subunit is composed of an N-terminal seven-bladed β-propeller followed by three β-sandwich domains. The β_3 subunit is composed of an N-terminal I-like domain (βA-domain), followed by an Ig-like domain, a PSI-domain, four EGF domains and a C-terminal α/β domain. The overall structure is bent as both the α_v and β_3 subunits fold back on themselves (Figure 4A) [58, 63].

The crystal structure of the extracellular region of $\alpha_v\beta_3$ was also determined in complex with an Arg-Gly-Asp (RGD)-containing pentapeptide (Figure 4A) [63]. The RGD sequence plays a central role in ligand recognition by integrins and is found in a large number of integrin ligands [64-67]. Peptides containing this sequence can mimic integrin ligands as they promote cell adhesion as well as competitively inhibit integrin binding by ligands. Since the $\alpha_v\beta_3$ integrin does not contain an I domain, this structure provided information about ligand binding in the absence of an I domain. The ligand-binding domain of $\alpha_v\beta_3$ is formed by the β-propeller in α_v and the βA-domain present in β_3. Ligand binding requires both α_v and β_3 and the RGD pentapeptide makes extensive contacts with both subunits (Figure 4A) [63]. As in the case of the isolated I domains, the Mn^{2+} ion is coordinated at the MIDAS in the βA-domain of β_3 with the aspartate residue of the RGD completing the coordination (Figure 4B). The arginine residue of the peptide contacts the α_v subunit and the glycine lies at the interface between α_v and β_3. The conformation of the RGD-containing peptide ligand in $\alpha_v\beta_3$ is in an almost identical conformation as in echistatin, a natural integrin ligand, suggesting that the structure of the integrin-RGD peptide complex is physiologically relevant [63].

Integrin Ligands Integrin ligands include extracellular matrix proteins as well as cell surface receptors. Although integrin ligands are diverse and can be unrelated, they all contain acidic residues required for integrin binding. Many integrin ligands belong to the IgSF superfamily of cell surface receptors. The crystal structure of a number of these proteins, including ICAM-1, ICAM-2, VCAM-1, MAdCAM-1 and Fn type III repeats 7-10, have been solved [68-73]. The overall structures of these different proteins are very similar: they are rod-like molecules composed of tandem Ig-like folds (Figure 5). In

addition many of them also have their critical integrin-binding residues positioned at the edge of the IgSF domains. Interestingly, for ICAM-1 and ICAM-2, molecules that bind to I domain-containing integrins, these critical residues are found on a flat surface in the context of a β-strand. For VCAM-1, MAdCAM-1 and Fn, proteins that bind to integrins that do not contain I domains, residues critical for integrin binding are located on protruding loops (Figure 5). These structural differences have been proposed to allow differentiation between integrins that contain or lack I domains [68].

Integrins are also a common target for pathogens. Many bacteria as well as viruses, including *Neisseria meningitis*, *Bordetella pertusis*, adenovirus, Coxsackievirus, bind to integrins directly or indirectly through integrin-binding proteins [74-77]. Integrins are expressed on the apical surface of M cells, a class of specialized epithelial cells overlaying Peyer's patches in the intestinal epithelium, and are thus ideally suited for bacterial internalization [78]. By binding to β1 integrins, invasin activates a reorganization of the host cytoskeleton to form pseudopods that envelop the bacterium (Figure 6) [79, 80]. Invasin receptors include the $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_6\beta_1$, and $\alpha_v\beta_I$ integrins. The $\alpha_2\beta_1$ integrin has been shown not to bind to invasin and this is attributed to the presence of an I-domain in α_2 that is not present in α_3, α_4, α_5, α_6, and α_v [32].

Integrin Binding by Invasin

The cell binding domain of invasin is not homologous to any known integrin ligand. However, competition assays and mutagenesis studies indicate that invasin and fibronectin (Fn), an $\alpha_5\beta_1$ ligand, bind to integrins at the same or overlapping sites [81]. A critical determinant of fibronectin binding to integrins is the RGD sequence found in Fn

type III repeat 10 (Fn-III 10) [64-67]. The *Yersinia pseudotuberculosis* invasin lacks the RGD sequence required for integrin binding by fibronectin although RGD peptides can inhibit invasin binding to the $\alpha_5\beta_1$ integrin [81]. Although the presence of the RGD sequence is not required for integrin binding, all integrin ligands contain an acidic residue that is critical for divalent cation-dependent integrin binding. This residue is required to complete the coordination of the metal ion by the integrin (Figure 4B and C) [57, 62]. A critical aspartate residue (Asp911) has been identified in the cell-binding region of invasin that probably plays the role of the Asp in the RGD sequence (Figure 1). Even a conservative mutation of this residue to Asn reduces cell binding 10-fold and abolishes cellular entry, while the mutation to Ala abolishes integrin binding activity [36].

A second region of invasin approximately 100 amino acids from Asp911 contains additional residues that are implicated in integrin binding (Figure 1). Mutations at residues centered around Asp811 resulted in reduced bacterial uptake [82]. The presence of a second region of invasin that plays a role in integrin binding is reminiscent of the fibronectin synergy region located in Fn-III 9 that is required for maximal $\alpha_5\beta_1$ integrin-dependent cell spreading (Figure 5B) [83-85].

Invasin vs. Fibronectin

Although invasin and fibronectin bind to a mutually exclusive site on integrins, the result of integrin binding by these two proteins is very different. While invasin-integrin binding results in bacterial uptake, fibronectin-integrin binding does not result in efficient Fn uptake [79]. Studies indicate that this is due to the different affinities of the two proteins for integrins. The affinity of invasin for integrins is two orders of magnitude

higher than that determined for fibronectin (K_d=5nM for invasin, Kd=800 nM for fibronectin [81, 86]. This high affinity binding is required for bacterial uptake because mutations that lower the affinity result in little or no uptake [79]. Integrin binding by invasin results in the activation of signaling pathways that cause the reorganization of the actin cytoskeleton and the formation of pseudopods on the cell surface (Figure 6). Actin polymerization has been shown to be required for bacterial uptake since cytochalasin D inhibits invasion without affecting attachment of the bacteria to host cells [87]. The high affinity integrin binding allows invasin to compete with natural ligands for binding to β_1 integrins, and as a result, *Yersinia* is able to "zipper" the host cell surface around itself leading to uptake (Figure 6).

Integrin-Mediated Bacterial Uptake

Integrins interact with the cytoskeleton through their cytoplasmic tails [88]. Mutations that reduce the interaction of integrins with the cytoskeleton increase bacterial uptake [80]. These mutations are thought to allow for greater mobility of the receptor resulting in the recruitment of more integrin receptors around the bacterial surface. Bacterial uptake also seems to share some common features with clathrin-mediated endocytosis. Mutations that result in a decreased ability for integrins to associate with clathrin-coated pits decrease bacterial uptake [80]. The exact role clathrin-coated pits play in bacterial uptake is not known.

Multimerization of invasin has also been shown to increase bacterial uptake. A region of the *Yersinia pseudotuberculosis* invasin, corresponding to residues 594-692, has been shown to undergo homotypic interactions [89]. This region of the protein is not

found in *Yersinia enterocolitica* invasin, which is thought to be why this bacterium is less efficient at promoting uptake than *Yersinia pseudotuberculosis.* In the absence of multimerization, increasing the concentration of invasin and integrins has been shown to increase the efficiency of uptake [79, 90].

Ligand binding by integrins generates a variety of intracellular signals that results in changes in tyrosine phosphorylation, intracellular pH and actin polymerization [52]. *Yersinia* uptake has been shown to involve activation of host tyrosine kinases including focal adhesion kinase and Src [91]. In addition, CAS, Crk, CDC42Hs, WASp, Arp2/3 complex, as well as Rac and Rho GTPases have been implicated in uptake [91-93].

Structural Characterization of Invasin

Despite the extensive use of antibiotics, bacterial pathogens continue to be a leading cause of morbidity and mortality worldwide. Thus, understanding the molecular mechanisms of microbial pathogenesis is of great importance. Many bacterial species have evolved specific structures that allow them to take advantage of important signaling pathways in mammalian cells. The enteropathogenic *Yersinia* species express the outer membrane protein invasin in order to bind β_1 integrins and thus utilize host machinery for initiating cytoskeletal events necessary for bacterial internalization. Determining the structure of invasin and comparing it to the structures of known integrin ligands is important for gaining an understanding of bacterial invasion of mammalian cells. In chapter 2, I describe the determination of the crystal structure of the extracellular integrin-binding region of invasin. Comparison of the structure with that of fibronectin shows how *Yersinia* is able to manipulate integrin binding by structural mimicry through

convergent evolution. In chapter 3, I describe the biochemical characterization and work toward crystallization of the transmembrane region of invasin. This region of the protein is required for outer membrane localization [31]. Our results show that this region of the protein can be refolded from inclusion bodies into a monomeric protein composed of mainly β-sheet structure that anchors the integrin-binding region of invasin to the bacterial cell surface.

References

1. Pepe, J.C. and Miller, V.L. The biological role of invasin during a *Yersinia enterocolitica* infection. Infectious Agents & Disease. **2** (1993) 236-41.
2. Marra, A. and Isberg, R.R. Invasin-dependent and invasin-independent pathways for translocation of *Yersinia pseudotuberculosis* across the Peyer's patch intestinal epithelium. Infect Immun. **65** (1997) 3412-21.
3. Autenrieth, I.B. and Firsching, R. Penetration of M cells and destruction of Peyer's patches by *Yersinia enterocolitica*: an ultrastructural and histological study. Journal of Medical Microbiology. **44** (1996) 285-94.
4. Skurnik, M. and Wolf-Watz, H. Analysis of the yopA gene encoding the Yop1 virulence determinants of *Yersinia spp*. Mol Microbiol. **3** (1989) 517-29.
5. Miller, V.L., Bliska, J.B., and Falkow, S. Nucleotide sequence of the *Yersinia enterocolitica* ail gene and characterization of the Ail protein product. J Bacteriol. **172** (1990) 1062-9.
6. Isberg, R.R., Voorhis, D.L., and Falkow, S. Identification of invasin: a protein that allows enteric bacteria to penetrate cultured mammalian cells. Cell. **50** (1987) 769-78.
7. Portnoy, D.A. and Martinez, R.J. Role of a plasmid in the pathogenicity of *Yersinia species*. Curr Top Microbiol Immunol. **118** (1985) 29-51.
8. Skurnik, M. and Toivanen, P. LcrF is the temperature-regulated activator of the yadA gene of *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. J Bacteriol. **174** (1992) 2047-51.

9. Mack, D., Heesemann, J., and Laufs, R. Characterization of different oligomeric species of the *Yersinia enterocolitica* outer membrane protein YadA. Med Microbiol Immunol (Berl). **183** (1994) 217-27.

10. Yang, Y. and Isberg, R.R. Cellular internalization in the absence of invasin expression is promoted by the *Yersinia pseudotuberculosis* yadA product. Infect Immun. **61** (1993) 3907-13.

11. Rosqvist, R., Skurnik, M., and Wolf, W.H. Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations. Nature. **334** (1988) 522-4.

12. El Tahir, Y. and Skurnik, M. YadA, the multifaceted *Yersinia* adhesin. Int J Med Microbiol. **291** (2001) 209-18.

13. Gripenberg-Lerche, C., Skurnik, M., and Toivanen, P. Role of YadA-mediated collagen binding in arthritogenicity of *Yersinia enterocolitica* serotype O:8: experimental studies with rats. Infect Immun. **63** (1995) 3222-6.

14. Skurnik, M., Bolin, I., Heikkinen, H., Piha, S., and Wolf-Watz, H. Virulence plasmid-associated autoagglutination in *Yersinia spp*. J Bacteriol. **158** (1984) 1033-6.

15. China, B., Sory, M.P., N'Guyen, B.T., De Bruyere, M., and Cornelis, G.R. Role of the YadA protein in prevention of opsonization of *Yersinia enterocolitica* by C3b molecules. Infect Immun. **61** (1993) 3129-36.

16. Heffernan, E.J., Harwood, J., Fierer, J., and Guiney, D. The *Salmonella typhimurium* virulence plasmid complement resistance gene rck is homologous to a family of virulence-related outer membrane protein genes, including pagC and ail. J Bacteriol. **174** (1992) 84-91.

17. Vogt, J. and Schulz, G.E. The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. Structure Fold Des. **7** (1999) 1301-9.

18. Bliska, J.B. and Falkow, S. Bacterial resistance to complement killing mediated by the Ail protein of *Yersinia enterocolitica*. Proc Natl Acad Sci USA. **89** (1992) 3561-5.

19. Pierson, D.E. and Falkow, S. The ail gene of *Yersinia enterocolitica* has a role in the ability of the organism to survive serum killing. Infect Immun. **61** (1993) 1846-52.

20. Yang, Y., Merriam, J.J., Mueller, J.P., and Isberg, R.R. The psa locus is responsible for thermoinducible binding of *Yersinia pseudotuberculosis* to cultured cells. Infect Immun. **64** (1996) 2483-9.

21. Pepe, J.C. and Miller, V.L. *Yersinia enterocolitica* invasin: a primary role in the initiation of infection. Proc Nat Acad Sci USA. **90** (1993) 6473-7.

22. Isberg, R.R., Swain, A., and Falkow, S. Analysis of expression and thermoregulation of the *Yersinia pseudotuberculosis* inv gene with hybrid proteins. Infect Immun. **56** (1988) 2133-8.

23. Pierson, D.E. Mutations affecting lipopolysaccharide enhance ail-mediated entry of *Yersinia enterocolitica* into mammalian cells. J Bacteriol. **176** (1994) 4043-51.

24. Cornelis, G.R., Biot, T., Lambert de Rouvroit, C., Michiels, T., Mulder, B., Sluiters, C., Sory, M.P., Van Bouchaute, M., and Vanooteghem, J.C. The *Yersinia* yop regulon. Mol Microbiol. **3** (1989) 1455-9.

25. Black, D.S. and Bliska, J.B. Identification of p130Cas as a substrate of *Yersinia* YopH (Yop51), a bacterial protein tyrosine phosphatase that translocates into mammalian cells and targets focal adhesions. EMBO Journal. **16** (1997) 2730-44.

26. Black, D.S., Montagna, L.G., Zitsmann, S., and Bliska, J.B. 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** (1998) 1263-74.

27. Hamid, N., Gustavsson, A., Andersson, K., McGee, K., Persson, C., Rudd, C.E., and Fallman, M. YopH dephosphorylates Cas and Fyn-binding protein in macrophages. Microb Pathog. **27** (1999) 231-42.

28. Persson, C., Carballeira, N., Wolf-Watz, H., and Fallman, M. The 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** (1997) 2307-18.

29. Black, D.S. and Bliska, J.B. The RhoGAP activity of the *Yersinia pseudotuberculosis* cytotoxin YopE is required for antiphagocytic function and virulence. Mol Microbiol. **37** (2000) 515-27.

30. Simonet, M., Riot, B., Fortineau, N., and Berche, P. Invasin production by *Yersinia pestis* is abolished by insertion of an IS200-like element within the inv gene. Infect Immun. **64** (1996) 375-9.

31. Leong, J.M., Fournier, R.S., and Isberg, R.R. Identification of the integrin binding domain of the *Yersinia pseudotuberculosis* invasin protein. Embo J. **9** (1990) 1979-89.

32. Isberg, R.R. and Leong, J.M. Multiple beta 1 chain integrins are receptors for invasin, a protein that promotes bacterial penetration into mammalian cells. Cell. **60** (1990) 861-71.

33. Mackman, N., Nicaud, J.M., Gray, L., and Holland, I.B. Identification of polypeptides required for the export of haemolysin 2001 from *Escherichia coli.* Mol Gen Genet. **201** (1985) 529-36.

34. Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T.F. Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature. **325** (1987) 458-62.

35. Klauser, T., Pohlner, J., and Meyer, T.F. The secretion pathway of IgA protease-type proteins in gram-negative bacteria. Bioessays. **15** (1993) 799-805.

36. Leong, J.M., Morrissey, P.E., Marra, A., and Isberg, R.R. An aspartate residue of the *Yersinia pseudotuberculosis* invasin protein that is critical for integrin binding. Embo J. **14** (1995) 422-31.

37. Leong, J.M., Fournier, R.S., and Isberg, R.R. Mapping and topographic localization of epitopes of the *Yersinia pseudotuberculosis* invasin protein. Infect Immun. **59** (1991) 3424-33.

38. Isberg, R.R. and Leong, J.M. Cultured mammalian cells attach to the invasin protein of *Yersinia pseudotuberculosis*. Proc Natl Acad Sci USA. **85** (1988) 6682-6.

39. Leong, J.M., Morrissey, P.E., and Isberg, R.R. A 76-amino acid disulfide loop in the *Yersinia pseudotuberculosis* invasin protein is required for integrin receptor recognition. Journal of Biological Chemistry. **268** (1993) 20524-32.

40. Isberg, R.R., Yang, Y., and Voorhis, D.L. Residues added to the carboxyl terminus of the *Yersinia pseudotuberculosis* invasin protein interfere with recognition by integrin receptors. J Biol Chem. **268** (1993) 15840-6.

41. Yu, J. and Kaper, J.B. Cloning and characterization of the eae gene of enterohemorrhagic *Escherichia coli* O157-H7. Mol. Microbiol. **6** (1992) 411-417.

42. Jerse, A.E., Yu, J., Tall, B.D., and Kaper, J.B. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. Proc Natl Acad Sci USA. **87** (1990) 7839-43.

43. Donnenberg, M.S., Tacket, C.O., James, S.P., Losonsky, G., Nataro, J.P., Wasserman, S.S., Kaper, J.B., and Levine, M.M. Role of the eaeA gene in experimental enteropathogenic *Escherichia coli* infection. J Clin Invest. **92** (1993) 1412-7.

44. Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A., and Finlay, B.B. Enteropathogenic *Escherichia coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell. **91** (1997) 511-20.

45. Liu, H., Magoun, L., Luperchio, S., Schauer, D.B., and Leong, J.M. The Tir-binding region of enterohaemorrhagic *Escherichia coli* intimin is sufficient to trigger actin condensation after bacterial-induced host cell signalling. Mol Microbiol. **34** (1999) 67-81.

46. Kelly, G., Prasannan, S., Daniell, S., Fleming, K., Frankel, G., Dougan, G., Connerton, I., and Matthews, S. Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. Nat. Struct. Biol. **6** (1999) 313-8.

47. Luo, Y., Frey, E.A., Pfuetzner, R.A., Creagh, A.L., Knoechel, D.G., Haynes, C.A., Finlay, B.B., and Strynadka, N.C. Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. Nature. **405** (2000) 1073-7.

48. Leahy, D.J. Implications of atomic-resolution structures for cell adhesion. Annu Rev Cell Dev Biol. **13** (1997) 363-93.

49. Weis, W.I. and Drickamer, K. Structural basis of lectin-carbohydrate recognition. Annu Rev Biochem. **65** (1996) 441-73.

50. Hartland, E.L., Batchelor, M., Delahay, R.M., Hale, C., Matthews, S., Dougan, G., Knutton, S., Connerton, I., and Frankel, G. Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. Mol Microbiol. **32** (1999) 151-8.

51. Hynes, R.O. Integrins: versatility, modulation, and signaling in cell adhesion. Cell. **69** (1992) 11-25.

52. Cary, L.A., Han, D.C., and Guan, J.L. Integrin-mediated signal transduction pathways. Histol Histopathol. **14** (1999) 1001-9.

53. Plow, E.F., Haas, T.A., Zhang, L., Loftus, J., and Smith, J.W. Ligand binding to integrins. J Biol Chem. **275** (2000) 21785-8.

54. Humphries, M.J. Integrin structure. Biochem Soc Trans. **28** (2000) 311-39.

55. Tuckwell, D., Calderwood, D.A., Green, L.J., and Humphries, M.J. Integrin alpha 2 I-domain is a binding site for collagens. J Cell Sci. **108** (1995) 1629-37.

56. Dickeson, S.K., Walsh, J.J., and Santoro, S.A. Contributions of the I and EF hand domains to the divalent cation-dependent collagen binding activity of the alpha 2 beta1 integrin. J Biol Chem. **272** (1997) 7661-8.

57. Lee, J.O., Rieu, P., Arnaout, M.A., and Liddington, R. Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). Cell. **80** (1995) 631-8.

58. Xiong, J.P., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D.L., Joachimiak, A., Goodman, S.L., and Arnaout, M.A. Crystal structure of the extracellular segment of integrin alpha V beta 3. Science. **294** (2001) 339-45.

59. Baldwin, E.T., Sarver, R.W., Bryant, G.L., Jr., Curry, K.A., Fairbanks, M.B., Finzel, B.C., Garlick, R.L., Heinrikson, R.L., Horton, N.C., Kelley, L.L., Mildner, A.M., Moon, J.B., Mott, J.E., Mutchler, V.T., Tomich, C.S., Watenpaugh, K.D., and Wiley, V.H. Cation binding to the integrin CD11b I domain and activation model assessment. Structure. **6** (1998) 923-35.

60. Emsley, J., King, S.L., Bergelson, J.M., and Liddington, R.C. Crystal structure of the I domain from integrin alpha 2 beta 1. J Biol Chem. **272** (1997) 28512-7.

61. Qu, A. and Leahy, D.J. Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin. Proc Natl Acad Sci USA. **92** (1995) 10277-81.

62. Emsley, J., Knight, C.G., Farndale, R.W., Barnes, M.J., and Liddington, R.C. Structural basis of collagen recognition by integrin alpha 2 beta 1. Cell. **101** (2000) 47-56.

63. Xiong, J.P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S.L., and Arnaout, M.A. Crystal structure of the extracellular segment of integrin alpha V beta 3 in complex with an Arg-Gly-Asp ligand. Science. **296** (2002) 151-5.

64. Pierschbacher, M.D. and Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Nature. **309** (1984) 30-3.

65. Pierschbacher, M.D., Ruoslahti, E., Sundelin, J., Lind, P., and Peterson, P.A. The cell attachment domain of fibronectin. Determination of the primary structure. J Biol Chem. **257** (1982) 9593-7.

66. Ruoslahti, E. and Pierschbacher, M.D. Arg-Gly-Asp: a versatile cell recognition signal. Cell. **44** (1986) 517-8.

67. D'Souza, S.E., Ginsberg, M.H., and Plow, E.F. Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. Trends Biochem Sci. **16** (1991) 246-50.

68. Casasnovas, J.M., Springer, T.A., Liu, J.H., Harrison, S.C., and Wang, J.H. Crystal structure of ICAM-2 reveals a distinctive integrin recognition surface. Nature. **387** (1997) 312-5.

69. Wang, J.-h., Pepinsky, R.B., Stehle, T., Liu, J.H., Karpausas, M., Browning, B., and Osborn, L. The crystal structure of an N-terminal two-domain fragment of vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on the domain 1 C-D loop can inhibit VCAM-1-alpha 4 integrin interaction. Proc. Natl. Acad. Sci. USA. **92** (1995) 5714-8.

70. Leahy, D.J., Aukhil, I., and Erickson, H.P. 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. Cell. **84** (1996) 155-64.

71. Tan, K., Casasnovas, J.M., Liu, J.H., Briskin, M.J., Springer, T.A., and Wang, J.H. The structure of immunoglobulin superfamily domains 1 and 2 of

MAdCAM-1 reveals novel features important for integrin recognition. Structure. **6** (1998) 793-801.

72. Casasnovas, J.M., Stehle, T., Liu, J.-h., Wang, J.-h., and Springer, T.A. A dimeric crystal structure for the N-terminal two domains of intercellular adhesion molecule-1. Proc. Natl. Acad. Sci. USA. **95** (1998) 4134-9.

73. Bella, J., Kolatkar, P.R., Marlor, C.W., Greve, J.M., and Rossmann, M.G. The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand. Proc. Natl. Acad. Sci. USA. **95** (1998) 4140-5.

74. Virji, M., Makepeace, K., and Moxon, E.R. Distinct mechanisms of interactions of Opc-expressing meningococci at apical and basolateral surfaces of human endothelial cells; the role of integrins in apical interactions. Mol Microbiol. **14** (1994) 173-84.

75. Relman, D., Tuomanen, E., Falkow, S., Golenbock, D.T., Saukkonen, K., and Wright, S.D. Recognition of a bacterial adhesion by an integrin: macrophage CR3 (alpha M beta 2, CD11b/CD18) binds filamentous hemagglutinin of *Bordetella pertussis*. Cell. **61** (1990) 1375-82.

76. Wickham, T.J., Mathias, P., Cheresh, D.A., and Nemerow, G.R. Integrins alpha V beta 3 and alpha V beta 5 promote adenovirus internalization but not virus attachment. Cell. **73** (1993) 309-19.

77. Roivainen, M., Piirainen, L., Hovi, T., Virtanen, I., Riikonen, T., Heino, J., and Hyypia, T. Entry of coxsackievirus A9 into host cells: specific interactions with alpha V beta 3 integrin, the vitronectin receptor. Virology. **203** (1994) 357-65.

78. Clark, M.A., Hirst, B.H., and Jepson, M.A. M-cell surface beta1 integrin expression and invasin-mediated targeting of *Yersinia pseudotuberculosis* to mouse Peyer's patch M cells. Infect Immun. **66** (1998) 1237-43.

79. Tran Van Nhieu, G. and Isberg, R.R. Bacterial internalization mediated by beta 1 chain integrins is determined by ligand affinity and receptor density. Embo J. **12** (1993) 1887-95.

80. Tran Van Nhieu, G., Krukonis, E.S., Reszka, A.A., Horwitz, A.F., and Isberg, R.R. Mutations in the cytoplasmic domain of the integrin beta1 chain indicate a role for endocytosis factors in bacterial internalization. J Biol Chem. **271** (1996) 7665-72.

81. Tran Van Nhieu, G. and Isberg, R.R. The *Yersinia pseudotuberculosis* invasin protein and human fibronectin bind to mutually exclusive sites on the alpha 5 beta 1 integrin receptor. Journal of Biological Chemistry. **266** (1991) 24367-75.

82. Saltman, L.H., Lu, Y., Zaharias, E.M., and Isberg, R.R. A region of the *Yersinia pseudotuberculosis* invasin protein that contributes to high affinity binding to integrin receptors. J Biol Chem. **271** (1996) 23438-44.

83. Bowditch, R.D., Hariharan, M., Tominna, E.F., Smith, J.W., Yamada, K.M., Getzoff, E.D., and Ginsberg, M.H. Identification of a novel integrin binding site in fibronectin. Differential utilization by β 3 integrins. J. Biol. Chem. **269** (1994) 10856-63.

84. Aota, S., Nomizu, M., and Yamada, K.M. The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function. Journal of Biological Chemistry. **269** (1994) 24756-61.

85. Ugarova, T.P., Zamarron, C., Veklich, Y., Bowditch, R.D., Ginsberg, M.H., Weisel, J.W., and Plow, E.F. Conformational transitions in the cell binding domain of fibronectin. Biochemistry. **34** (1995) 4457-66.

86. Akiyama, S.K. and Yamada, K.M. The interaction of plasma fibronectin with fibroblastic cells in suspension. J Biol Chem. **260** (1985) 4492-500.

87. Finlay, B.B. and Falkow, S. Comparison of the invasion strategies used by *Salmonella cholerae-suis*, *Shigella flexneri* and *Yersinia enterocolitica* to enter cultured animal cells: endosome acidification is not required for bacterial invasion or intracellular replication. Biochimie. **70** (1988) 1089-99.

88. Reszka, A.A., Hayashi, Y., and Horwitz, A.F. Identification of amino acid sequences in the integrin beta 1 cytoplasmic domain implicated in cytoskeletal association. J Cell Biol. **117** (1992) 1321-30.

89. Dersch, P. and Isberg, R.R. An immunoglobulin superfamily-like domain unique to the *Yersinia pseudotuberculosis* invasin protein is required for stimulation of bacterial uptake via integrin receptors. Infect Immun. **68** (2000) 2930-8.

90. Isberg, R.R. Discrimination between intracellular uptake and surface adhesion of bacterial pathogens. Science. **252** (1991) 934-8.

91. Alrutz, M.A. and Isberg, R.R. Involvement of focal adhesion kinase in invasin-mediated uptake. Proc Natl Acad Sci USA. **95** (1998) 13658-63.

92. Weidow, C.L., Black, D.S., Bliska, J.B., and Bouton, A.H. CAS/Crk signalling mediates uptake of *Yersinia* into human epithelial cells. Cell Microbiol. **2** (2000) 549-60.

93. Wiedemann, A., Linder, S., Grassl, G., Albert, M., Autenrieth, I., and Aepfelbacher, M. *Yersinia enterocolitica* invasin triggers phagocytosis via beta 1 integrins, CDC42Hs and WASp in macrophages. Cell Microbiol. **3** (2001) 693-702.

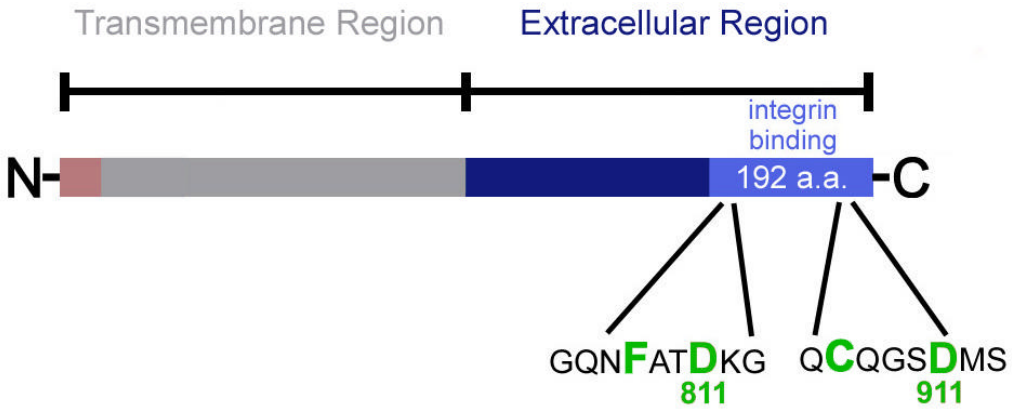


Figure 1. Schematic diagram of the *Yersinia pseudotuberculosis* invasin protein. The predicted transmembrane region of invasin is shown in gray with the Sec-dependent signal sequence shown in pink. The C-terminal 192 amino acid region of invasin, required for integrin binding, is shown in light blue while the rest of the extracellular region is shown in navy. Residues that have been shown to be required for integrin binding are shown in single letter code. Residues that have been mutated and that eliminate cell adhesion (C^{907} and D^{911}) or drastically reduce it (F^{808} and D^{811}) are shown in green. Adapted from Isberg and Barnes, 2001.

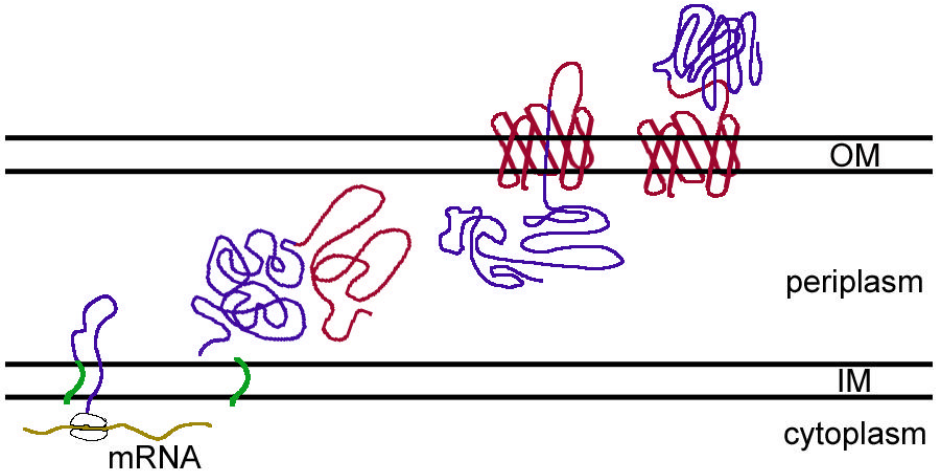


Figure 2. Model of surface display by autotransporters in Gram-negative bacteria. The nascent peptide chain of the autotransporter protein is translocated across the inner membrane (IM) in a Sec-dependent manner. After the signal sequence is cleaved from the precursor protein, the C-terminal transporter inserts into the outer membrane (OM) forming a translocation pore through which the N-terminal extracellular domain can pass. The signal peptide is shown in green, the C-terminal transport domain in red and the N-terminal extracellular domain in blue. Adapted from Klauser et al., 1993.

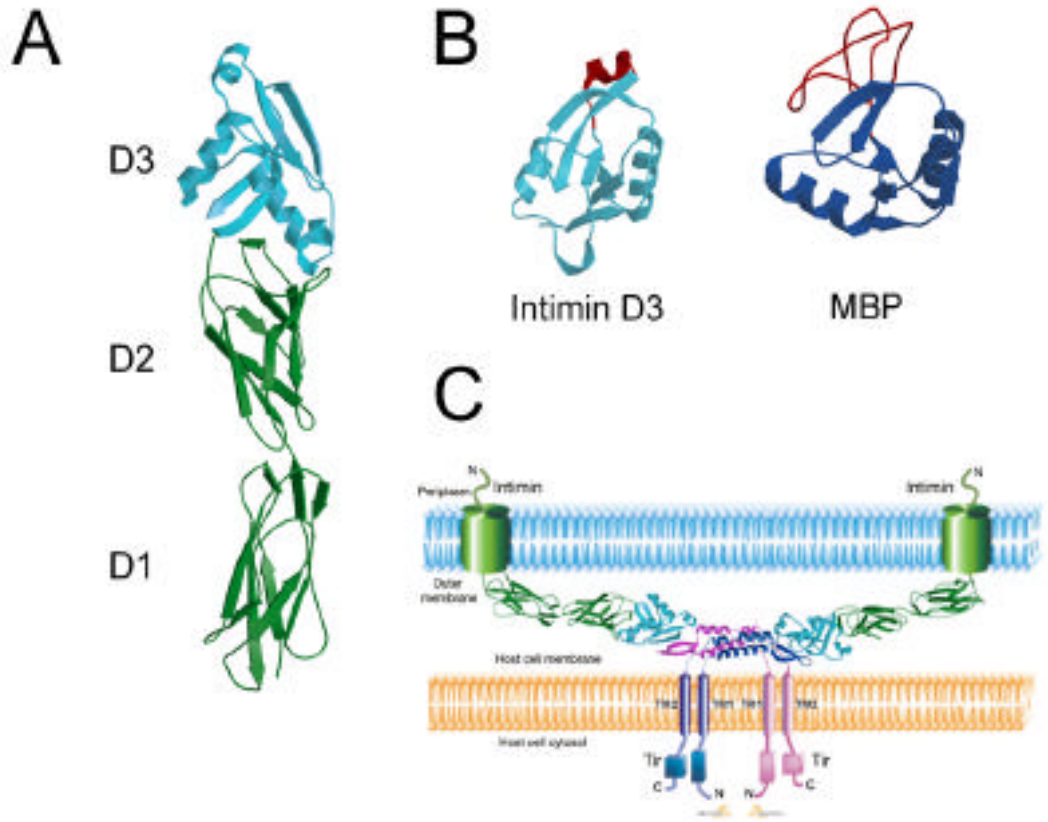


Figure 3. Crystal structure of intimin and Tir. A. Ribbon diagram of a fragment of the extracellular domain of intimin (Luo et al., 2000). D1 and D2 (green) resemble IgSF-like domains while D3 (cyan) is a C-type lectin-like domain. B. D3 of intimin is shown next to a canonical C-type lectin domain from mannose-binding protein (MBP) (Weis and Drickamer, 1994). The calcium-binding loop in MBP and its truncated counterpart in intimin are shown in red. C. A model for the intimin and Tir-mediated interface between enteropathogenic *E. coli* and a host cell. Intimin is shown in green and cyan and a Tir dimer is shown in pink and blue. Adapted from Luo et al., 2000.

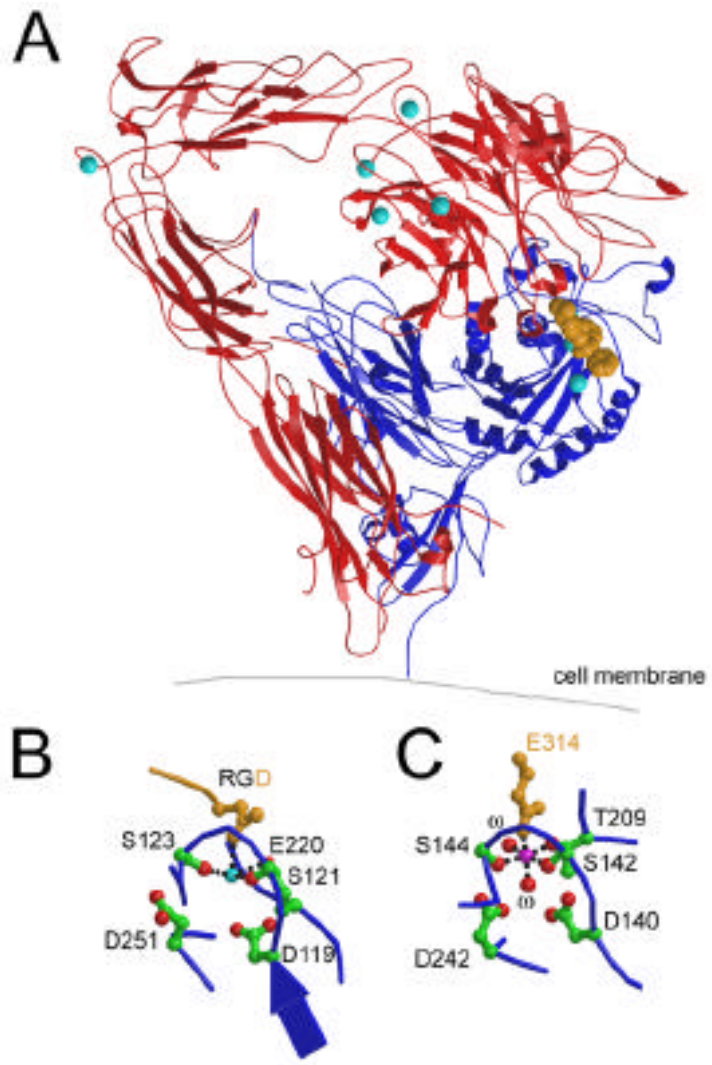


Figure 4. Structural features of integrins. A. Ribbon diagram of the extracellular segment of $\alpha_v\beta_3$ integrin in complex with an RGD-containing pentapeptide (Xiong et al., 2002). α_v is shown in red, β_3 in blue, the peptide in gold and the Mn^{2+} ions in cyan. B. Diagram of the MIDAS domain in β_3. C. Diagram of the MIDAS motif from the I domain-containing integrin, CR3 (Lee et al., 1995). Residues involved in metal coordination are indicated (ω, water). Hydrogen bonds are shown by dashed lines, oxygen residues are red, Mn^{2+} is cyan and Mg^{2+} is magenta.

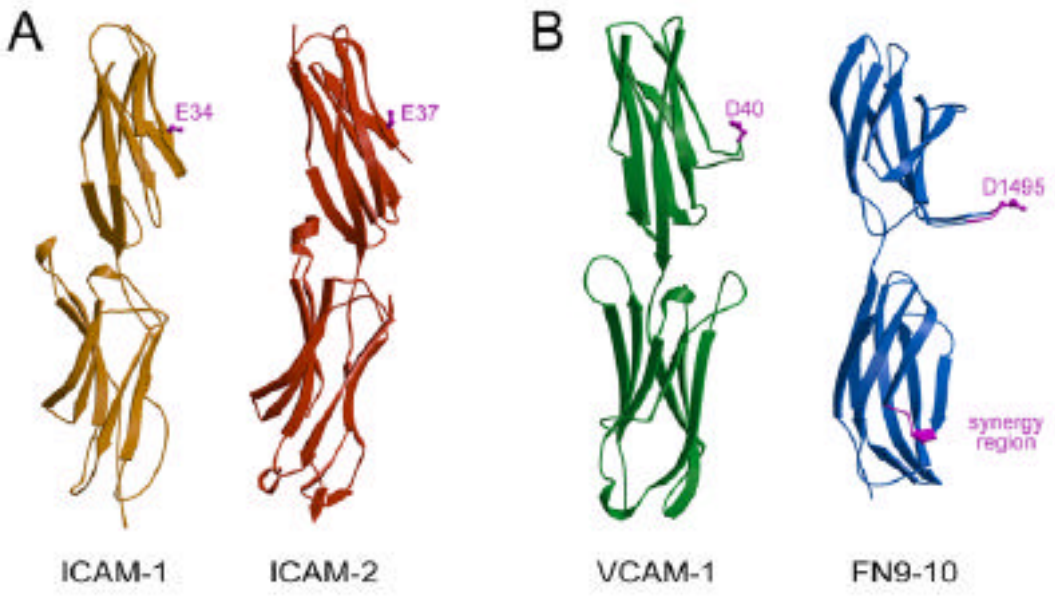


Figure 5. Mammalian integrin ligands. Ribbon diagram representation of the structures of the mammalian integrin ligands ICAM-1 (Bella et al., 1998), ICAM-2 (Casasnovas et al., 1997), VCAM-1 (Wang et al., 1995), FN9-10 (Leahy et al., 1996). Residues critical for integrin binding are shown in magenta. A. Integrin ligands that bind to I domain-containing integrins have their acidic residue on a flat surface in the context of a β-strand. B. Integrin ligands that bind to I-domain lacking integrins have their critical integrin-binding residues on a surface exposed loop. For Fn9-10, the synergy region is also shown.

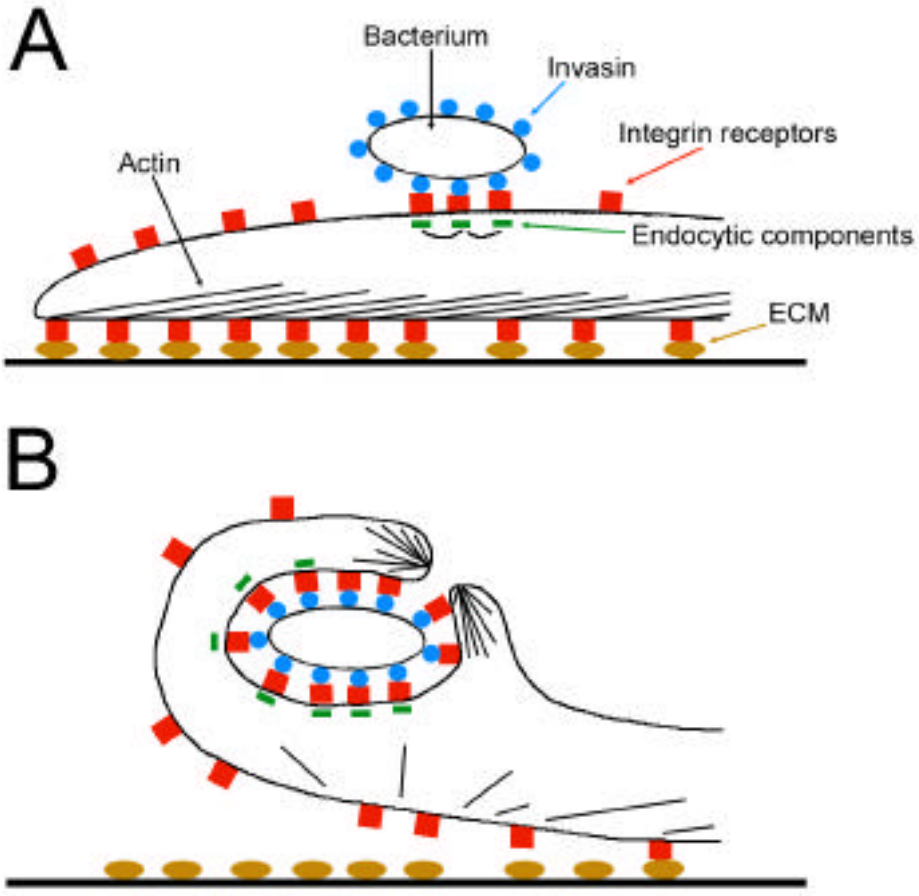


Figure 6. *Yersinia* internalization via a zipper-like mechanism. A. Bacteria expressing invasin bind to β_1 integrins on mammalian cells. B. Interaction between invasin and integrins results in the reorganization of the actin cytoskeleton and the formation of pseudopods on the mammalian cell surface. High affinity interaction between invasin and integrin allows zippering of the host cell membrane around the bacterium leading to uptake. ECM, extracellular matrix. Adapted from Isberg and Tran Van Nhieu, 1995.

Chapter 2:

This chapter was published as Z. A. Hamburger, M. S. Brown, R. R. Isberg and P. J. Bjorkman, “Crystal structure of invasin: a bacterial integrin-binding protein.” *Science* **286**, 291-295 (1999).

Crystal Structure of Invasin: a Bacterial Integrin-Binding Protein

Zsuzsa A. Hamburger,[1] Michele S. Brown,[3] Ralph R. Isberg,[3]

Pamela J. Bjorkman[1,2,*]

[1]Division of Biology 156-29 and [2]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, U.S.A.

[3]Department of Microbiology and Molecular Biology, Howard Hughes Medical Institute, Tufts University School of Medicine, Boston, MA 02111, U.S.A.

[*] To whom correspondence should be addressed. E-mail: bjorkman@its.caltech.edu

The *Yersinia pseudotuberculosis* invasin protein promotes bacterial entry by binding to host cell integrins with higher affinity than natural substrates such as fibronectin. The 2.3 Å crystal structure of the invasin extracellular region reveals five domains that form a 180 Å rod with structural similarities to tandem fibronectin-III domains. The integrin-binding surfaces of invasin and fibronectin include similarly located key residues, but in the context of different folds and surface shapes. The structures of invasin and fibronectin provide an example of convergent evolution, in which invasin presents an optimized surface for integrin binding compared with host substrates.

Many bacterial pathogens bind and enter eukaryotic cells to establish infection. *Y. pseudotuberculosis* and *Y. enterocolitica* are enteropathogenic Gram-negative bacteria that cause gastroenteritis when they are translocated across the intestinal epithelium at Peyer's patches by way of M cells. Translocated bacteria enter the lymphatic system and colonize the liver and spleen, where they grow mainly extracellularly (*1*). Invasin is an outer membrane protein required for efficient uptake of *Yersinia* into M cells (*2*,*3*). Invasin mediates entry into eukaryotic cells by binding to members of the β1 integrin family that lack I domains, such as $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_6\beta_1$, and $\alpha_v\beta_1$ (*3*). Integrins are heterodimeric integral membrane proteins that mediate communication between the extracellular environment and the cytoskeleton by binding to both cytoskeletal components and either extracellular matrix proteins or cell surface proteins (*4*). Invasin binding to β1 integrins is thought to activate a reorganization of the host cytoskeleton to form pseudopods that envelop the bacterium (*5*). Another family of enteropathogenic bacterial proteins related to invasin, the intimins, mediate attachment of the bacteria to host cells by binding to a bacterially secreted protein Tir, which upon secretion becomes inserted into the host membrane (*6*).

Y. pseudotuberculosis invasin is a 986 residue protein. The N-terminal ~500 amino acids, which are thought to reside in the outer membrane (*7*), is related (~36% sequence identity) to the analogous regions of intimins (*8*). The C-terminal 497 residues of invasin, which make up the extracellular region, can be expressed as a soluble protein (Inv497) that binds integrins and promotes uptake when attached to bacteria or beads (*9*). The shortest invasin fragment capable of binding integrins consists of the C-terminal 192 amino acids (*7*). This fragment is not homologous to the integrin binding domains of

fibronectin (the fibronectin type III repeats 9 and 10; Fn-III 9-10) (*8*), although mutagenesis studies and competition assays indicate that invasin and fibronectin bind to $\alpha_3\beta_1$ and $\alpha_5\beta_1$ integrins at the same or overlapping sites (*10*). The integrin binding region of invasin also lacks significant sequence identity with the corresponding regions of intimins (~20% identity) (*8*). To gain insight into enteric bacterial pathogenesis and to compare the structural basis of integrin binding by invasin and Fn-III domains, we solved the crystal structure of Inv497.

Inv497 was expressed in *Escherichia coli* and purified (*9*). The structure was solved to 2.3 Å by multiple isomorphous replacement with anomalous scattering (MIRAS) (Table 1) (*11,12*). Inv497 is a rod-like molecule with overall dimensions of ~180 Å x 30 Å x 30 Å (Fig. 1A), consistent with analytical ultracentrifugation analyses that suggest the fragment has an extended monomeric structure in solution (*13*). The Inv497 structure bears an overall resemblance to that of another $\alpha_5\beta_1$ binding fragment, Fn-III repeats 7-10 (Fn-III 7-10) (*14*), as they are both elongated molecules composed of tandem domains. The first four Inv497 domains (D1, D2, D3, and D4) are composed mainly of β-structure, while the fifth domain (D5) includes α-helices and β-sheets. Despite only 20% sequence identity (*8*), the D3 to D5 region of Inv497 is structurally similar to a 280 residue fragment of the extracellular portion of enteropathogenic *E. coli* intimin (*14*).

The four N-terminal domains of Inv497 adopt folds resembling eukaryotic members of the immunoglobulin superfamily (IgSF) (*15*), although the Inv497 domains do not share significant sequence identity with IgSF domains and lack the disulfide bond and core residues conserved in IgSF structures (*8,15*). D1 belongs to the I2 set of the

IgSF, and D2 and D3 belong to the I1 set (*15*). D4 adopts the folding topology of the C1 set of IgSF domains, a fold seen in the constant domains of antibodies, T cell receptors and MHC molecules (*15*). Unlike these C1 domains, D4 of Inv497 includes a 15 amino acid insertion between strands A and B that forms two additional β-strands (A'' and A''') (Fig. 1B). D1 and D2 of the intimin fragment are also Ig-like, and the second domain includes an insertion similar to that found in Inv497 D4 (*14*).

D5 of Inv497 has a folding topology related to that of C-type lectin-like domains (CTLDs) (Fig. 1B) (*16*). This superfamily includes true C-type lectins such as mannose binding protein (*14*) and E-selectin (*14*), which contain carbohydrate recognition domains (CRDs) that bind carbohydrates in a calcium dependent manner, and evolutionarily related proteins such as the Ly49 family of natural killer cell receptors, which bind ligands in the absence of calcium and may not recognize carbohydrates (*16*). A characteristic feature of C-type lectin CRDs is a long stretch of extended structure including one or two calcium binding sites, which is required for carbohydrate recognition (*16*). The C-terminal domains of both Inv497 and intimin lack these calcium binding loops (Fig. 1B) (*14,16*). Inv497 is not known to bind carbohydrates (*17*), thus the significance of the CTLD fold remains to be determined. By analogy with Ly49A, which recognizes a carbohydrate-independent epitope on its class I MHC ligand (*16*), Inv497 may recognize an unglycosylated region of integrins.

Like CTLDs and structurally related proteins such as the COOH-terminal domain of intimin (D3) (*14*), Inv497 D5 is composed of two antiparallel β-sheets with interspersed α-helical and loop regions, and includes a disulfide bond linking helix 1 to β-strand 5 (Fig. 1B). An additional disulfide bond linking β-strand 3 and the loop

following strand 4 is found in CTLDs and CRDs but is absent in Inv497 D5 and intimin D3. Whereas C-type lectin CRDs contain two α-helices located between the first and second β-strands, the region corresponding to the second helix is replaced by a loop in Inv497 D5 (Fig. 1B) and CD94, a component of the CD94/NKG2 natural killer cell receptor (*14*). In Inv497 D5, the loop is preceded by a two turn α-helix (residues 917-921 and 931-936) interrupted by a nine residue loop (residues 922-930) (Fig. 1C) (*18*). The corresponding region in intimin was not interpretable in the NMR structure (*14*).

Extensive interactions between Inv497 D4 and D5 create a superdomain that is composed of the 192 residues identified as necessary and sufficient for integrin binding (*7*). The interface between D4 and D5 is significantly larger than the interfaces between tandem IgSF domains and between the Ig-like invasin domains (D4-D5 buried surface area is 1925 $Å^2$ compared to ~ 500 $Å^2$ for IgSF interfaces) (*19*). The D4-D5 interface is predominantly hydrophobic although a number of hydrogen bonds are also present (Fig. 2). The interrupted helix in D5 and strands A'' and A''' in D4 (Fig. 1B) play a major role in the interaction between these two domains. In particular, a portion of the loop within the D5 interrupted helix contacts the A''' strand in D4 (Fig. 2). In addition, strand A''' hydrogen bonds with strand one of D5, extending the second β-sheet of the CTLD (Fig. 1B). The large buried surface area at the D4-D5 interface and the consequent rigidity of this portion of invasin contrasts with the flexibility between the integrin-binding portions of fibronectin, inferred from lower than average interdomain buried surface areas at these interfaces (Fn-III 9-10, Fn-III 12-13) (*14, 19*). Interdomain flexibility in fibronectin was proposed to facilitate integrin binding (*15*), and is also observed in the structures of two other integrin binding proteins, ICAM-1 (*14*) and

VCAM-1 (*14*). However, invasin, which shows little or no interdomain flexibility in its integrin binding region, binds at least five different integrins and binds $\alpha_5\beta_1$ with about 100-fold higher affinity than does fibronectin (*5,10*). High affinity binding of invasin is necessary for bacterial internalization, as studies have shown that bacteria coated with lower affinity ligands for $\alpha_5\beta_1$ bind, but do not penetrate, mammalian cells (*5,10*).

Invasin residues important for integrin binding include 903-913 (*7,20*), which form helix 1 and the loop after it in D5. The disulfide bond between Cys^{906} and Cys^{982}, conserved in all CTLDs (Fig. 1B), is required for integrin binding (*20*), presumably because it is necessary for correct folding. Although invasin lacks an Arg-Gly-Asp (RGD) sequence, which is critical for the interaction of Fn-III 10 with integrins (*4*), an aspartate in Inv497 D5 (Asp^{911}) is required for integrin binding (*7,20*). Like the aspartate in the Fn-III RGD sequence, Asp^{911} is located in a loop (Fig. 1A,3B). Other host proteins, such as VCAM-1 and MAdCAM-1, which bind integrins that lack I domains, also contain a critical aspartate residue on a protruding loop (*15*). By contrast, ligands of I domain-containing integrins, such as the ICAM proteins, present their acidic integrin-binding residue in the context of a β-strand rather than a loop (*15*). A second region of invasin approximately 100 amino acids from Asp^{911} contains additional residues that are implicated in integrin binding, including Asp^{811} (Fig. 1A, 3B) (*20*). This region of invasin is reminiscent of the fibronectin synergy region located in Fn-III 9 that is required for maximal $\alpha_5\beta_1$ integrin-dependent cell spreading (*21*). Invasin Asp^{811} is located in D4 between strands A'' and A''' and lies on the same surface as Asp^{911}, separated by 32 Å (measured between carbon-α atoms). The distance between Fn-III 10 Asp^{1495} in the RGD sequence and Fn-III 9 Asp^{1373} in the synergy region is also 32 Å (*14*), although the side

chain orientation of Asp^{1373} differs from that of Asp^{811} in invasin (Fig. 3). Within the Fn-III synergy region, a critical residue for integrin binding is Arg^{1379} (32 Å from Asp^{1495}; Fig. 3B) (*21*). The invasin synergy-like region also includes a nearby arginine (Arg^{883}; 32 Å from Asp^{911}; Fig. 3B). The overall similarity in the relative positions of these three residues suggests that invasin and host proteins share common integrin-binding features.

The transmembrane regions of outer membrane proteins of known structure are β-barrels, as represented by the structures of porins (*7*). Assuming that the membrane associated region of invasin is also a β-barrel (*7*), the structure of intact invasin may resemble the model shown in Fig. 1D, in which the cell binding region projects about 180 Å away from the bacterial surface, ideally positioned to contact host cell integrins. Similarities between invasin and fibronectin demonstrate convergent evolution of common integrin binding properties. However, the integrin binding surface of invasin does not include a cleft, as found on the binding surface of fibronectin (Fig. 3), thus invasin may bind integrins with a larger interface. Together with the restricted orientation of the invasin integrin binding domains, a larger binding interface provides a plausible explanation for the increased integrin binding affinity of invasin compared with fibronectin. Differences between the integrin binding properties of invasin and fibronectin illustrate how a bacterial pathogen is able to efficiently compete with host proteins to establish contact and subsequent infection, thereby exploiting a host receptor for its own purposes.

Table 1. Summary of data collection and refinement statistics for Inv497. Inv497 crystals (space group $P2_1$, a = 60.1 Å, b = 50.7 Å, c = 97.9 Å, β = 98.3°; one molecule per asymmetric unit) were grown at 22 °C in hanging drops by combining 1 μl protein solution (5 to 10 mg/ml Inv497, 20 mM Hepes pH 7.0, 1 mM EDTA) with 1 μl precipitant solution (20 mM sodium citrate pH 5.6, 20% polyethylene glycol 4K, 20% isopropanol). Crystals were improved by microseeding. SeMet crystals, derived from selenomethionine-substituted Inv497 protein (*9*), grew under similar conditions. For cryoprotection, 5 μl of mother liquor containing 25% isopropanol was added to the crystals immediately before transferring them to liquid nitrogen. A cryo-cooled xenon derivative was prepared by mounting a cryoprotected crystal in a nylon loop and subjecting it to 200 psi of xenon for 2.5 minutes in a xenon pressure cell (*11*). A small microfuge tube containing excess mother liquor was placed in the pressurization chamber to maintain vapor pressure and prevent cracking of the crystals. Immediately after depressurization, the crystals were transferred to liquid nitrogen. The PIP derivative was prepared by the addition of 1 grain of PIP to a drop containing several crystals followed by soaking for 5 hours. Data from the native and the xenon derivative crystals were collected at –170 °C using a wavelength of 0.98 Å on a MAR Research image plate detector at beam line 9-1 at the Stanford Synchrotron Radiation Laboratory. Data from the PIP and SeMet derivatives were collected at –170 °C on an RAXIS IIC image plate using a Rigaku rotating anode. Statistics in parentheses refer to the highest resolution bin. Phasing, model building, and refinement were done as described (*11*,*12*).

Data Set	Resolution (Å)	*Complete (%)	†R_{merge} (%)	I/σI	‡rms f_h/E
Native	2.30 (2.34-2.30)	96.3 (97.2)	5.1 (29.6)	17.9 (2.9)	
Xenon	2.75 (2.85-2.75)	90.4 (76.3)	4.6 (10.0)	17.1 (7.9)	1.2
PIP§	2.80 (2.90-2.80)	87.7 (63.8)	7.4 (23.8)	11.1 (2.8)	1.4
SeMet‖	3.00 (3.11-3.00)	93.1 (91.6)	15.9 (40.0)	7.4 (2.9)	1.6

Refinement Statistics:			
Resolution (Å)	30.0-2.3	Number of non-hydrogen atoms:	
Reflections in working set	24256	Protein	3593
Reflections in test set	1216	Water	92
R_{free} (%)¶	28.7	Citrate	13
R_{cryst} (%)#	23.5	Non-glycine residues in most favorable	
R.m.s. deviations from ideality		region of Ramachandran plot (%)	86.6
Bond lengths (Å)	0.008	as defined (*12*)	
Bond angles (deg)	1.51		

*Complete = (number of independent reflections)/total theoretical number. †R_{merge} (I)= ($\Sigma|I(i) - \langle I(h)\rangle|/\Sigma I(i)$), where I(i) is the *i*th observation of the intensity of the hkl reflection and <I> is the mean intensity from multiple measurements of the h,k,l reflection. ‡Rms f_h/E (phasing power), where f_h is the heavy-atom structure factor amplitude and E is the residual lack of closure error. §PIP: di-μ-iodobis(ethylenediamine)diplatinum nitrate. ‖SeMet: selenomethionine.

¶R_{cryst} (F) = $\Sigma_{\mathbf{h}}||\mathbf{F_{obs}}(\mathbf{h})| - |\mathbf{F_{calc}}(\mathbf{h})||/\Sigma_{\mathbf{h}}|\mathbf{F_{obs}}(\mathbf{h})|$, where $|\mathbf{F_{obs}}(\mathbf{h})|$ and $|\mathbf{F_{calc}}(\mathbf{h})|$ are the observed and calculated structure factor amplitudes for the h,k,l reflection. #R_{free} is calculated over reflections in a test set not included in atomic refinement (*12*).

Fig. 1. (A) Ribbon diagram of the structure of *Yersinia pseudotuberculosis* Inv497. Residues implicated in integrin binding (Asp^{911}, Asp^{811} (*7*,*20*) and possibly Arg^{883}; see text) are highlighted in green. α-helical regions in D5 and a 3_{10} helix in D4 are red. The disulfide bond in D5 is yellow. **(B)** Topology diagrams for domains of invasin and related proteins. Inv497 D5 is shown beside a canonical C-type lectin CRD (from E-selectin (*14*)); Inv497 D4 is shown beside a C1-type IgSF domain. β-strands are blue, helices are red, and disulfide bonds are yellow. The calcium binding loop in E-selectin (residues 54-89) and its truncated counterpart in Inv497 (residues 956-959) are green. **(C)** Left: Hydrogen bonding pattern of the interrupted helix (*18*) in D5. Main chain atoms are shown for residues in the α-helix. Side chains are shown for those residues in which main chain atoms form hydrogen bonds (dashed cyan lines) across the break in the helix. Other side chains have been omitted for clarity. The carbon-α trace of the loop is shown in gray. Right: The Inv497 model in the region of the loop (gray in left panel) of the interrupted helix superimposed on a 2.3 Å σ_A-weighted $2F_o$-F_c annealed omit electron density map contoured at 1.0 σ (map radius = 3.5 Å) (*12*). **(D)** Schematic model of the structure of intact invasin in which the ~500 N-terminal residues reside in the *Yersinia* outer membrane (OM; yellow) in a porin-like structure (*7*) (red) and the Inv497 portion of invasin (green and blue) projects ~180 Å from the outer membrane.

Fig. 2. Close-up comparison of interdomain interfaces in integrin binding regions of Inv497 (D4 and D5), fibronectin (Fn-III 9-10) (*14*) and VCAM-1 (D1 and D2) (*14*). Hydrogen bonds are shown as dashed yellow lines. Additional hydrogen bonds, van der

Waals contacts and a three- to five-fold larger interdomain surface area (*19*) stabilize Inv497 D4-D5 and restrict interdomain flexibility compared to the other interfaces.

Fig. 3. Comparison of integrin binding regions of invasin and fibronectin. Despite different folding topologies and surface structures, the relative positions of several residues implicated in interactions with integrins are similar (Asp^{811}, Asp^{911}, Arg^{883} in Inv497; Asp^{1373}, Asp^{1495}, Arg^{1379} in Fn-III 9-10; aspartates in red; arginines in blue). **(A)** Surface representations (*12*) of the structures of Inv497 and Fn-III 7-10 (*14*). **(B)** Ribbon representations of Inv497 D4-D5 and Fn-III 9-10. Addition of one or more residues to the COOH-terminus of invasin (indicated as "COO^-") interferes with integrin binding (*22*), suggesting that the rather flat region between Asp^{811} and Asp^{911} is at the integrin-binding interface. By contrast, the integrin-binding surface of fibronectin contains a cleft resulting from the narrow link between Fn-III 9 and Fn-III 10.

References and Notes.

1. J. C. Pepe, V. L. Miller, *Infectious Agents and Disease* **2**, 236 (1993); A. Marra, R. R. Isberg, *Infect. Immun.* **65**, 3412 (1997); I. B. Autenrieth, R. Firsching, *J. Med. Microbiol.* **44**, 285 (1996).

2. R. R. Isberg, D. L. Voorhis, S. Falkow, *Cell* **50**, 769 (1987).

3. R. R. Isberg, J. M. Leong, *Cell* **60**, 861 (1990); R. R. Isberg, G. Tran Van Nhieu, *Trends Microbiol.* **2**, 10 (1994).

4. R. O. Hynes, *Cell* **69**, 11 (1992).

5. G. Tran Van Nhieu, E. S. Krukonis, A. A. Reszka, A. F. Horwitz, R. R. Isberg, *J. Biol. Chem.* **271**, 7665 (1996); G. Tran Van Nhieu, R. R. Isberg, *EMBO J.* **12**, 1887 (1993).

6. H. Liu, L. Magoun, J. M. Leong, *Infect. Immun.* **67**, 2045 (1999); B. Kenny *et al*., *Cell* **91**, 511 (1997).

7. J. M. Leong, R. S. Fournier, R. R. Isberg, *EMBO J*. **9**, 1979 (1990); M. J. Worley, I. Stojiljkovic, F. Heffron, *Mol. Microbiol*. **29**, 1471 (1998). Using a neural network algorithm that predicts the membrane topology of integral outer membrane proteins [K. Diederichs, J. Freigang, S. Umhau, K. Zeth, and J. Breed, *Protein Sci*. **7**, 2413 (1998)], residues 142 to 494 are predicted to contain 22 β-strands. Thus the structure of this portion of invasin may resemble those of the membrane-spanning portions of porins (16-18 β-strands) [reviewed in B. K. Jap, P. J. Walian, *Physiol. Rev.* 76, 1073 (1996)] or FhuA (22 β-strands) [K. P. Locher *et al.*, *Cell* **95**, 771 (1998); A. D. Ferguson, E. Hofmann, J. W. Coulton, K. Diederichs, W. Welte, *Science* **282**, 2215 (1998)].

8. J. Yu, J. B. Kaber, *Mol. Microbiol*. 6, 411 (1992). A sequence identity of 30% has been established as the threshold for guaranteed three-dimensional similarity [C. Chothia, A. M. Lesk, *EMBO J.* 5, 823-826 (1986)]. Length-dependent sequence identity thresholds are discussed in R. A. Abagyan, S. Batalov, *J. Mol. Biol*. 273, 355-368 (1997) and references therein. By these criteria, the cell binding regions of invasin and intimin do not share significant sequence identity, and individual invasin domains do not share significant sequence similarity with Fn-III, IgSF, CRD, CTLD or CTLD-related proteins.

9. Protein expression: The C-terminal 497 amino acid region of invasin was produced in *E. coli* fused to maltose binding protein, cleaved by Factor Xa, and purified as previously described [J. M. Leong, P. E. Morrissey, A. Marra, R. R. Isberg, *EMBO J.* **14**, 422 (1995)]. N-terminal sequencing indicated that the first residue of the cleaved protein was Ser^{495}, five residues shorter than the predicted site of cleavage (R. R. Isberg, unpublished data). A selenomethionine-substituted version of Inv497 was produced following the method of W. A. Hendrickson, J. R. Horton, D. M. LeMaster, *EMBO J*. **9**, 1665 (1990), and purified under the same conditions as the native protein. Amino acid composition analysis showed ~100% replacement of the eight methionines by selenomethionine (data not shown).

10. G. Tran Van Nhieu, R. R. Isberg, *J. Biol. Chem.* **266**, 24367 (1991); E. S. Krukonis, P. Dersch, J. A. Eble, R. R. Isberg, *J. Biol. Chem.* **273**, 31837 (1998); J. A. Eble *et al*., *Biochem.* **37**, 10945 (1998); Y. Takada, J. Ylanne, D. Mandelman, W. Puzon, M. H. Ginsberg, *J. Cell Biol.* **119**, 913 (1992).

11. Phasing and model building: A cryo-cooled xenon derivative was prepared using the apparatus described by S. M. Soltis, M. H. B. Stowell, M. C. Wiener, G. N. Phillips,

D. C. Rees, *J. Appl. Crystallogr.* **30**, 190 (1997). Data were processed and scaled with the HKL package [Z. Otwinowski, W. Minor, *Meth. Enzymol.* **276**, 307 (1997)]. Heavy-atom refinement and phasing were performed with the program SHARP [E. De La Fortelle, G. Bricogne, *Meth. Enzymol.* **276**, 472 (1997]. Difference Patterson maps for the xenon and PIP derivatives were interpreted using XTALVIEW [D. E. McRee, *Practical Protein Crystallography*. Academic Press Inc., San Diego, CA. (1993)] and one xenon, three platinum and two iodine sites were refined using SHARP. An initial MIRAS electron density map was calculated to 3.6 Å and solvent flattened using Solomon [J. P. Abrahams, A. G. W. Leslie, *Acta Crystallogr.* **D52**, 30 (1996)] as implemented in SHARP. A skeleton of the map [G. J. Kleywegt, T. A. Jones, *Acta Crystallogr.* **D52**, 826 (1997)] served as a starting point for model building using the program O [T. A. Jones, M. Kjeldgaard, *Meth. Enzymol.* **277**, 173 (1997)]. The initial electron density map revealed the Ig-like domain structures of the first four domains, but definitive assignment of side chains was not possible and the connectivity in D5 was ambiguous. Using MIRAS phases, eight selenium sites were found in a difference Fourier map calculated for the SeMet derivative, which allowed identification of methionines that were used as markers for the assignment of the rest of the sequence. After including the SeMet sites in heavy atom positional refinement using SHARP, an improved solvent flattened MIRAS electron density map was calculated to 3.2 Å resolution with a mean figure of merit of 0.509.

12. Refinement and structure analysis: After rigid body refinement of the five domains of Inv497, refinement was carried out using the simulated annealing and energy minimization protocols in the program CNS [A. T. Brünger *et al.*, *Acta Crystallogr. D* **54**,

905 (1998))] using bulk solvent and anisotropic temperature-factor corrections (B11=9.730, B22=-20.203, B33=10.473, B12=0.000, B13=-3.975, B23=0.000) and protocols that minimized R_{free} [A. T. Brünger, *Nature* **355**, 472 (1992)]. In each round of model building, a combination of σ_A weighted [R. J. Read, *Acta Crystallogr. A* **48**, 851 (1992)] 2|Fo|-|Fc| and |Fo|-|Fc| maps (calculated with model phases combined with experimental phases or model phases alone) and simulated annealing omit maps [A. Hodel, S.-H. Kim, A. T. Brünger, *Acta Crystallogr*. *A* **48**, 851 (1992)] were used. In later rounds of refinement, water molecules were built into peaks greater than 3σ in |Fo|-|Fc| maps. The eight N-terminal residues were not visible in electron density maps, thus the final model includes residues 503 to 986 of the Inv497 construct (*9*), one citrate and 92 water molecules, with an overall B factor of 41.4 $Å^2$. Several regions include residues with real space correlation values [T. A. Jones, M. Kjeldgaard, *Meth. Enzymol*. **277**, 173 (1997)] below 1σ from the mean (residues 531-534, 582-586, 647-650, 676-679, 779-780, 892-899, 955-957, 969-977). Ramachandran plot statistics (Table 1) are as defined by PROCHECK [R. A. Laskowski, M. W. McArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr*. **26**, 283 (1993)]. Figures were made using MOLSCRIPT [P. J. Kraulis, *J. Appl. Crystallogr*. **24**, 946 (1991)], RASTER-3D [E. A. Merritt, M. E. P. Murphy, *Acta Crystallogr*. **D. 50**, 869 (1994)]. Molecular surfaces were generated using GRASP [A. Nicholls, R. Bharadwaj, B. Honig, *Biophys. J*. **64**, A166 (1993)].

13. Equilibrium analytical ultracentrifugation analyses establish that Inv497 is monomeric at μM concentrations in solution [P. Dersch, R. R. Isberg, *EMBO J*. 18, 1199 (1999)]. Sedimentation velocity analytical ultracentrifugation experiments suggest that

Inv497 is elongated in solution [Su *et al*., *Science* **281**, 991 (1998)], thus the extended conformation does not result from crystal packing forces.

14. Fn-III 7-10 [Protein Data Bank (PDB) code 1FNF], D. J. Leahy, I. Aukhil, H. P. Erickson, *Cell* **84**, 155 (1996); Fn-III 12-14 (PDB code 1FNH), A. Sharma, J. A. Askari, M. J. Humphries, E. Y. Jones, D. I. Stuart, *EMBO J.* **18**, 1468 (1999); intimin (coordinates obtained from S. Matthews), G. Kelly *et al*., *Nat. Struct. Biol.* **6**, 313 (1999); mannose binding protein (PDB code 1RTM), W. I. Weis, K. Drickamer, *Structure* **2**, 1227 (1994); E-selectin (PDB code 1ESL), B. J. Graves *et al*., *Nature* **367**, 532 (1994); CD94 (coordinates obtained from P. D. Sun), J. C. Boyington *et al.*, *Immunity* **10**, 75 (1999); VCAM-1 (PDB code 1VSC), E. Y. Jones *et al.*, *Nature* **373**, 539 (1995); J.-h. Wang *et al*., *Proc. Natl. Acad. Sci. USA* **92**, 5714 (1995); ICAM-1 (PDB code 1IC1), J. M. Casasnovas, T. Stehle, J.-h. Liu, J.-h. Wang, T. A. Springer, *Proc. Natl. Acad. Sci. USA* **95**, 4134 (1998); J. Bella, P. R. Kolatkar, C. W. Marlor, J. M. Greve, M. G. Rossmann, *Proc. Natl. Acad. Sci. USA* **95**, 4140 (1998).

15. IgSF domains were previously classified into V, C1, C2, I1 and I2 sets based on similarities in sequence and structure [Y. Harpaz, C. Chothia, *J. Mol. Biol.* 238, 528 (1994); J. M. Casasnovas, T. Stehle, J.-h. Liu, J.-h. Wang, T. A. Springer, *Proc. Natl. Acad. Sci. USA* 95, 4134 (1998)]. The V and C1 sets are similar to antibody variable and constant domains, respectively. The V set consists of two β-sheets: one containing β-strands ABED and the other containing strands A'GFCC'C". The C1 set contains an ABED and a GFC sheet. The two sheets of the C2 set are ABE and GFCC'. I set domains are intermediate between the V and C1 sets. The I1 set contains ABED and A'GFC sheets, while the I2 set contains ABE and A'GFCC' sheets. D1-D4 of Inv497

adopt folding topologies that resemble IgSF domains, but lack the core residues and disulfide bonds conserved in IgSF members.

16. The CTLDs in natural killer cell receptors share many features of the C-type lectin CRD fold, but differ substantially from canonical C-type lectin domains in their ligand binding characteristics since they lack most of the calcium coordinating residues that are critical for carbohydrate recognition in CRDs [W. I. Weis, M. E. Taylor, K. Drickamer, *Immunol. Rev.* **163**, 19 (1998)]. Ly49A, a natural killer cell receptor, recognizes a carbohydrate-independent epitope on its class I MHC ligand [N. Matsumoto, R. K. Ribaudo, J. P. Abastado, D. H. Marguilies, W. M. Yokoyama, *Immunity* **8**, 245 (1998)]. Other proteins such as the TSG-6 Link module (*14*), intimin D3 (*14*), and invasin D5 are not related by obvious sequence similarity to CTLDs and CRDs (for example, they do not contain the characteristic "WIGL" sequence in β-strand 2 or the "inner" disulfide bond linking strand 3 to the loop following strand 4; Fig. 1B), but share a similar folding topology. The TSG-6 Link module lacks the calcium binding loops present in CRDs, but is believed to bind hyaluronan using an exposed patch of hydrophobic and charged residues (*14*). The CTLD-related domain of intimin contains an analogous patch of residues (Lys^{227}, Tyr^{228}, Tyr^{230} and Tyr^{231}) (*14*). The corresponding residues of Inv497 are not conserved with intimin (Thr^{931}, Leu^{932}, Gly^{934}, Glu^{935}).

17. Although direct binding of invasin to a carbohydrate has not been demonstrated, high concentrations of N-acetylneuraminic acid (IC_{50} = 20 mM) inhibit mammalian cell adhesion to immobilized invasin. A variety of other acetylated sugars showed no such inhibition (R. R. Isberg, unpublished results).

18. Interrupted α-helices have been observed in other protein structures, including subtilisin [reviewed in J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1981)] and fibritin [Y. Tao, S. V. Strelkov, V. V. Mesyanzhinov, M. G. Rossmann, *Structure* **5**, 789 (1997)].

19. Surface areas buried between domains in the Inv497, Fn-III 7-10, Fn-III 12-14, and VCAM-1 structures (*14*) were calculated using XPLOR (*12*) with a 1.4 Å probe radius. Inv497 D1-D2: 411 $Å^2$; Inv497 D2-D3: 454 $Å^2$; Inv497 D3-D4: 564 $Å^2$; Inv497 D4-D5: 1925 $Å^2$; Fn-III 7-8: 608 $Å^2$; Fn-III 8-9: 481 $Å^2$; Fn-III 9-10: 342 $Å^2$; Fn-III 12-13: 450 $Å^2$; Fn-III 13-14: 696 $Å^2$; VCAM-1 D1-D2 (molecule B) 696 $Å^2$.

20. J. M. Leong, P. E. Morrissey, R. R. Isberg, *J. Biol. Chem.* **268**, 20524 (1993); L. H. Saltman, Y. Lu, E. M. Zaharias, R. R. Isberg, *J. Biol. Chem.* **271**, 23438 (1996).

21. Bowditch *et al*., *J. Biol. Chem.* **269**, 10856 (1994); S.-I. Aota, M. Nomizu, K. M. Yamada, *J. Biol. Chem.* **269**, 24756 (1994); T. P. Ugarova *et al*., *Biochem.* **34**, 4457 (1995).

22. R. R. Isberg, Y. Yang, D. L. Voorhis, *J. Biol. Chem.* **268**, 15840 (1993).

23. We thank S. M. Soltis and the staff at SSRL for help with xenon derivatization and data collection, M. J. Bennett, A. J. Chirino, L. M. Sánchez, D. E. Vaughn, and A. P. Yeh for discussions and help with crystallographic software, S. Matthews for intimin coordinates, P. D. Sun for CD94 coordinates, W. I. Weis for helpful discussions about C-type lectin structures, and W. I Weis, John M. Leong and members of the Bjorkman lab for critical reading of the manuscript. Inv497 coordinates have been deposited in the PDB (PDB code 1CWV).

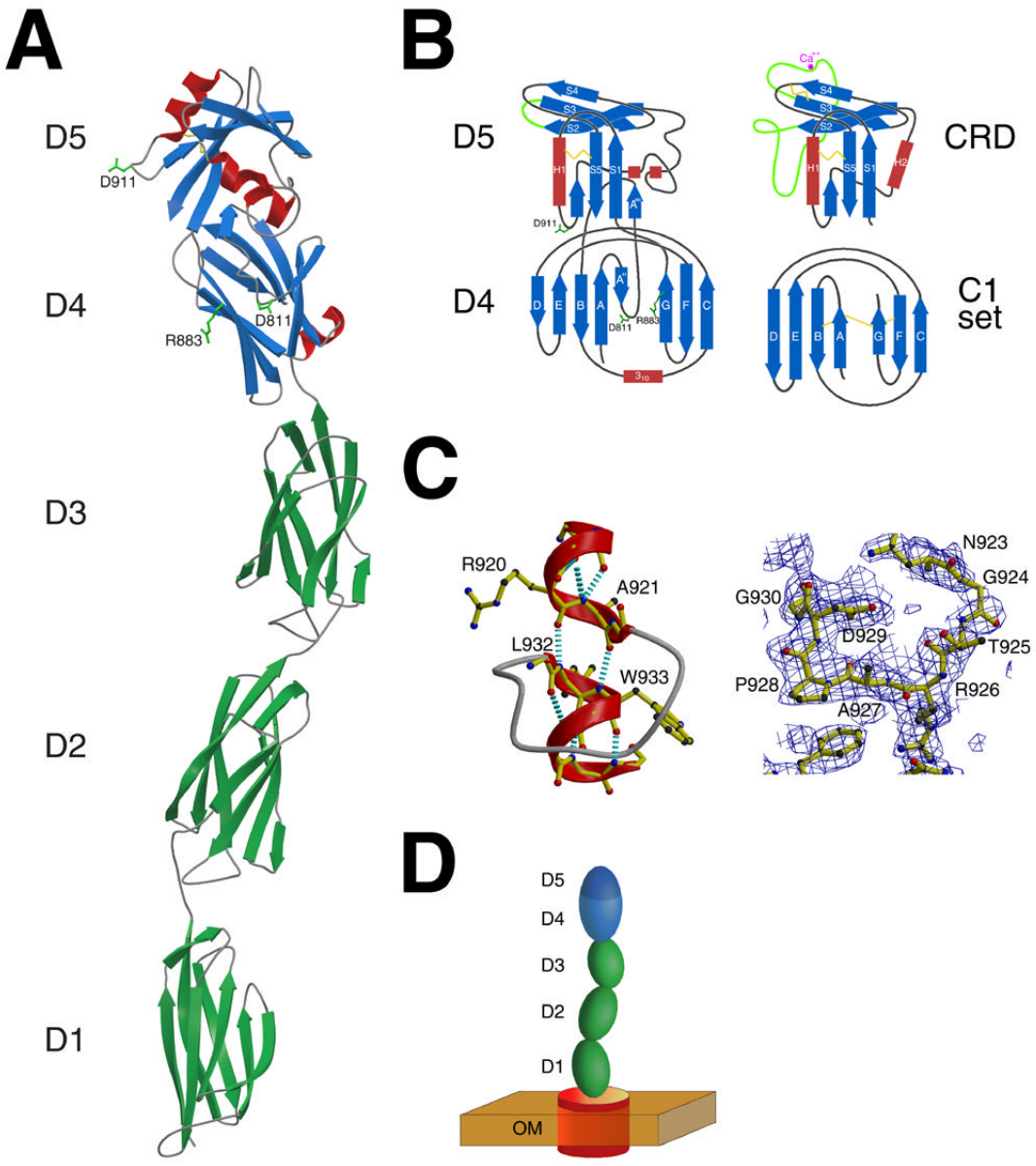


Figure 1. Hamburger et al.

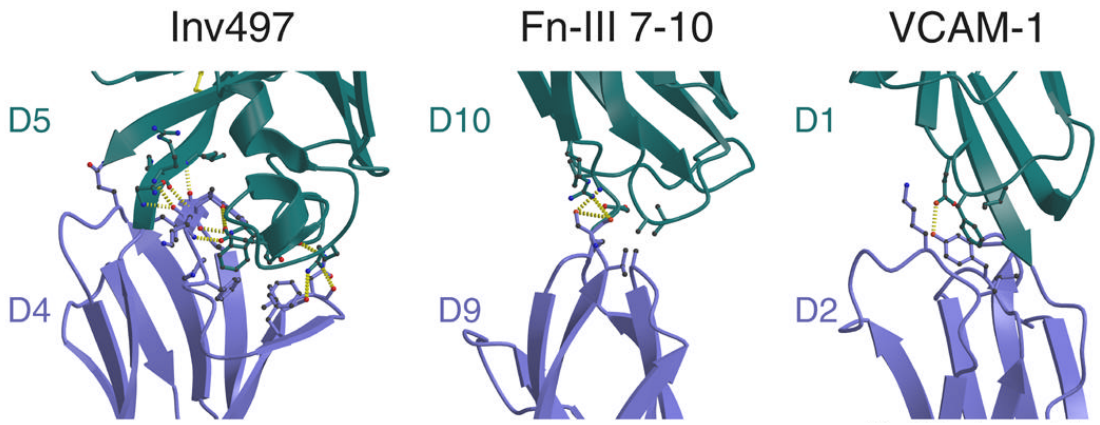


Figure 2. Hamburger et al.

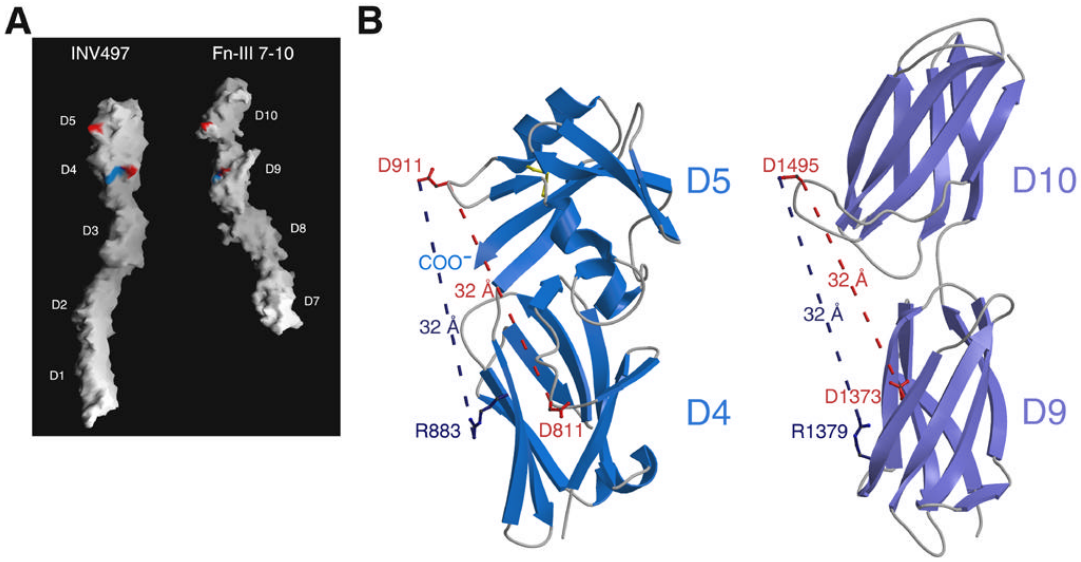


Figure 3. Hamburger et al.

Chapter 3:

Refolding and Biochemical Characterization of the Transmembrane Region of Invasin

Abstract

The N-terminal ~500 amino acids of invasin has been shown to be required for membrane localization and for presentation of the integrin-binding region of invasin on the cell surface [1]. We have initiated structural analyses in order to gain insight into the function of this region of invasin. We overexpressed different invasin constructs in the outer membranes of *E. coli* cells. Our results show that expression of the invasin transmembrane constructs occurs at low levels in the outer membrane and purification of these protein products is very inefficient. On the other hand, expression of the transmembrane region of invasin as inclusion bodies in the cytoplasm of *E. coli* cells resulted in high levels of protein expression. This region of the protein can be refolded in the presence of detergents and we have obtained microcrystals of this membrane protein. Circular dichroism studies indicate that this region of invasin is composed mainly of β structure, as expected since the transmembrane regions of outer membrane proteins of known structure are β-barrels. Proteolysis experiments suggest that the N-terminal 70 amino acids of invasin may form a separate, probably periplasmic domain analogous to that found in another outer membrane protein, the sucrose-specific porin ScrY. Equilibrium sedimentation analytical ultracentrifugation studies indicate the invasin transmembrane region is monomeric in solution. Black lipid bilayer experiments suggest that this region of the protein does not contain a pore, suggesting a role as an outer membrane anchor for the presentation of the integrin-binding domain on the cell surface.

Introduction

Invasin is a 986-residue outer membrane protein expressed on the surface of the Gram-negative enteropathogenic bacterium *Y. pseudotuberculosis.* As described in Chapters 1 and 2, the C-terminal 497 residues of invasin comprise the extracellular region of the protein and promote bacterial uptake by binding to a number of different β1 integrins on host cells [2]. Less is known about the structure and the function of the N-terminal ~500 amino acids of invasin. This region of the protein has been shown to play a role in membrane localization and presentation of the integrin-binding region of invasin on the cell surface [1]. Removal of the N-terminal 148 amino acids of the protein prevents invasin expression on the bacterial surface. However, when the periplasmic maltose binding protein (MBP) is fused to the construct lacking residues 1-148, the fusion protein is exported to the outer membrane and bacteria are able to bind and enter host cells [1]. This suggests that MBP, which normally does not cross the outer membrane, mediates export of the fusion protein to the periplasm, and it is the C-terminal 838 residues of invasin that direct localization to the outer membrane [1]. Invasin residues 1-148 presumably carry a Sec-dependent signal sequence that allows export through the inner membrane. Cleavage of the signal peptide is predicted to occur after Ala49 using the program SignalP [3].

The N-terminal transmembrane domain of invasin is not homologous to any proteins of known structure. However, it has significant sequence identity (~36%) to the corresponding region of another family of enteropathogenic bacterial proteins related to invasin, the intimins. Intimins mediate attachment of bacteria to host cells by binding to a bacterially secreted protein Tir, which upon secretion becomes inserted into the host

membrane [4, 5]. Interestingly, exchanging the N-terminal 489 amino acids of invasin for the corresponding N-terminal 539 amino acids of intimin results in proper membrane localization and presentation of invasin epitopes on the bacterial cell surface [6].

The large size of invasin (106 kDa) is reminiscent of other secreted proteins from Gram-negative bacteria including IgA protease. These proteins belong to the autotransporter family and contain a C-terminal ~500 residue region predicted to form an amphipathic pore in the outer membrane, which mediates the secretion of the N-terminal protease domain (Figure 1) [7-9]. Analogously, the N-terminal ~500 amino acids of invasin have been predicted to form an amphipathic channel that translocates the C-terminal integrin binding domain of invasin across the outer membrane [1]. This predicted functional similarity is despite the fact that invasin and the IgA protease share no significant sequence similarity and the relative orientations of the pore domain and secreted domains are reversed. Alternatively, the N-terminal transmembrane region of invasin may play a more passive role by acting as an anchor that attaches the extracellular region to the outer membrane with other proteins involved in the secretion of the C-terminal half of invasin across the outer membrane.

To address the function of the invasin transmembrane region, we initiated a structural and biochemical characterization of the protein in purified form. Like other integral outer membrane proteins (reviewed in [10], the N-terminal ~500 amino acids of invasin likely have a β-barrel fold [1]. However, determining the exact three-dimensional structure of the membrane-spanning domain of invasin will provide information about the function of this region of the protein. Two distinct architectures may be envisioned for the putative β-barrel of invasin, which has its N-terminus located in the periplasm due to

the Sec-dependent mechanism for protein export, and its C-terminus located extracellularly. The β-barrel may have an odd number of strands, or alternatively an even number of strands that forms a pore through which the C-terminal region can be threaded for extracellular localization. The presence of an odd number of β-strands would suggest that the role of this domain is that of an anchor that tethers the integrin-binding region to the cell surface. If invasin contains an odd number of strands, it will be the first time this architecture is seen in the outer membrane since the structures of all known β-barrels are composed of an even number of antiparallel β-strands. If the invasin transmembrane domain contains an even number of strands, it is probable that it will also contain a pore for translocation of the extracellular region across the outer membrane. In this case the structure would suggest that this transmembrane domain plays the role of an autotransporter.

To obtain the milligram amounts of protein necessary for crystallization, we overexpressed eight different invasin constructs in the outer membranes of *E. coli* cells. We have found that expression of the invasin transmembrane constructs occurs at low levels and purification of these protein products is very inefficient. On the other hand, expression of the transmembrane domain of invasin as inclusion bodies in the cytoplasm of *E. coli* cells results in high levels of protein production. Refolding was accomplished in the presence of detergents and microcrystals of the protein have recently been obtained. We also biophysically characterized the transmembrane domain of invasin using analytical ultracentrifugation, circular dichroism, limited proteolysis, and black lipid bilayer experiments.

Materials and Methods

Expression vectors

The vector pRI203, encoding full length invasin, was kindly provided by Ralph Isberg at Tufts University. Fourteen different constructs were generated using PCR and are designated cell surface constructs: A, B, C, D, IBA-2-D1, IBA-6-D1, IBA-2-F1, IBA-6-F1 and inclusion body constructs: △ssInv, △ssInv-NHis, △ssInv-NHis-D1, △ssInv-NHis-Full, △ssInv-NHis-127, △ssInv-NHis-149. See also Figure 2.

Cell surface constructs

Protein expression and purification BL21△Omp cells were transformed with the IBA-6-D1 construct and grown in 1 L Terrific Broth to an $O.D._{600}$ of 0.5 in the presence of 50 mg/L kanamycin and 200 mg/L ampicillin. BL21(DE3) cells were transformed with the IBA-2-F1 construct and grown in 1 L Terrific Broth to an $O.D._{600}$ of 0.5 in the presence of 200 mg/L ampicillin. Both the BL21(DE3) and the BL21△Omp cells were induced for 2 hours with 200 μg/L anhydrotetracycline (Sigma Genosys). The cells were pelleted at 6000 rpm in an Avanti JLA 8.1 rotor and resuspended in 50 mM Tris pH 8, 1 mg/ml lysozyme and 50 ug/ml DNAse I. Cells were lysed using a French press. Inclusion bodies and unbroken cells were pelleted in a tabletop centrifuge at 3000 rpm at 4 °C for 15 min. To pellet both the outer and inner membranes, the cleared lysate was spun in a Ti60 rotor at 40, 000 rpm for 30 min at 4 °C.

Purification of the IBA-6-D1 construct Isolated membranes were resuspended in 50 mM Tris pH 8 in the presence of 2% Triton X-100 and 6M Urea and incubated at room temperature for 30 min. This incubation dissolves the inner membrane, leaving the

outer membrane intact [11]. The outer membranes were pelleted at 40, 000 rpm for 30 min at 4 °C in a Ti60 rotor. To remove the Triton X-100 and Urea, the outer membrane pellet was washed by incubation at room temperature for 30 min in 25 mM sodium phosphate pH 7.6 and 0.15 % n-Octylpolyoxyethylene, followed by centrifugation as above. In order to extract proteins from the purified outer membranes, the sample was resuspended in 25 mM sodium phosphate, pH 7.6 and 1 % n-Octylpolyoxyethylene and incubated for 1 hour at room temperature. The final supernatant contained extracted proteins from the outer membrane.

Purification of the IBA-2-F1 construct The membranes, isolated as described above, were resuspended in 3 ml 20 mM Tris pH 8 and 3 % n-Octylpolyoxyethylene and incubated for 1 h at room temperature. The extracted membranes were pelleted in a tabletop centrifuge at 4 °C for 40 min at 14,000 rpm. The lysate, containing the extracted outer membrane proteins was loaded onto a StrepTactin column (Sigma Genosys) and purified according to the manufacturer's instructions.

Inclusion Body Constructs

Protein expression and refolding BL21(DE3) cells transformed with the △ssInv-NHis construct were grown to an $O.D._{600}$ of ~1.1 in Terrific Broth and expression was induced with 1 mM IPTG for 2 hours. The cells were pelleted at 6000 rpm in an Avanti JLA 8.1 rotor and lysed in 10 mM Tris pH 8, 0.5% Triton X-100, 3 mM $CaCl_2$,1 mg/ml lysozyme and 50 ug/ml DNAse I using a hand held Omni TH homogenizer. Inclusion bodies were pelleted at 10,000 rpm at 4 °C for 30 min in a Beckman SS-34 rotor. Inclusion bodies were washed twice with 10 mM Tris pH 8 in the presence of 0.5% Triton X-100 and twice in 10 mM Tris pH 8. The purified inclusion bodies were

denatured in 8 M Urea and added to 200 ml of Ni-NTA resin at a concentration of 1 mg/ml. The sample was incubated at RT for 30 minutes with gentle agitation and poured into a Pharmacia XK50/20 column. Refolding of the denatured protein was achieved by applying a linear gradient on the Ni-NTA column in the presence of 1% n-Octylpolyoxyethylene, 50 mM Tris pH 8, 200 mM NaCl from 8 M to 0 M Urea over 6 column volumes at a flow rate of 5 ml/min. The refolded protein was eluted from the column with 250 mM imidazole. The eluted sample (~100 ml) was passed over a Hi-Prep 26/10 desalting column in 0.5% n- Octylpolyoxyethylene, 50 mM Tris pH 8, 200 mM NaCl in order to remove excess imidazole. In order to concentrate the protein, the sample was subsequently loaded onto an ~20 ml Ni-NTA column. After elution, the protein was subjected to gel filtration on a Hi-Prep 26/60 Sephacryl S-300 sizing column (10 ml each) in 0.5% n-Octylpolyoxyethylene, 50 mM Tris pH 8, 150 mM NaCl.

Analytical ultracentrifugation Equilibrium ultracentrifugation studies were done with a Beckman-Coulter XL-I Ultima analytical ultracentrifuge equipped with absorption optics. △ssInv-NHis samples were in 150 mM NaCl, 50 mM Tris pH 8, 0.5% n-Octylpolyoxyethylene. Sedimentation equilibrium experiments were conducted at 16,000, 21,000 and 28,000 rpm at 20 °C. Protein concentrations were 4.6 μM, 1.5 μM and 0.8 μM. Data files were analyzed as described in [12]. The effects of the detergent were also studied at a protein concentration of 3 μM comparing buffers containing 0.5% n-Octylpolyoxyethylene and 1.25% n-Octylpolyoxyethylene.

Circular dichroism CD spectroscopy was performed at 20 °C with an Aviv 62A DS spectrometer. Protein concentrations were determined by tryptophan and tyrosine absorbance at 280 nm using ε=69840 M-1cm-1 (calculated from [13]) using the amino

acid sequence of △ssInv-NHis. Protein concentration was 1mg/ml in 10 mM Tris pH 8, 150 mM NaCl and 0.5% n-Octylpolyoxyethylene. Three wavelength scans (250 to 190 nm) were collected with a 1 sec averaging time per data point.

Reconstitution of △ssInv-NHis into vesicles In order to reconstitute the △ssInv-NHis protein into lipid vesicles, 43.6 mg of Soybean lectin phosphatidylcholine (Sigma) and 22.6 mg of cholic acid (Sigma) were dissolved by water-bath sonication in 950 µl of 50 mM Tris pH 8 and 200 ml NaCl (as recommended by Amika Corp). 0.5 µg of △ssInv-NHis solubilized in 0.5% n-Octylpolyoxyethylene was added to the lipid mixture. Alternatively, vesicles were formed by dissolving soybean lectin phosphatidylcholine and 3% n-Octylpolyoxyethylene in 950 µl of 50 mM Tris pH 8 and 200 ml NaCl by water-bath sonication. 25 µg or 250 µg of △ssInv-NHis was added to the lipid mixture. Detergent was removed to allow vesicle formation by dialyzing the samples against 500 ml of 50 mM Tris pH 8 and 200 ml NaCl overnight at RT.

Black lipid bilayer experiments The black lipid bilayer experiments were carried out essentially as described [14]. The bilayer apparatus was assembled with the CV-5B-100GU headstage, small chamber with a 250 µm hole between the chambers (polysulphone cuvette), stir bars, isolated Ag/AgCl electrodes in 3 M KCl, and salt bridges to both the cis and trans chambers. The cis chamber served as a virtual ground. Prior to bilayer formation each well was filled with 1 mL of 200 mM KCl, 10 mM HEPES pH 7.2, 40 mM $MgCl_2$. When the protein was reconstituted into lipid vesicles,16 µM $CaCl_2$ was added to each chamber to facilitate fusion of the vesicles to the bilayer.

The bilayer forming solution was prepared by mixing 56 µL of 25 mg/ml POPE in $CHCl_3$ and 24 µL of 25 mg/ml POPS in $CHCL_3$ in a V-shaped vial. The solvent was

evaporated and the resulting lipid film dissolved in 100 µl of decane. A bilayer was painted over the 250 µm hole separating the two chambers using a glass rod. The bilayer capacitance was determined using a triangular wavefunction (pClamp 8). Bilayers were smoothed with an air bubble if needed. When the bilayer broke, it was either reformed using the glass rod or an air bubble. Bilayers exhibiting a capacitance of > 85 pF were used for the experiments.

Protease digestion 7 µg of △ssInv-NHis and 0.5 µg of chymotrypsin were mixed and incubated at 37 °C in 100 mM Tris pH8, 15 mM $CaCl_2$, 0.5% n-Octylpolyoxyethylene. After 10 min, 30 min and 60 min an aliquot was removed and the reaction was stopped with 1% trifluoroacetic acid. The last sample was subjected to SDS-PAGE and blotted to a PVDF membrane for N-terminal sequencing.

Protein crystallization The purified, refolded protein was concentrated at 15 °C in a Millipore BioMax concentrator (50 kDa cut-off) to 10-30 mg/ml. The concentrated sample was dialyzed against different detergents in the presence of 10 mM Tris pH 8 and 150 mM NaCl. Crystallization trials were set up using the sitting drop vapor diffusion method.

Results and discussion

Cell surface expression To produce large amounts of the transmembrane region of *Yersinia pseudotuberculosis* invasin protein for crystallization, eight different constructs (construct A, B, C, D, IBA-2-D1, IBA-6-D1, IBA-2-F1 and IBA-6-F1) were designed for expression in the outer membrane of *E. coli* cells (Figure 2A). These constructs contain the putative transmembrane region of invasin (residues 50-502) as well as a signal sequence that allows the protein to be exported across the inner membrane [15]. Upon translocation across the inner membrane, the signal sequence is cleaved off from the exported protein. In the case of invasin, the exact residue with which the invasin protein begins is not known because expression of invasin in cells results in the synthesis of the full-length protein as well as many smaller species [16]. Met18 and Met20 are two candidates as the DNA sequence preceding these methionines contains a potential Shine-Delgarno sequence. Using the program SignalP, the cleavage site for the signal sequence is predicted to be after residue 49 [3]). Thus it is likely that the signal sequence of invasin is composed of residues 18-49. To ensure export, constructs A, IBA-2-D1, IBA-6-D1, IBA-2-F1 and IBA-6-F1 were designed to contain the signal sequence of OmpA. Construct B contains the OmpT signal sequence while constructs C and D contain the native invasin signal sequence (Figure 2A).

The presence of soluble regions in membrane proteins have been shown to facilitate membrane protein crystallization as these regions are able to form important lattice contacts in the crystals [17, 18]. In order to increase our chances of obtaining protein crystals of the invasin transmembrane region, the IBA-2-D1, IBA-6-D1, IBA-2-F1 and IBA-6-F1 constructs were designed to express not only the transmembrane region

but also part or the entire extracellular region of invasin. The IBA-2-D1 and IBA-6-D1 constructs contain, in addition to the membrane bound region, the first extracellular domain of invasin as determined by x-ray crystallography (see Chapter 2). The IBA-2-F1 and IBA-6-F1 constructs contain both the membrane-bound as well as the full extracellular region of invasin (Figure 2A).

For cell surface expression, two *E. coli* strains (BL21(DE3) and BL21△Omp) were used. BL21△Omp cells are derivatives of the BL21 cells lines and have their lamB (maltoporin), OmpA, and OmpC (osmoporin) genes deleted. In addition, the OmpF gene is inactivated by the presence of a transposon conferring kanamycin resistance [19]. BL21△Omp cells have been successfully used to express a number of outer membrane proteins including OmpF and OmpC. The lack of the most abundant outer membrane proteins in this strain is advantageous for two reasons. First, it results in higher expression levels of the target protein because there is physically more room in the outer membrane. Second, it facilitates protein purification due to the lack of outer membrane proteins that would contaminate the extracted protein prep. A potential problem of the BL21△Omp strain is that these cells are fragile as they are missing many of the proteins that are important for maintaining the structural stability of the outer membrane [20].

In our experiments, BL21△Omp cells transformed with constructs B, C, IBA-2-F1 or IBA-6-F1 were not able to grow in liquid culture. Although cells transformed with construct D were able to grow, protein expression was not observed when the cells were induced under a number of different conditions. Only cells transformed with constructs A, IBA-2-D1 and IBA-6-D1 had detectable protein expression levels (Figures 2 and 3). Since BL21△Omp cells have outer membranes that are less stable, expression of the

transmembrane region of invasin on the surface of these cells, due to leakage at the promoters of the vectors, could have resulted in lysis of the cells. Therefore, the more sturdy BL21(DE3) cells were transformed with the eight invasin constructs. Once again cells electroporated with construct C did not grow under a number of growth conditions. Although cells transformed with constructs B and D did grow, protein expression was not observed under a number of different induction conditions (Figure 2A). However, cells expressing constructs A, IBA-2-D1, IBA-6-D1, IBA-2-F1 and IBA-6-F1 were able to grow and express the transmembrane region of invasin (Figure 3).

Purification of invasin from the cell surface Because degradation products for constructs A and IBA-2-D1 as well as very low protein expression levels for IBA-6-F1 were observed (Figure 3), only constructs IBA-6-D1 and IBA-2-F1 were used for protein purification. BL21△Omp cells were transformed with the IBA-6-D1 construct and induced with anhydrotetracycline (Sigma Genosys). Cells were lysed and the outer membranes were isolated as described [11]. Proteins from the outer membranes were extracted with n-Octylpolyoxyethylene. In order to determine if IBA-6-D1 (predicted MW 63 kDa) was expressed in the outer membrane and to monitor our purification, each step of the purification was analyzed by SDS-PAGE analysis (Figure 4). Lanes 1 and 2 show that the IBA-6-D1 protein was expressed in the BL21△Omp cells. Some of the protein had expressed as inclusion bodies (lane 3) and none of the protein was found in the cytoplasm (lane 4). Surprisingly, much of the protein was found in the inner membrane and/or periplasmic fraction (lane 5). This could indicate that the protein had not correctly inserted into the outer membrane such that during solubilization of the inner membranes using Triton X-100/Urea [11] much of the IBA-6-D1 was extracted. The

overall effect was that only minimal amounts of the protein could be extracted from the outer membrane by extraction with n-Octylpolyoxyethylene (lane 8) and that very little protein remained in the final outer membrane pellet (lane 9).

To try to increase protein yields, we decided to utilize the *Strep*-tag II that was expressed on the transmembrane constructs. The membrane fraction of the BL21△Omp cells expressing the IBA-2-F1 construct was treated with n-Octylpolyoxyethylene and the supernatant containing the extracted membrane proteins was loaded onto a StrepTactin column (Sigma Genosys). Figure 5 shows the SDS-PAGE analysis of the purification of the IBA-2-F1 protein (predicted MW 104 kDa) on the StrepTactin column. Lanes 1 and 2 show the extracted outer membrane proteins and the fraction of the sample that did not bind to the StrepTactin column (flow through), respectively. The wash fractions are shown in lanes 3-7, and the eluted protein samples are shown in lanes 8-13. Once again only low levels of protein were purified. Also, the purification was not efficient as many contaminating bands, as well as potential degradation products of IBA-2-F1, eluted from the column (lanes 10-12). As many milligrams of protein are needed for crystallization and only low levels of protein was obtained from the work described above, large scale expression and further purification of these constructs was not pursued.

Refolding of the transmembrane region of invasin from inclusion bodies Previous work has shown that outer membrane proteins can be expressed at very high levels as insoluble aggregates (inclusion bodies) in the cytoplasm of *E. coli* cells, and subsequently refolded to their native state in the presence of detergents [10, 18, 21, 22]. We therefore designed a number of different constructs, all lacking the predicted invasin signal

sequence (residues 1-49), for expression as inclusion bodies (Figure 2B). The inclusion bodies were isolated from BL21 cells, washed and dissolved in 8 M Urea.

After preliminary studies, large scale refolding was attempted by binding the denatured protein to either DEAE or Ni-NTA column materials. The △ssInv construct was refolded on a 200 ml DEAE column by removal of the urea in the presence of 1% n-Octylpolyoxyethylene. The sample was eluted from the column with 200 mM NaCl in the presence of 1% n-Octylpolyoxyethylene (data not shown). To assess the yield of refolded protein, the sample was run on an S-200 Sephacryl sizing column (Amersham Pharmacia). Figure 6 shows the results from two different refolding experiments. Even though some of the protein eluted in the soluble fraction of the sizing column, suggesting successful refolding, the experiments were not reproducible and yielded different ratios of the refolded △ssInv species. This is likely due to the fact that when △ssInv is applied to the DEAE column under denaturing conditions, any part of the protein backbone may be immobilized by interactions with the beads. Certain regions of △ssInv may be important in the correct folding of the protein and immobilization of these regions may interfere with the refolding process.

To promote correct folding, new constructs (△ssInv-NHis, △ssInv-NHis-D1, △ssInv-NHis-Full) containing N-terminal 6 x His tag were designed and expressed as inclusion bodies (Figure 2B). The N-terminal tag allows tethering of only the N-terminus of the protein to Ni-NTA beads (Amersham Pharmacia) leaving the rest of the molecule free for correct folding. The △ssInv-NHis-D1 and △ssInv-NHis-Full constructs contain, in addition to the transmembrane region of invasin, the first extracellular domain or the full extracellular domain of invasin, respectively. The three inclusion body samples were

bound to Ni-NTA columns and refolded in the presence of 1% n-Octylpolyoxyethylene. The samples were eluted, concentrated and run on a Sephacryl S-300 sizing column in the presence of 0.5% n-Octylpolyoxyethylene. Refolding of the △ssInv-NHis -Full construct failed and the refolding efficiency of the △ssInv-NHis-D1 construct was very low (data not shown). However, refolding of the △ssInv-NHis protein was successful as it was reproducible and had 10% efficiency. This sample eluted from the sizing column with an apparent molecular weight of ~ 100 kDa (predicted MW of △ssInv-NHis is 53 kDa) (Figure 7).

Analytical ultracentrifugation To determine the oligomeric state of the transmembrane region of invasin, we used equilibrium sedimentation analytical ultracentrifugation in the presence of 0.5% n-Octylpolyoxyethylene. The use of n-Octylpolyoxyethylene allows determination of molecular masses at gravitational transparency because this detergent has the same density as water (ρ =1 g/cm^3) [23]. The results of the analytical ultracentrifugation suggested that the △ssInv-NHis protein was monodisperse with a molecular weight of 50.3 kDa, indicating that the refolded protein was monomeric in solution (Figure 8). To study if the detergent had any effect on the equilibrium sedimentation analysis, the sample was also tested in the presence of 1.25% n-Octylpolyoxyethylene. We found that the detergent had no effect since the molecular weight was still fitted to be 50 kDa (data not shown). Thus, the apparent molecular mass of 100 kDa of the refolded protein on the sizing column is likely due to the presence of detergent micelles or an elongated shape of the molecule. Integrin receptor clustering as well as high affinity integrin binding are required for bacteria to be internalized by host cells [24, 25]. Domain 2 of the invasin extracellular region has been shown to play a role

in homo-oligomerization of invasin [24]. Even at millimolar protein concentrations in the invasin crystals we did not observe multimerization of invasin (Chapter 2). We also do not find multimerization of the N-terminal transmembrane region of invasin when this portion of the protein is purified in the presence of n-Octylpolyoxyethylene.

Circular dichroism To obtain information about the secondary structure of the transmembrane region of invasin, we analyzed the refolded △ssInv-NHis protein in the presence of 0.5% n-Octylpentaoxyethylene by circular dichroism (CD) (Figure 9). The CD spectrum of △ssInv-NHis is indicative of β-sheet secondary structure with a minimum at ~215 nm indicating that the transmembrane region of invasin is composed mostly of β-structure. The data below 200 nm was not interpretable due to the presence of the detergent. The transmembrane regions of outer membrane proteins of known structure are β-barrels (reviewed in [10]. Based on this and the CD spectrum of △ssInv-NHis, we believe that this portion of the protein also forms such a structure.

Black lipid bilayer experiments To determine whether the transmembrane region of invasin forms a pore across the outer membrane, we performed black lipid bilayer experiments to study the permeability properties of △ssInv-NHis. Black lipid bilayer experiments require the reconstitution of the target membrane protein into a lipid bilayer that is suspended between two chambers containing salt solutions. If the protein contains a pore, the flow of ions can be measured between the two chambers. Different methods can be used to reconstitute membrane proteins into the bilayer. In the case of porins, detergent-solubilized protein (1-100 ng/ml) can be directly added to the chamber solution since porins are able to self-insert into the membrane [26]. Another method uses

membrane proteins reconstituted into lipid vesicles which then fuse with the lipid bilayer (reviewed in [27].

For initial experiments, detergent-solubilized △ssInv-NHis was added to one of the two chambers surrounding the lipid bilayer. Conductance was not observed when 5 ng, 50 ng nor 5 μg of protein was added to the bilayer (data not shown). The absence of conductance is an ambiguous result. It could suggest that the protein does not contain a pore or that the △ssInv-NHis did not insert into the bilayer. To facilitate insertion into the bilayer, △ssInv-NHis was reconstituted into phosphatidylcholine vesicles and the vesicles were added to the bilayer. Once again we did not observe conductance across the lipid bilayer either when the protein to lipid ratio was 1:200, 1:2000 nor 1:100,000 suggesting that △ssInv-NHis does not form a pore. As with the previous experiment, the negative result can also be attributed to the lack of protein insertion into the membrane.

Crystallization trials for △ssInv-NHis Refolding of the △ssInv-NHis construct on the Ni-NTA column was reproducible and yielded milligrams of protein necessary for crystallization. Crystallization of membrane proteins is often more difficult than that of soluble proteins [28, 29]. One of the key factors influencing membrane protein crystallization are detergent properties, such as micelle size and dynamics, and it is often necessary to screen a number of different detergents before crystals are obtained [30]. n-Octylpolyoxyethylene, the detergent used for the refolding of △ssInv-NHis, has a high critical micelle concentration (CMC) and can be dialyzed away from the protein with relative ease, allowing different detergents to be screened for crystallization. Table 1A shows the detergents, screens and temperatures tested for crystallization of △ssInv-NHis.

Although a large number of conditions were tested, reproducible crystals were not obtained.

Proteolysis of △ssInv-NHis Proteolytic cleavage of the △ssInv-NHis construct was observed after the sample was stored at 4 °C for several weeks. Based on SDS-PAGE analysis, the loss of approximately 10 kDa was observed (Figure 10A). Size exclusion chromatography showed that the proteolysed sample was still monodisperse and monomeric (Figure 10A). N-terminal sequencing of the proteolysed sample detected three cleavage sites starting with residues Ala123, Ser128, Val139, respectively.

Limited proteolysis can be used to determine the boundaries of separate domains present in proteins. Chymotrypsin digestion of △ssInv-NHis resulted in approximately 100% cleavage of the protein after Leu121, as determined by N-terminal sequencing (Figure 10B). This suggested that the N-terminal ~70 amino acids forms a separate domain from the rest of the △ssInv-NHis protein, reminiscent of the periplasmically located N-terminal 70 residues of the sucrose-specific porin ScrY. In ScrY, this region is thought to affect the interaction with the peptidoglycan layer, because its deletion results in a much tighter interaction between the protein and the murein [31].

Previous work by Leong et al. has shown that replacement of the N-terminal 148 amino acids of invasin with MBP results in expression of invasin on the cell surface, suggesting that this region of the protein is not required for membrane insertion [1]. The presence of flexible domains can inhibit crystal formation since they can interfere with crystal packing. Based on the findings by Leong et al. as well as on the results of the proteolysis, two new inclusion body constructs were designed and produced in *E. coli* BL21(DE3) cells to increase the chances of obtaining crystals (Figure 2B). The △ssInv-

NHis-127 construct begins with residue Ala127 and the △ssInv-NHis-149 construct begins with residue Arg149. Both of these constructs expressed at approximately the same level as △ssInv-NHis, as judged by SDS-PAGE and Western blot analysis using an anti-His tag antibody (data not shown). Both constructs also refolded under the same conditions and at the same efficiency as △ssInv-NHis (Figure 11). A number of detergents were screened for crystallization of both constructs (Table 1B and C). Using the PEG/Ion screen (Hampton Research), several conditions were identified that resulted in microcrystal formation. Figure 12 shows protein crystals of the △ssInv-NHis-127 grown in PEG and Ammonium Chloride. The crystals are ~50 X 10 X 2 microns in length. Although no diffraction could be observed in initial tests at the home x-ray source, optimization of growth conditions is in progress.

Acknowledgments

I would like to thank Ralph Isberg for kindly providing the pRI203 vector; Kaspar Locher for guidance and helpful discussions regarding membrane protein expression and refolding; Adam Politzer for help with cloning and protein expression; Andrew Herr for help with analytical ultracentrifugation, and Joshua Maurer for help with the black lipid bilayer experiments.

References:

1. Leong, J.M., Fournier, R.S., and Isberg, R.R. Identification of the integrin binding domain of the *Yersinia pseudotuberculosis* invasin protein. EMBO J. **9** (1990) 1979-89.
2. Isberg, R.R. and Leong, J.M. Multiple beta 1 chain integrins are receptors for invasin, a protein that promotes bacterial penetration into mammalian cells. Cell. **60** (1990) 861-71.
3. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. **10** (1997) 1-6.
4. Liu, H., Magoun, L., and Leong, J.M. beta 1 chain integrins are not essential for intimin-mediated host cell attachment and enteropathogenic *Escherichia coli*-induced actin condensation. Infect. Immun. **67** (1999) 2045-2049.
5. Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A., and Finlay, B.B. Enteropathogenic *Escherichia coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell. **91** (1997) 511-520.
6. Liu, H., Magoun, L., Luperchio, S., Schauer, D.B., and Leong, J.M. The Tir-binding region of enterohaemorrhagic *Escherichia coli* intimin is sufficient to trigger actin condensation after bacterial-induced host cell signalling. Mol. Microbiol. **34** (1999) 67-81.
7. Mackman, N., Nicaud, J.M., Gray, L., and Holland, I.B. Identification of polypeptides required for the export of haemolysin 2001 from *Escherichia coli*. Mol. Gen. Genet. **201** (1985) 529-36.

8. Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T.F. Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature. **325** (1987) 458-62.

9. Klauser, T., Pohlner, J., and Meyer, T.F. The secretion pathway of IgA protease-type proteins in gram-negative bacteria. Bioessays. **15** (1993) 799-805.

10. Buchanan, S.K. Beta-barrel proteins from bacterial outer membranes: structure, function and refolding. Curr. Opin. Struct. Biol. **9** (1999) 455-61.

11. Schnaitman, C.A. Outer membrane proteins of *Escherichia coli*. I. Effect of preparative conditions on the migration of protein in polyacrylamide gels. Arch. Biochem. Biophys. **157** (1973) 541-52.

12. West, A.P., Jr., Giannetti, A.M., Herr, A.B., Bennett, M.J., Nangiana, J.S., Pierce, J.R., Weiner, L.P., Snow, P.M., and Bjorkman, P.J. Mutational analysis of the transferrin receptor reveals overlapping HFE and transferrin binding sites. J. Mol. Biol. **313** (2001) 385-97.

13. Gill, S.C. and von Hippel, P.H. Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182** (1989) 319-326.

14. Schindler, H. Planar lipid-protein membranes: strategies of formation and of detecting dependencies of ion transport functions on membrane conditions. Methods Enzymol. **171** (1989) 225-53.

15. von Heijne, G. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173** (1984) 243-51.

16. Isberg, R.R. and Leong, J.M. Cultured mammalian cells attach to the invasin protein of *Yersinia pseudotuberculosis*. Proc. Natl. Acad. Sci. USA. **85** (1988) 6682-6.

17. Ostermeier, C., Iwata, S., Ludwig, B., and Michel, H. Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. Nat. Struct. Biol. **2** (1995) 842-6.

18. Pautsch, A., Vogt, J., Model, K., Siebold, C., and Schulz, G.E. Strategy for membrane protein crystallization exemplified with OmpA and OmpX. Proteins. **34** (1999) 167-72.

19. Prilipov, A., Phale, P.S., Van Gelder, P., Rosenbusch, J.P., and Koebnik, R. Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from *Escherichia coli*. FEMS Microbiol. Lett. **163** (1998) 65-72.

20. Sonntag, I., Schwarz, H., Hirota, Y., and Henning, U. Cell envelope and shape of *Escherichia coli*: multiple mutants missing the outer membrane lipoprotein and other major outer membrane proteins. J. Bacteriol. **136** (1978) 280-5.

21. Schmid, B., Kromer, M., and Schulz, G.E. Expression of porin from *Rhodopseudomonas blastica* in *Escherichia coli* inclusion bodies and folding into exact native structure. FEBS Lett. **381** (1996) 111-4.

22. Prince, S.M., Feron, C., Janssens, D., Lobet, Y., Achtman, M., Kusecek, B., Bullough, P.A., and Derrick, J.P. Expression, refolding and crystallization of the OpcA invasin from *Neisseria meningitidis*. Acta Crystallogr. D. Biol. Crystallogr. **57** (2001) 1164-6.

23. Fleming, K.G., Ackerman, A.L., and Engelman, D.M. The effect of point mutations on the free energy of transmembrane alpha-helix dimerization. J. Mol. Biol. **272** (1997) 266-75.

24. Dersch, P. and Isberg, R.R. An immunoglobulin superfamily-like domain unique to the *Yersinia pseudotuberculosis* invasin protein is required for stimulation of bacterial uptake via integrin receptors. Infect. Immun. **68** (2000) 2930-8.

25. Tran Van Nhieu, G. and Isberg, R.R. Bacterial internalization mediated by beta 1 chain integrins is determined by ligand affinity and receptor density. EMBO J. **12** (1993) 1887-95.

26. Benz, R., Janko, K., Boos, W., and Lauger, P. Formation of large, ion-permeable membrane channels by the matrix protein (porin) of *Escherichia coli*. Biochim. Biophys. Acta. **511** (1978) 305-19.

27. Benz, R. and Bauer, K. Permeation of hydrophilic molecules through the outer membrane of gram-negative bacteria. Review on bacterial porins. Eur. J. Biochem. **176** (1988) 1-19.

28. Garavito, R.M., Picot, D., and Loll, P.J. Strategies for crystallizing membrane proteins. J. Bioenerg. Biomembr. **28** (1996) 13-27.

29. Sowadski, J.M. Introduction: crystallization of membrane proteins--in need of a new focus? J Bioenerg. Biomembr. **28** (1996) 3-5.

30. Sowadski, J.M. Crystallization of membrane proteins. Curr. Opin. Struct. Biol. **4** (1994) 761-764.

31. Schulein, K., Andersen, C., and Benz, R. The deletion of 70 amino acids near the N-terminal end of the sucrose- specific porin ScrY causes its functional similarity to LamB in vivo and in vitro. Mol. Microbiol. **17** (1995) 757-67.

Table 1. Crystallization trials for the refolded invasin constructs

The △ssInv-NHis, △ssInv-NHis-127 and △ssInv-NHis-149 constructs were refolded in the presence of n-Octylpolyoxyethylene and dialyzed against a number of different detergents to be screened for crystallization. The table shows the different crystallization conditions that were tested for crystallizing these invasin transmembrane constructs.

Table 1. Crystallization trials for the refolded invasin constructs

A. △ssInv-NHis

Detergent	Concentration	Hampton Screens	Temp
n-Octyl-β-D-glucoside	0.9%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT
n-Octyl-β-D-glucoside	1.8%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT 4C
2-Hydroxyethyloctylsulfoxide	0.6%	Crystal Screen I, Crystal Screen 2, MembFac	RT
n-Octylpentaoxyethylene	0.5%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT
n-Octylpentaoxyethylene	1%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	4C
N,N-Dimethyldecylamine-N-oxide	0.8%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT
n-decyl-ß-D-maltoside	0.2%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT 4C

B. △ssInv-NHis-127

Detergent	Concentration	Hampton Screens	Temp
n-Octyl-β-D-glucoside	1.8%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT
n-Octylpentaoxyethylene	1%	Crystal Screen I, Crystal Screen 2, MembFac, PEG/Ion	RT
N,N-Dimethyldecylamine-N-oxide	0.8%	Crystal Screen I, Crystal Screen 2, PEG/Ion	RT
n-decyl-ß-D-maltoside	0.1%	Crystal Screen I, Crystal Screen 2, PEG/Ion	RT
Cymal-5	0.1%	Crystal Screen I, Crystal Screen 2, PEG/Ion	RT
Fos-choline 12	0.05%	Crystal Screen I, Crystal Screen 2, PEG/Ion	RT

C. △ssInv-NHis-149

Detergent	Concentration	Hampton Screens	Temp
n-decyl-ß-D-maltoside	0.1%	Crystal Screen I, Crystal Screen 2, PEG/Ion	RT

Figure 1. Model of surface display by autotransporters in Gram-negative bacteria

The nascent peptide chain of the autotransporter protein is translocated across the inner membrane (IM) in a Sec-dependent manner. After the signal sequence is cleaved from the precursor protein, the C-terminal transporter inserts into the outer membrane (OM) forming a translocation pore through which the N-terminal extracellular domain can pass. The signal peptide is shown in green, the C-terminal transport domain in red and the N-terminal extracellular domain in blue. Adapted from Klauser et al., 1993.

Figure 2. The invasin transmembrane constructs

A. For expression on the surface of BL21(DE3) and BL21△Omp cells, the expression constructs A, B, C, D, IBA-2-D1, IBA-6-D1, IBA-2-F1 and IBA-6-F1 were designed. The invasin signal sequence (SS) is shown in bright pink, while the transmembrane domain (TM) of invasin is shown in purple and gray. Domain 1 (D1) of the extracellular region of invasin is shown in light green while the rest of the extracellular region (D2-D5) is shown in dark green and blue. Degradation products observed during expression are designated by "deg."

B. For expression as inclusion bodies, the expression constructs △ssInv, △ssInv-NHis, △ssInv-NHis-D1, △ssInv-NHis-Full, △ssInv-NHis-127 and △ssInv-NHis-149 were designed. The putative periplasmic domain (PD), identified by proteolysis, is shown in purple while the invasin transmembrane domain is shown in gray. Domain 1 of the extracellular region of invasin is shown in light green while the rest of the extracellular domain is shown in dark green and blue.

Figure 3. Western blot analysis to check for expression of the transmembrane region of invasin on the surface of BL21△Omp and BL21(DE3) cells

Expression of the transmembrane region of invasin was analyzed by Western blotting using an anti *Strep*-tag II antibody. Cells expressing the invasin transmembrane region were separated by SDS-PAGE and electrotransferred to a nitrocellulose membrane. BL21△Omp cells transformed with construct A after induction with anhydrotetracycline (lane 1); BL21△Omp cells transformed with IBA-2-D1 (lanes 2 and 3) and IBA-6-D1 (lanes 4 and 5) before and after induction, respectively; BL21(DE3) cells transformed with IBA-2-F1 (lanes 6 and 7) and IBA-6-F1 (lanes 8 and 9) before and after induction, respectively.

Figure 4. SDS-PAGE analysis of the purification of IBA-6-D1

Each step of the purification of the IBA-6-D1 construct was analyzed by SDS-PAGE. Lanes 1 and 2 show induced cells and lysed cells expressing the IBA-6-D1 protein, respectively. The predicted molecular weight of the IBA-6-D1 construct is 63 kDa. After cell lysis, a low speed spin was performed to pellet unbroken cells and inclusion bodies (lane 3). The cytoplasmic fraction is shown in lane 4. The inner membrane is dissolved in a buffer containing Triton X-100 and urea (lane 5). To wash away any remaining Triton X-100 and urea, the outer membranes are washed in 0.15% n-Octylpolyoxyethylene (lane 6). Proteins are extracted from the purified outer membranes with 3% n-Octylpolyoxyethylene (lanes 7 and 8). The final pellet contains extracted outer membranes (lane 9).

Figure 5. SDS-PAGE analysis of the purification of IBA-2-F1

Purification of the IBA-2-F1 construct from BL21△Omp cells was analyzed by SDS-PAGE. The predicted molecular weight of IBA-2-F1 is 104 kDa. IBA-2-F1 extracted with 3% n-Octylpolyoxyethylene (lane 1) was applied to a StrepTactin column. Lane 2 shows the flow through sample and lanes 3-7 show the wash fractions. The eluted IBA-2-F1 protein is shown in lanes 8-13.

Figure 6. Size exclusion chromatography of the refolded △ssInv

The refolded △ssInv protein was applied to a Sephacryl S-200 gel-filtration column in the presence of 1% n-Octylpolyoxyethylene. Two separate refolding experiments are shown.

Figure 7. Size exclusion chromatography and SDS-PAGE analysis of △ssInv-NHis

The refolded △ssInv-NHis protein was applied to a Sephacryl S-300 gel-filtration column in the presence of 0.5% n-Octylpolyoxyethylene. Molecular weight standards are shown. The recovered peak was concentrated and analyzed by SDS-PAGE (lane 1).

Figure 8. Equilibrium sedimentation analytical centrifugation of △ssInv-NHis

Equilibrium sedimentation studies for △ssInv-NHis were performed at three different concentrations and three different speeds in the presence of 0.5% n-Octylpolyoxyethylene. The resulting nine curves were globally fitted to determine the molecular weight of △ssInv-NHis. Three representative curves at 4.6 μM concentration

at 16,000 (blue diamonds), 21,000 (green circles) and 28,000 (red squares) rpm are shown. The residuals for the fit are also shown.

Figure 9. Circular dichroism spectrum for △ssInv-NHis

The secondary structure of △ssInv-NHis was analyzed by circular dichroism in the presence of 0.5% n-Octylpolyoxyethylene. The spectrum of a wavelength scan at 20 °C is shown. Each data point represents the average of two measurements with the signal for the buffer subtracted.

Figure 10. Proteolysis of △ssInv-NHis

A. Storage of the refolded △ssInv-NHis protein at 4 °C resulted in proteolysis of the sample with an apparent loss of 10 kDa as observed by SDS-PAGE analysis. Size exclusion chromatography of the protein along with Blue Dextran (peak V) demonstrated that the proteolysed sample (peak I) was monodisperse and monomeric. Molecular weight standards are shown.

B. △ssInv-NHis was digested by chymotrypsin for 10, 30 or 60 min. The samples were analyzed by SDS-PAGE. Undigested protein (lanes 1 and 2); protein samples incubated at 37 °C in the presence (lanes 5, 8, 11) or absence (lanes 4, 7, 10) of chymotrypsin; chymotrypsin alone in the absence of △ssInv-NHis (lanes 3, 6, 9).

Figure 11. Size exclusion chromatography of △ssInv-NHis, △ssInv-NHis-127 and △ssInv-NHis-149

The refolded △ssInv-NHis, △ssInv-NHis-127 and △ssInv-NHis-149 proteins were applied to a Sephacryl S-300 gel-filtration column in the presence of 0.5% n-Octylpolyoxyethylene. Molecular weight standards are shown.

Figure 12. △ssInv-NHis-127 crystals

Crystals of △ssInv-NHis-127 were grown by the vapor diffusion method and photographed in the drop.

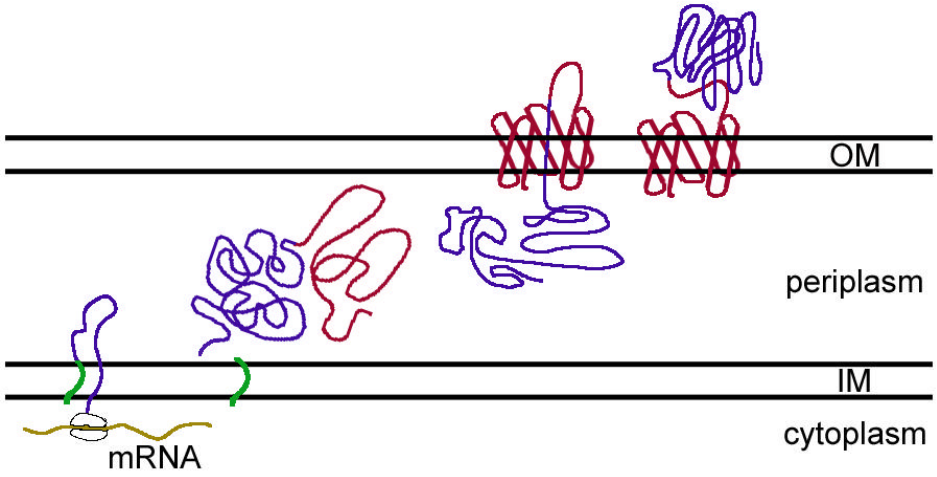


Figure 1. Model of surface display by autotransporters in Gram-negative bacteria

A

construct		invasin residues	vector	growth?		expression?
				BL21 (DE3)	BL21 △Omp	
A	ompA-ss – strep tag	S50-Q502	IBA-2	+	+	+ (deg)
B	ompT-ss	S50-Q502	pET12b	+	-	-
C		M20-Q502	pET23a	-	-	-
D	his tag	M18-Q502	pET21d	+	+	-
IBA-2-D1	ompA-ss – strep tag	S50-K595	IBA-2	+	+	+ (deg)
IBA-6-D1	ompA-ss – strep tag	S50-K595	IBA-6	+	+	+
IBA-2-F1	ompA-ss – strep tag	S50-I986	IBA-2	+	-	+
IBA-6-F1	ompA-ss – strep tag	S50-I986	IBA-6	+	-	low

B

construct		invasin residues	vector	cell line	refold?
△ssInv		S50-Q502	pET21c	BL21(DE3)	+
△ssInv-NHis	his tag	S50-Q502	pET28b	BL21(DE3)	+
△ssInv-NHis-D1	his tag	S50-K595	pET28b	BL21pLysS	low
△ssInv-NHis-Full	his tag	S50-I986	pET28b	BL21pLysS	-
△ssInv-NHis-127	his tag	A127-Q502	pET28b	BL21(DE3)	+
△ssInv-NHis-149	his tag	R149-Q502	pET28b	BL21(DE3)	+

Figure 2. The invasin transmembrane constructs

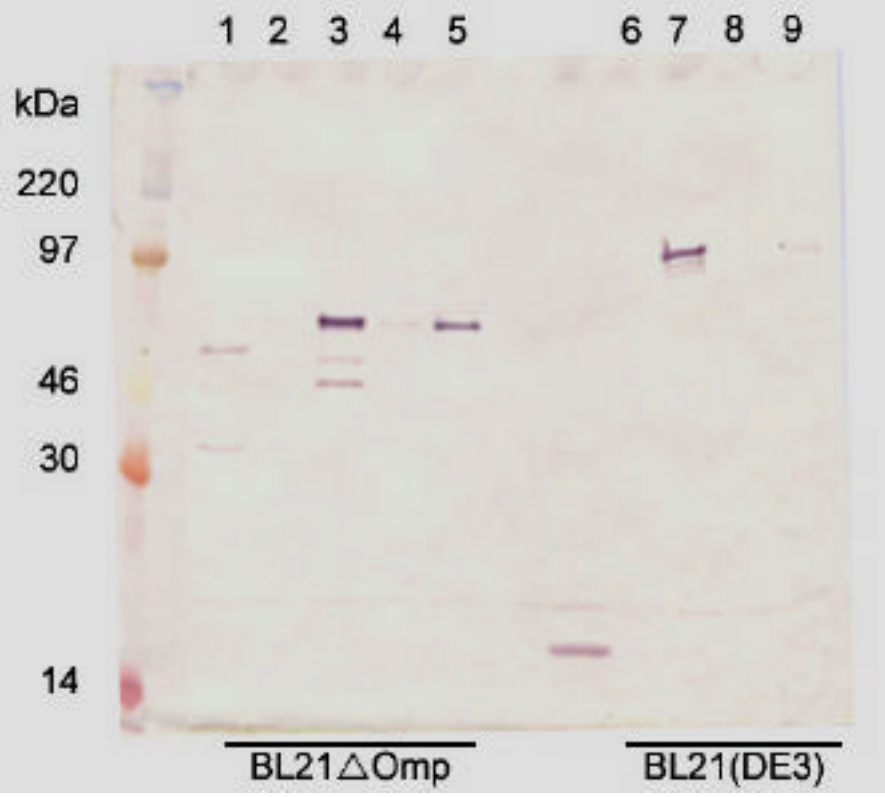


Figure 3. Western blot analysis to check for expression of the transmembrane region of invasin on the surface of BL21△Omp and BL21(DE3) cells

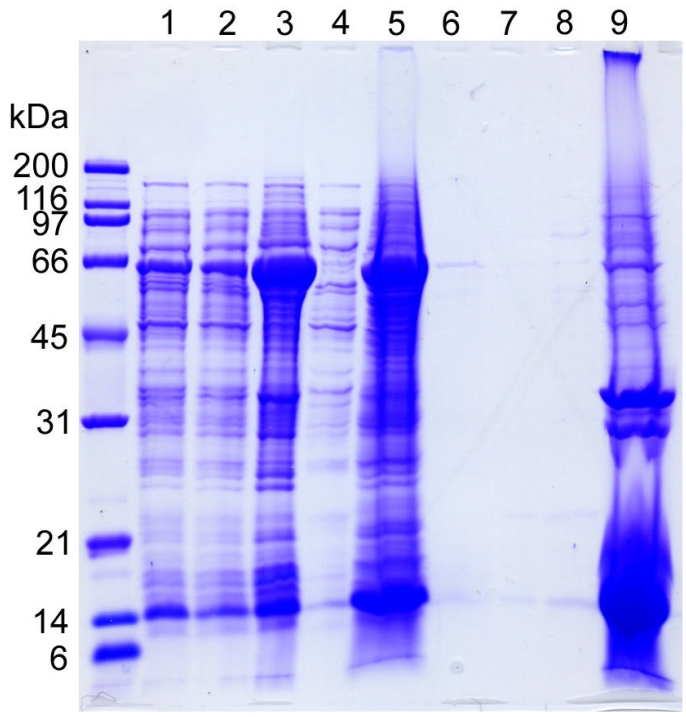


Figure 4. SDS-PAGE analysis of the purification of IBA-6-D1

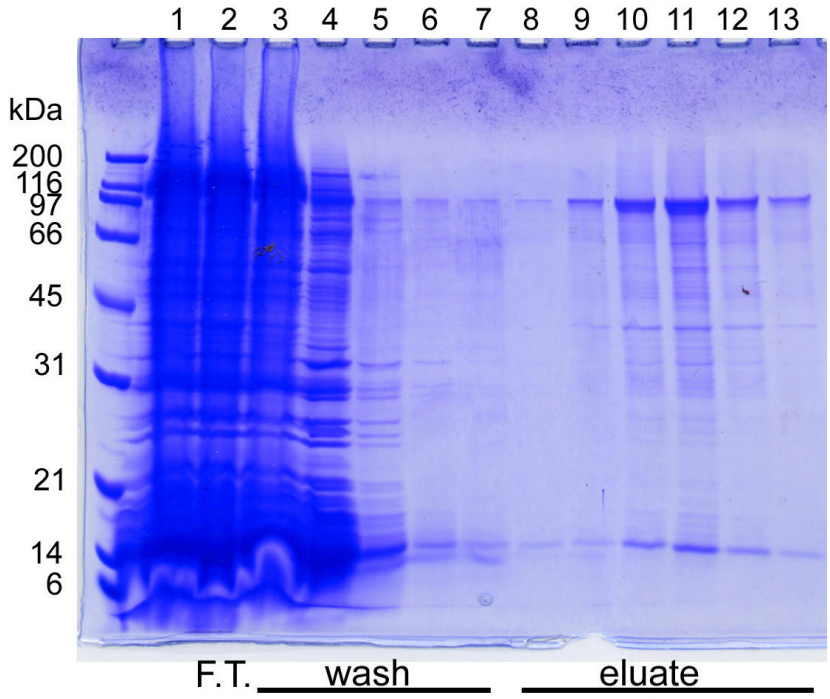


Figure 5. SDS-PAGE analysis of the purification of IBA-2-F1

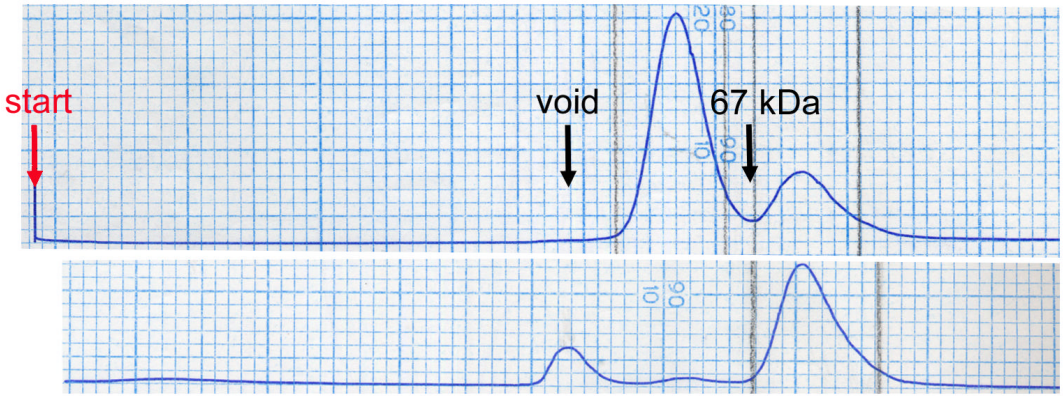


Figure 6. Size exclusion chromatography of the refolded △ssInv

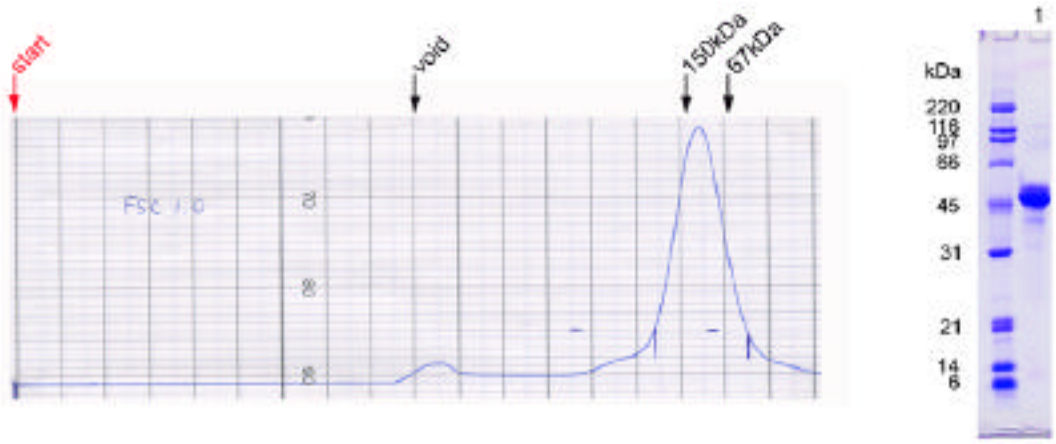


Figure 7. Size exclusion chromatography and SDS-PAGE analysis of △ssInv-NHis

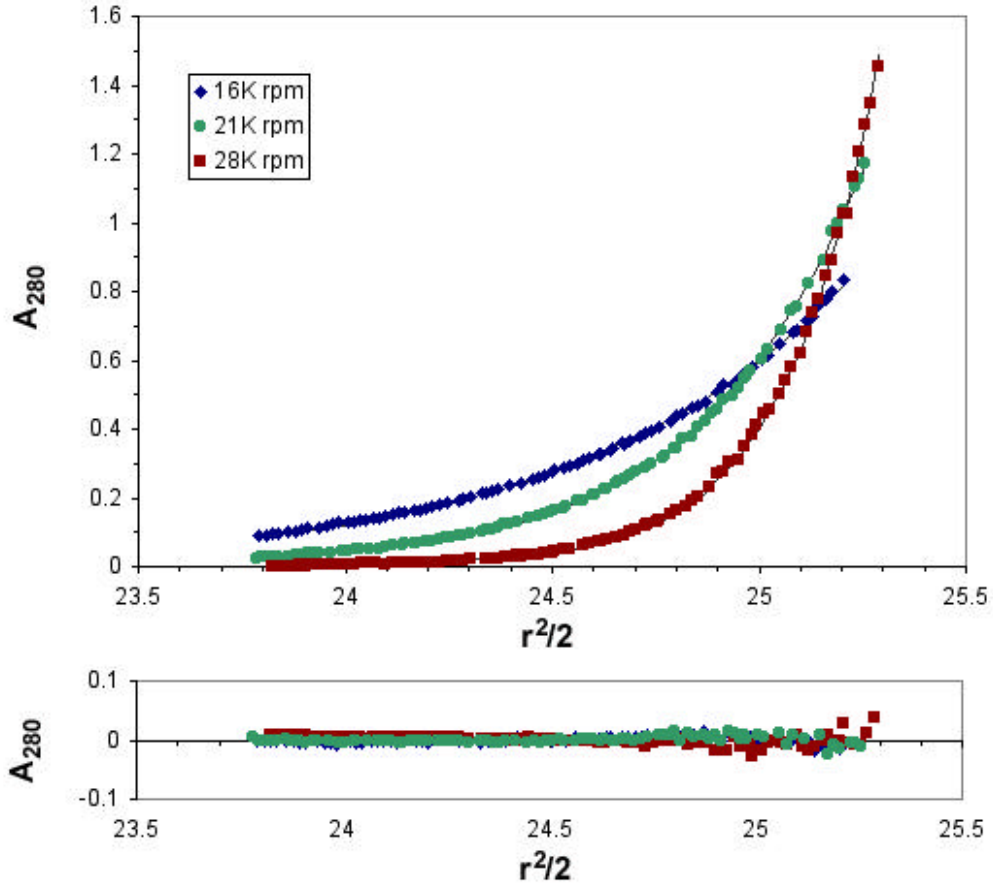


Figure 8. Equilibrium sedimentation data for △ssInv-NHis

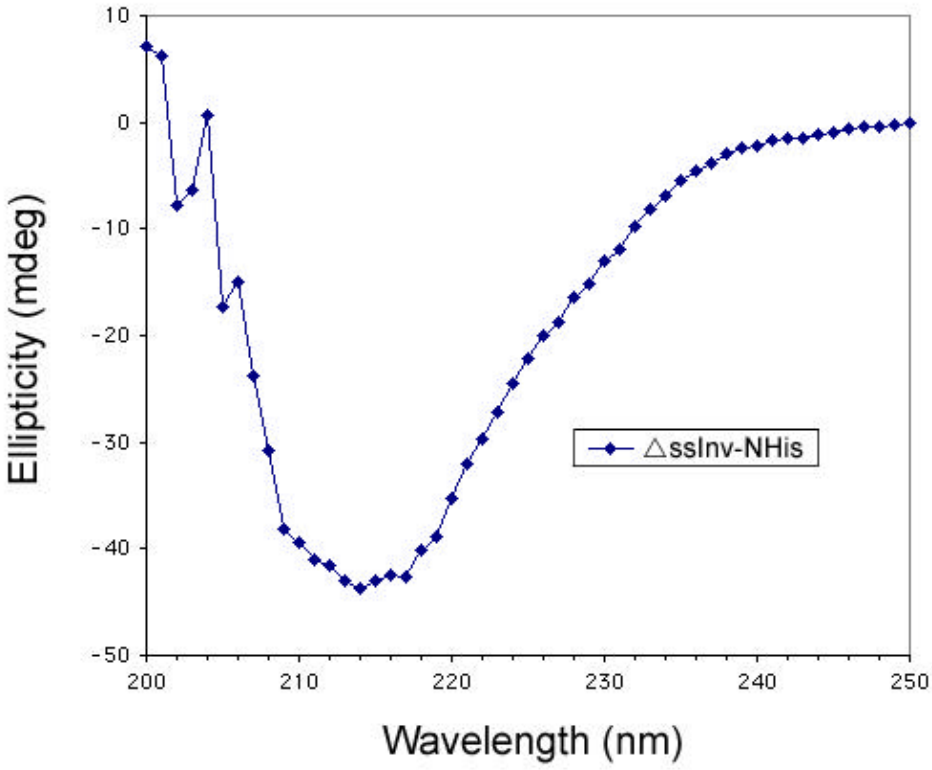
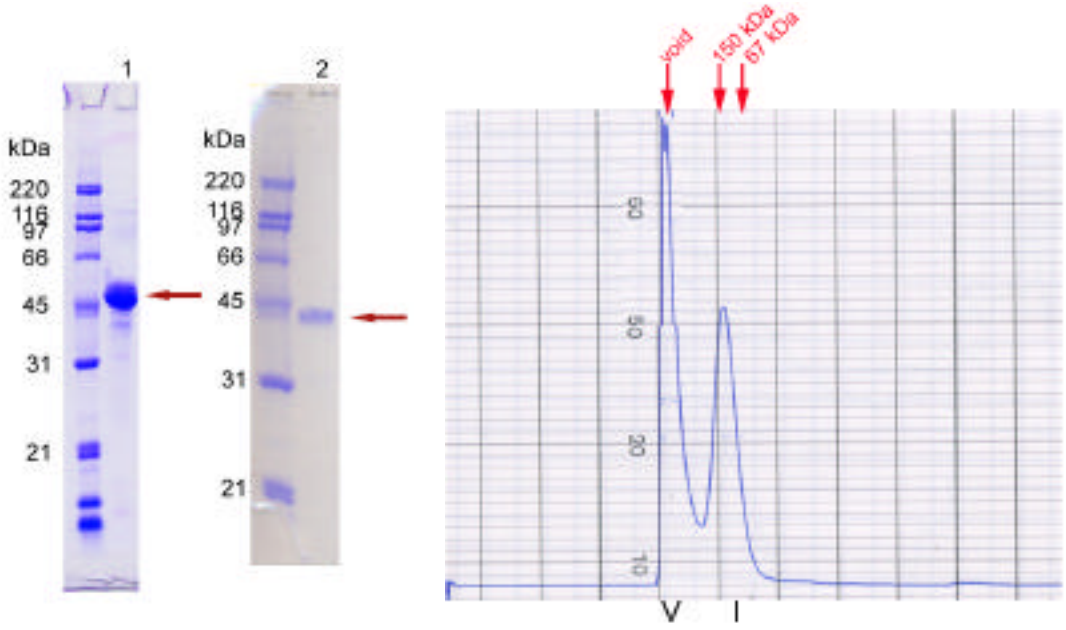


Figure 9. Circular dichroism spectrum of △ssInv-NHis

A.



B.

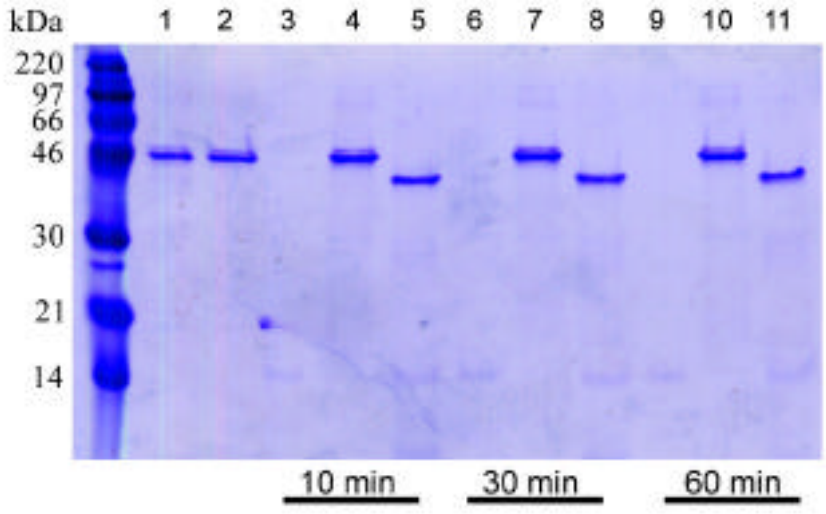


Figure 10. Proteolysis of △ssInv-NHis

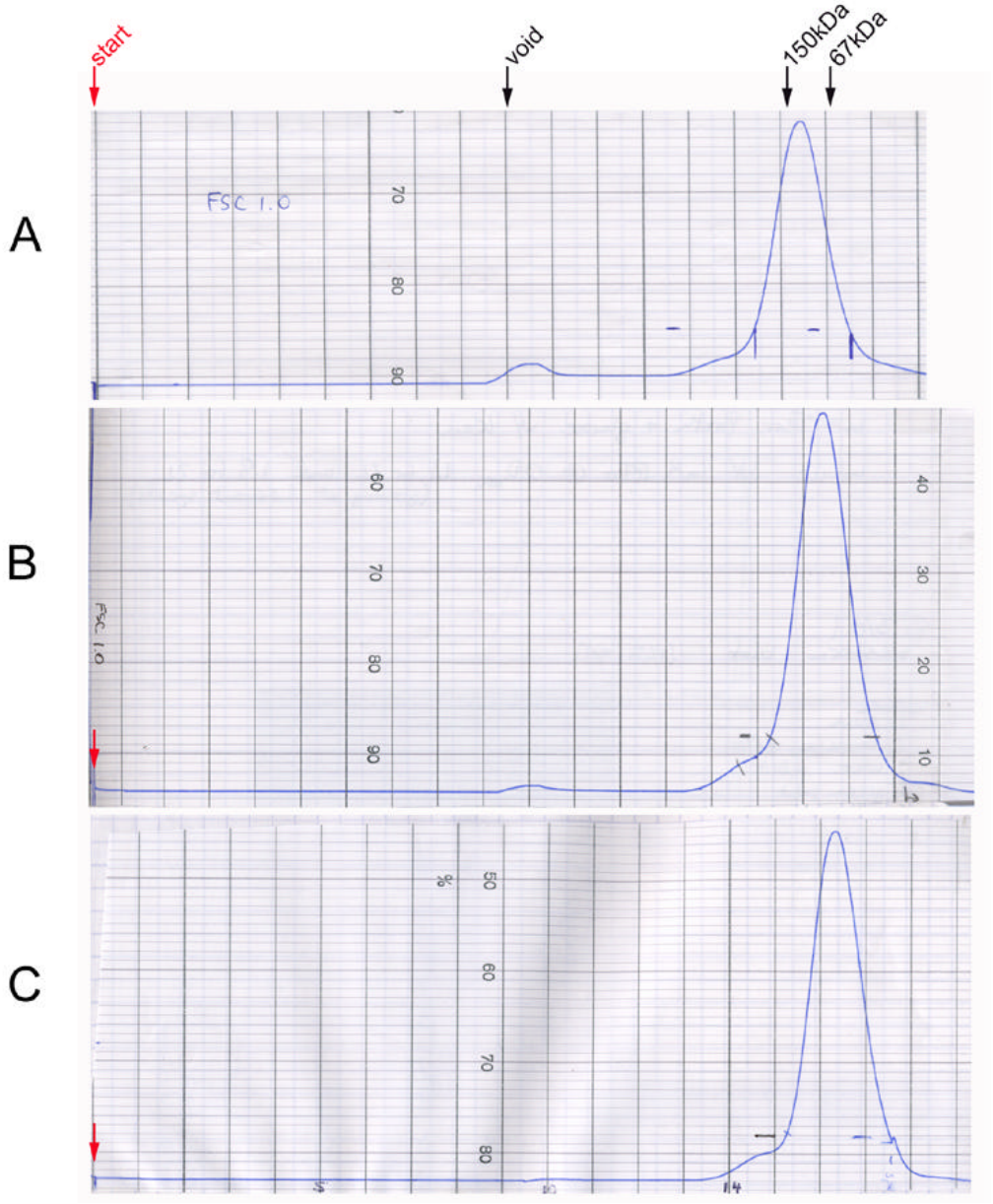


Figure 11. Size exclusion chromatography of △ssInv-NHis, △ssInv-NHis-127 and △ssInv-NHis-149

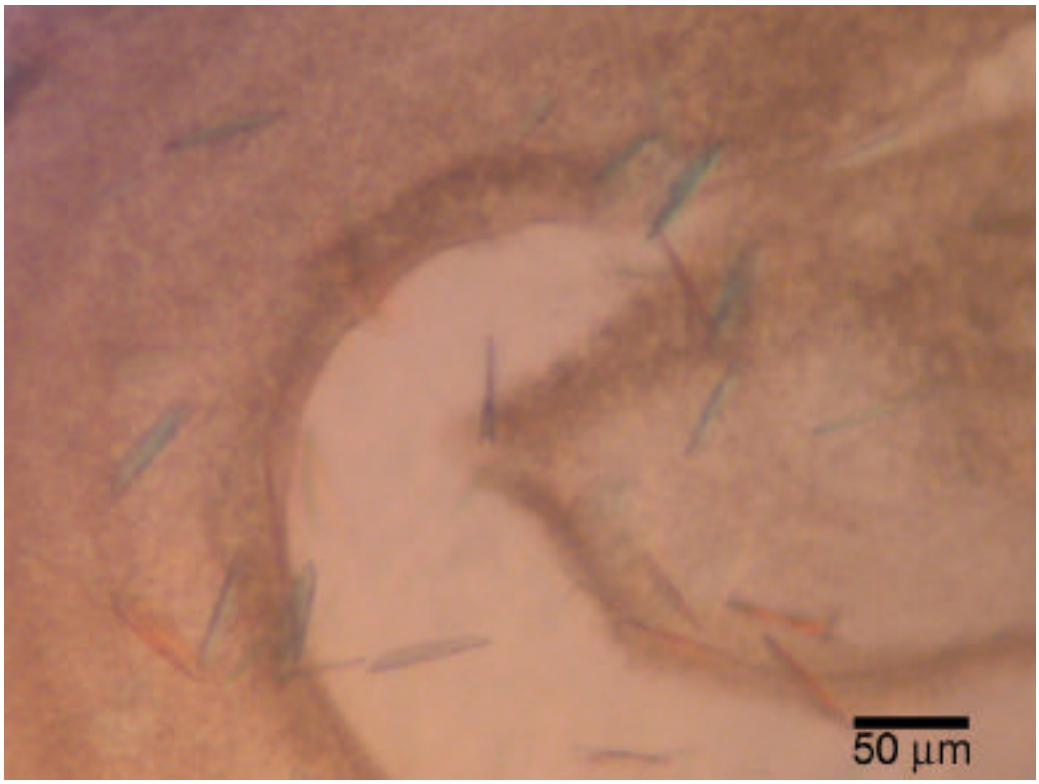


Figure 12. △ssInv-NHis-127 crystals

Appendix A:

This chapter was published as R. R. Isberg, Z. Hamburger and P. Dersch, “Signaling and invasin-promoted uptake via integrin receptors.” *Microbes and Infection* **2**, 793-801 (2000).

Signaling and Invasin-Promoted Uptake via Integrin Receptors

For: Microbes and Infection Forum, G. Tran Van Nhieu and D. Ojcius, Eds.

Ralph R. Isberg [1,2], Zsuzsa Hamburger[3] and Petra Dersch[2,4]

[1]Howard Hughes Medical Institute

[2]Department of Molecular Biology and Microbiology

Tufts University School of Medicine

136 Harrison Ave., Boston, MA 02111

[3]Division of Biology 156-29

California Institute of Technology

Pasadena, CA 91125

[4]Current Address: Institute of Microbiology, Freie Universität, Königin-Luise Str. 12-16,

14195 Berlin, Germany

E-mail: risberg@opal.tufts.edu

Abstract

The invasin protein encoded by enteropathogenic *Yersinia* allows entry of bacteria into intestinal M cells by binding to integrin receptors. In cultured cells, invasin mediated uptake requires proteins involved in endocytosis and signaling to the cell cytoskeleton. At least four different factors have been demonstrated to play a role in regulating the efficiency of invasin-promoted uptake. These include receptor-ligand affinity, receptor clustering, signaling through focal adhesion kinase, and stimulation of cytoskeletal rearrangements by small GTP binding proteins.

Introduction

A number of bacterial pathogens are able to enter normally non-phagocytic cultured cells [1]. There are numerous potential roles that cellular entry plays in establishing disease, and each pathogen has a distinct reason for assuming its intracellular localization. Some organisms, such as *Shigella* or *Chlamydia* species, must replicate within host cells to cause disease [2]. Alternatively, organisms such as the enteropathogenic *Yersinia* and *Salmonella* appear to utilize cellular entry to gain access to subepithelial regions. In *Salmonella* and enteropathogenic *Yersinia* species, it is likely that entry within epithelial cells is a transient event in systemic disease that leads to infection of regional lymph nodes. On the other hand, organisms such as *Neisseria meningitidis* may use specific uptake factors to directly access deep tissue sites [2]. Once microorganisms translocate across the epithelium, the routes that different organisms take to promote disease may diverge significantly from one another. In fact, once found within tissues, *Yersinia* and *Salmonella* species interact with macrophages in very different fashions from one another, with the former avoiding phagocytosis, and the latter probably growing within phagocytic cells [3].

The enteropathogenic *Yersinia, Y. pseudotuberculosis* and *Y. enterocolitica*, initiate disease after bacterial translocation across M cells into the Peyer's patches of the small intestine [4-6]. In humans, the infection usually results in localized mesenteric lymphadenitis, with reactive arthritis in HLA-B27 patients being the most common systemic disease [7]. In individuals having the genetic disorder hemochromatosis, however, a systemic infection may ensue that affects multiple organ sites [8]. Bacteria spread to the liver and spleen in mouse models, with death occurring within nine days

after initial ingestion [6]. It is generally believed that systemic disease in the mouse is the result of spread from the Peyer's patch, but there is no evidence that this is the case. The disease process within the mouse is quite similar for the two organisms, although in oral infection models, *Y. enterocolitica* has a lower LD_{50} and grows to much higher titers in the Peyer's patch than does *Y. pseudotuberculosis*.

Efficient entry of enteropathogenic *Yersinia* into both cultured mammalian cells and M cells overlying Peyer's patches requires the bacterial invasin protein, which binds multiple members of the integrin receptor superfamily [9]. Integrins are large heterodimeric transmembrane proteins that are capable of transmitting signals to the cell cytoskeleton after adhesion to substrates. The five integrins that bind invasin ($\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_5\beta_1$, $\alpha_6\beta_1$, and $\alpha_v\beta_1$) are either receptors for extracellular matrix proteins or are involved in cell-cell interactions [9, 10]. The best characterized of this group is $\alpha_5\beta_1$, which is also known as the fibronectin (Fn) receptor. Integrins are primarily localized on the basal surfaces of cells in the epithelium, and are not normally available for contact with bacteria in the lumen of the intestine. The exception to this rule is the M cell, which has at least two integrin invasin receptors uniformly distributed about its surface ([11]; Marra, A. and Isberg, R., unpublished results). This provides a likely reason for why enteropathogenic *Yersinia* are found exclusively within these cells in the epithelial layer.

Much circumstantial evidence indicates that after entry and translocation across M cells, *Yersinia* are largely extracellularly adherent. Undoubtedly, this is the result of bacterial proteins that antagonize internalization, most notably YopH and YopE (and YopT in *Y. enterocolitica*), which are translocated from adherent bacteria to target cells

via a Type III secretion system [12-15]. In cultured mammalian cells, bacteria that express YopE and YopH are found tightly adherent to host cells, initiating frustrated phagocytic events. YopE and YopH are clearly disrupting signaling events that are required to complete invasin-promoted uptake.

Structure and binding of the Y. pseudotuberculosis invasin

Invasin is sufficient to allow standard *E. coli* laboratory strains to be internalized by most cultured cell lines that express the appropriate integrin receptors [9]. The protein is part of a family of adhesins encoded by enteropathogenic bacteria that includes the intimins, which are involved in promoting attaching and effacing lesions in enterohemmorhagic and enteropathogenic *E. coli*, as well as *Citrobacter freundii* [16]. All members of this family appear be involved in binding receptors inserted into the mammalian cell plasma membrane that signal to the host cell cytoskeleton [17]. The most significant region of similarity between these family members is located in the amino terminal 500 amino acids of the *Y. pseudotuberculosis* invasin (36% identity), which is required for outer membrane localization and export of the carboxyl termini of these proteins. It is likely that this region forms a β-stranded barrel in the outer membrane, allowing it to act as an autotransporter domain [18].

The region of the protein that is the most highly divergent at the primary sequence level from the intimins is the carboxyl terminal 192 amino acids of the *Y. pseudotuberculosis* protein, which is the shortest region of the protein able to bind integrin receptors [18]. The analysis of point mutations in this region indicates that two noncontiguous sites at the primary sequence level are involved in binding integrin

receptors. The first region is localized in a disulfide loop and centered around the residue Asp^{911}. Sidechain changes of this asp have drastic affects on integrin binding and uptake, and even a conservative change to a glu residue results in total loss of bacterial uptake into host cells [19]. It has been proposed that this residue performs a similar role to the Asp^{1495} located in the arg-gly-asp sequence of Fn, which is critical for the binding of this protein to the integrin receptor $\alpha_5\beta_1$ [20]. The second region is centered around residue Asp^{811}, although residue changes in this region have much weaker affects than seen with Asp^{911} [21]. Fibronectin also has a binding determinant not contiguous to arg-gly-asp, called the synergy region, that appears to play a role in supporting integrin binding [22]. Thus, although there is no detectable sequence similarity between invasin and fibronectin, sequence determinants involved in receptor recognition appear to be similar.

The crystal structure has been determined for residues 503-986 of the *Y. pseudotuberculosis* invasin (Fig. 1), which is a region of the protein capable of promoting uptake when attached to the surface of latex beads or bacteria [23, 24]. The structure reveals a 180 Å-long tandem array of five domains (D1-D5; Fig. 1). The four most amino terminal domains (D1-D4) each have folding topologies that are similar to immunoglobulin superfamily (IgSF) domains, although there is no apparent sequence similarity to proteins having this folding structure. The most carboxyl terminal domain, D5, has a topology similar to C-type lectin-like domains (CTLDs), including carbohydrate binding proteins and natural killer cell receptors for which the carbohydrate recognition properties are unclear [25]. The folding topology of the disulfide-containing D5 is much more closely related to the latter proteins, as both are missing a loop involved in calcium binding and carbohydrate recognition. Extensive interaction of D4 and D5,

comprising the 192 residue integrin binding region, establish a large interdomain interface, which is unusual for interfaces involving IgSF family members. The interface creates a cell adhesion superdomain predicted to have little flexibility between domains (Fig. 1). The fact that a two domain structure is involved in integrin binding is also reminiscent of Fn, in which two separate sequence determinants of domains III9-10 are involved in receptor recognition.

The structure of D4-D5 supports the model that the Asp^{911} and Asp^{811} residues play similar roles to the arg-gly-asp and synergy regions in fibronectin. Both Asp^{911}, located in D5, and arg-gly-asp are found in protruding loops in the carboxyl terminal domain of two-domain adhesion modules (Fig. 2; [20]). Asp^{811}, located in invasin D4, and $Asp^{911}$are separated by 32 Å. In fibronectin, Asp^{1373} found in the synergy region, is also 32 Å separated from the Asp^{1495} in arg-gly-asp (Fig. 2; [20, 23]). In addition, the invasin and fibronectin synergy regions have arginine residues that are located the identical distances from the critical asp residues found in their respective carboxyl terminal cell adhesion domains ([22];Fig. 2). One striking difference is that the large buried surface between D4-D5 probably ensures significant rigidity in the region between $Asp^{911}$and Asp^{811}. In contrast, the interface between the two Fn domains involved in integrin recognition is approximately 25% that of D4-D5, presumably allowing interdomain flexibility. As invasin binds integrins with 100-fold higher affinity than fibronectin, interdomain rigidity may facilitate receptor recognition or stabilize ligand-receptor interaction.

Role of clustering and affinity in uptake

The crystal structure suggests striking similarities in the receptor recognition properties of Fn and invasin. In fact, the two substrates appear to bind the same site on the $\alpha_5\beta_1$ receptor, based on competitive inhibition studies, mutational analysis and monoclonal antibody inhibition results [26]. In spite of these similarities, invasin is much more efficient than Fn at promoting uptake of either bacteria or latex beads into cultured cells, particularly under conditions in which the amount of receptor on the target cell is relatively low [27, 28]. The much higher receptor binding affinity of invasin appears to be the reason for this difference. To demonstrate this, bacteria were coated with mAbs directed against $\alpha_5\beta_1$ and tested for the ability to be internalized by cultured cells [28]. Uptake of the bacteria was found to be a function of the affinity of these artificial integrin ligands for receptor, with uptake linearly proportional to the logarithm of the effective dissociation constant. In addition, a point mutation in invasin with 100-fold lower affinity than wild type is completely defective for promoting bacterial uptake [19].

There are two plausible explanations for the structural basis of higher affinity invasin binding relative to fibronectin. First, the rigidity of the D4-D5 cell adhesion module may lock the protein in a binding-optimal conformation that stabilizes ligand-receptor interaction. Fn, in contrast, appears to have great flexibility between the two domains involved binding integrins, and it may assume several conformations that are not optimal for binding. Secondly, in the region between the $Asp^{911}$and Asp^{811} residues, there exists a patch of 5 aromatic amino acids that forms a protrusion on the surface of D4 (Fig. 2). This region could facilitate either hydrophobic interactions or hydrogen bonding with the receptor and contribute greatly to the binding energy. Fn has a deep cleft in the

corresponding region that may not contribute significantly to binding.

Self-association of invasin appears to play a role in uptake, presumably resulting in multimerization of the integrin receptor. This is in spite of the fact that the soluble carboxyl terminal 497 residue fragment of invasin is primarily a monomer in solution, and the crystallized form is monomeric [23, 24]. Latex beads are unable to be internalized if they have been coated by a gel fractionated monomeric fragment of the *Y. pseudotuberculosis* invasin containing only the D4 + D5 superdomain bound to Fab [24]. In contrast, beads coated by either the D1-D5 region of invasin or a D4 + D5 fragment that has been dimerized by binding to antibody are readily internalized by cultured cells. This result indicates that there is a region in domains D1-D3 that enhances uptake of cell-associated beads, and that the enhancement can occur simply by multimerization. D1-D3 appears to promote stimulation by undergoing homotypic interaction. The IgSF-like domain D2, which is necessary to give full stimulation of uptake, has the ability to self associate, and invasin derivatives expressed on the bacterial cell surface having this domain can be chemically self-crosslinked [24]. It is interesting to note that the *Y. enterocolitica* invasin, which is less active at promoting uptake on a per molecule basis than the *Y. pseudotuberculosis* protein, is lacking D2 and has little ability to promote self-association (Dersch, P. and Isberg, R.R. submitted). The lowered activity of the *Y. enterocolitica* protein can be largely, but not entirely, overcome by expressing the protein at high levels on the bacterial cell surface. This may allow individual invasin molecules to promote localized multimerization of the integrin receptor after bacterial binding.

The presumed reason that invasin self-interaction plays an important role in promoting uptake is because after bacterial binding to target cells, individual integrin

receptors must associate with one another to generate the appropriate signals within the host cell cytoplasm necessary to direct phagocytosis. Similarly, many other receptors have been shown to be unable to generate downstream signals in the absence of dimerization, and on a number of occasions monomeric substrates have been demonstrated to be unable to stimulate integrin-dependent cytoskeletal rearrangements [29]. The surprising aspect of this result is that most of the studies demonstrating the importance of dimerization or clustering of receptor in response to ligand binding have involved the association of receptors with soluble substrates. In the case of invasin binding to host cells, the substrate is presented as an array on the surface of the bacterium or bead. This array may not be able to promote uptake because it does not bring individual integrin molecules in sufficiently close proximity to send an intracellular signal. Perhaps signaling requires a protein located in the mammalian cell cytoplasm that must bind multiple integrin receptors simultaneously to promote activation of downstream effectors involved in bacterial uptake. If the concentration of invasin is sufficiently high, however, invasin self-interaction may become dispensable because bacterial binding would allow receptors to be close to one another.

High affinity binding appears to be required for reasons other than multimerizing receptors, as there is evidence that much larger numbers of receptors must be recruited to the phagocytic cup than can be explained by simply dimerization of a population of receptors [28]. For instance, bound bacteria are not efficiently internalized if the target mammalian cell is immobilized on high affinity integrin substrates, indicating that sequestration of receptors away from the bound bacteria inhibits internalization. Furthermore, bacteria allowed to adhere to mammalian cells at low temperatures can be

blocked from internalization after raising the incubation temperature to 37 ^{o}C, if an activity blocking anti-invasin mAb is included during the 37 ^{o}C incubation (Fig. 3). This implies that adhesion to the mammalian surface is not sufficient for invasin to promote uptake, and that integrin receptors must bind circumferentially about the surface of the bacterium. Presumably, internalization involves the formation of a zipper-like structure about the bacterium (Fig. 3). Therefore, although signals must be sent from the phagocytic cup to host cell cytoplasm, the bacterial surface must also serve as a template for migration of the mammalian cell plasma membrane.

Role of integrin cytoplasmic domain in uptake

Signals from ligated integrins must be transmitted via either the cytoplasmic or transmembrane domains of the receptors. Both the α and β_1 integrin chains of the invasin receptors have short cytoplasmic tails that are known to regulate function of the receptor [30]. The β_1 chain cytoplasmic domain also interacts directly with a variety of cytoskeletal proteins such as talin, α-actinin, focal adhesion kinase (FAK), paxillin and ICAP-1 [31-33]. Invasin-mediated internalization can take place in the absence of the α chain cytoplasmic domain, but residues within the β_1 chain cytoplasmic domain are critical for uptake [34]. Surprisingly, most alterations of the cytoplasmic domain that lower the affinity of the receptor for cytoskeletal elements increase the efficiency of invasin-mediated internalization [34]. This is based on the observation that many mutations in the cytoplasmic domain that interfere with the ability of the integrin to localize in focal adhesions simultaneously stimulate the efficiency at which these receptors are able to promote uptake [30].

The most likely explanation for why lowered affinity for the cytoskeleton stimulates uptake is that integrin receptors bound tightly to cytoskeletal proteins can become immobilized, restricting the ability of receptors to move to the site of bacterial adhesion. Mutant receptors that have lowered affinity for the cytoskeleton may have hightened mobility in the plasma membrane, allowing them to be efficiently sequestered to the nascent phagosome. An alternate explanation is that lowered affinity for the cytoskeleton may stimulate uptake because a high concentration of either a particular cytoskeletal protein or a group of them at the site of bacterial binding may force the bacterium to remain surface adherent as is seen with enteropathogenic *E. coli.* Perhaps aggregation of critical cytoskeletal proteins at the site of bacterial binding, such as paxillin, vinculin or α-actinin, results in an immobilized phagosome that can initiate, but not complete the phagocytic process. By this model, lowered affinity for the integrin allows cytoskeletal proteins to interact with the nascent phagocytic cup without resulting in a rigid structure that favors extracellular adhesion of the bacterium.

Signaling events involved in integrin-promoted uptake

Based on the results that lowered affinity of the integrin for the cytoskeleton stimulates uptake, it is reasonable to question whether actin or cytoskeleton-associated factors are required for uptake. Circumstantial pieces of evidence exist that point to a general need for rearrangement of cytoskeletal-associated elements. These include the use of inhibitors and studies on the activities of antiphagocytic factors, both of which have shortcomings in their individual lack of specificities. Invasin-promoted uptake is highly sensitive to inhibitors of cytoskeletal polymerization (such as cytochalasin D

[35]). Inhibitors of tyrosine kinase activity also cause lowered uptake rates [35]. Many tyrosine kinases are involved in regulating cytoskeletal rearrangements, and it is likely that this inhibition is a result of blocking these events. In addition, the two Yops that inhibit phagocytosis, YopE and YopH, have activities that clearly antagonize the function of cytoskeleton-associated proteins.

YopH protein is one of the most active tyrosine phosphatases ever identified, with little substrate specificity *in vitro* [36]. After injection into host cells by the *Yersinia* Type III secretion system, its affects bear great similarity to what is seen after the addition of tyrosine kinase inhibitors such as genistein, with invasin-mediated uptake inhibited by 3-10 fold. The specific target that is inactivated by YopH, resulting in lowered bacterial uptake, is unclear. It is possible that YopH antagonism of uptake is the cumulative effect of targeting multiple signaling proteins. Within host cells, there appears to be a hierarchy of preferred YopH substrates (p130CAS, FAK, and Paxillin) that may give some insight into how YopH alters the uptake efficiency [13, 14]. The substrate that is dephosphorylated with the most rapid kinetics is the focal adhesion-associated protein p130CAS, and loss of focal adhesions in the cell occurs relatively soon after this protein loses tyrosine phosphorylation [14]. Approximately 60 min after p130 is dephosphorylated, loss of tyrosine phosphates from FAK occurs. This makes it unlikely that dephosphorylation of FAK by YopH plays a direct role in inhibiting uptake. Furthermore, since loss of focal adhesions may precede the dephosphorylation of FAK, it is unclear whether it is the absence of focal adhesions or the presence of YopH that causes loss of tyrosine phosphates from FAK.

YopE has much more drastic affects on cells than is observed with YopH, with

general loss of actin filaments observed in cells injected by this protein. Since wholesale changes are taking place in these cells, attributing loss of uptake to loss of a single signaling pathway may be difficult. YopE is highly homologous to SptP, a protein encoded by *S. typhimurium* that appears to be a fusion between YopE and YopH and has a highly specific GTPase activating protein activity (GAP) for the Rho family GTPases Rac1 and CDC42 [37, 38]. Both YopE and SptP have a consensus sequence called an arginine finger that is found in Rho GAPs, so it is likely that YopE targets Rho family members. YopE causes loss of stress fibers from cells, implying that at the very least it targets RhoA, or a related protein necessary for stress fiber formation [39]. As will be pointed out below, however, inactivation of RhoA cannot explain why YopE depresses uptake.

Focal Adhesion Kinase: direct role or a regulator of phagocytic capability?

Focal adhesion kinase (FAK) is a large tyrosine kinase that has autophosphorylation properties as well as the ability to interact with a variety of signaling and cytoskeleton-associated proteins [40]. In response to adhesion of a cultured cell to extracellular matrix proteins, FAK becomes tyrosine phosphorylated at several key sites, including Tyr397 located amino terminal to the kinase, which when phosphorylated is recognized by the SH2 domain of Src kinase [41, 42]. The carboxyl terminal of the protein is called the focal adhesion targeting region (FAT) because it binds a number of cytoskeletal proteins and is responsible for localization of the protein in focal adhesions. The analysis of dominant inhibitory mutations in FAK and a fak (-/-) cell line are consistent with this protein playing a role in the related processes of cell migration and

turnover of focal adhesions [43]. Although the current data are inconsistent with dephosphorylation of FAK playing a key role in YopH inhibition of invasin-promoted internalization, there are still multiple reasons to consider that this tyrosine kinase plays a role in uptake. FAK has been shown to bind the integrin β_1 chain, and may directly signal from the phagocytic cup to the cytoskeleton [32]. Furthermore, FAK appears to regulate the stability of focal adhesions in the cell, which may have great significance for invasin-promoted uptake [34].

Dominant inhibitory forms of FAK, including one having little more than the FAT region and another containing a substitution at Tyr397 that prevents its phosphorylation, significantly reduce invasin mediated uptake [44]. Furthermore, the fak (-/-) cell that lacks all detectable FAK is highly defective for internalization of invasin-coated particles, indicating that loss of uptake is not an artifact of using dominant inhibitory forms of the protein.

There are three straightforward models for how FAK could be involved in invasin-promoted uptake (Fig. 4): 1) it could signal directly to the cytoskeleton after binding integrin that is clustered at the site of bacterial binding; 2) it could promote turnover of cytoskeleton proteins at the site of bacterial binding, ensuring that an immobile adhesion complex antagonistic to uptake does not form at this site; or 3) it could allow dissociation of adhesion complexes containing integrin bound to the extracellular matrix, facilitating recruitment of receptor to the bound bacteria. Although all of these models could be correct, the analysis of the fak (-/-) cell line indicates that there is evidence in favor of the last model. This cell line is highly adherent and broadly spread on extracellular matrix with large numbers of focal adhesions, presumably reflecting a defect in their turnover.

After transfection of such cells with the wild type FAK cDNA, the cells are significantly reduced in their spreading and their number of focal adhesions, while their ability to internalize invasin-coated beads is greatly enhanced. Interestingly, a fragment of FAK lacking its kinase domain introduced into these cells reduces the number of focal adhesions and the amount of spreading compared to the untransfected control. There is considerable invasin-promoted uptake in these cells as well, suggesting that FAK kinase activity is not directly required to signal from the phagocytic cup to the cytoskeleton. Rather, because introduction of this FAK fragment reduces cell spreading and focal adhesion formation, it appears that turnover of focal adhesions is required for efficient invasin-promoted uptake. Therefore, FAK may regulate the phagocytic capability of the pool of integrins found in the cell.

The analysis of FAK appears to parallel the results observed regarding mutations in the integrin β_1 chain cytoplasmic domain. In each case, conditions that rigidify the integrin receptor into an immobilized structure result in lowered efficiencies of uptake. Lowering the affinity of the cytoplasmic domain for cytoskeleton-associated proteins, or increasing the turnover of adhesion structures in the cell via the action of FAK, result in a stimulation of the efficiency of invasin mediated uptake (Fig. 4; Model 3). It should be emphasized that the global effects that ensue from changing the duration a receptor remains associated with a focal adhesion structure may not be the sole reason that the presence of FAK stimulates uptake. The other models in Fig. 4, therefore, have not in any way been discounted.

GTP binding and invasin promoted uptake

Invasin-promoted uptake has a strong requirement for GTP, and is stimulated by the presence of the nonhydrolyzable analog GTP-γ-S. In addition, mammalian cells treated with *C. difficile* Toxin B, which glucosylates multiple members of the Rho family of small GTP binding proteins, are highly defective for invasin-promoted uptake. Members of the Rho family are known to regulate a number of cytoskeletal, morphological and membrane trafficking activities in mammalian cells, so Toxin B treatment has wide-ranging affects on the host cells [45, 46]. In addition, studies have shown that phagocytosis mediated by Fc receptor, complement receptor and *S. typhimurium* each require the activity of at least one of these proteins.

Further analysis of Rho family members indicates that invasin-mediated uptake has a specific requirement for Rac1. Uptake is highly stimulated by activated Rac1, while interfering Rac1 mutants result in cells highly defective for internalization of bound bacteria (M. Alrutz and R. Isberg, unpublished results). Furthermore, activated Rac1 colocalizes with the nascent phagosome formed about *Yersinia pseudotuberculosis*, consistent with the protein directly signaling from the site of bacterial entry. In contrast, inactivation or introduction of dominant interfering forms of several other Rho family members does not depress uptake. In fact, inactivation of RhoA, by treatment with *C. botulinum* exoenzyme C3 appears to mildly stimulate uptake of *Y. pseudotuberculosis*. One possible explanation for this latter result is that RhoA is required for formation of stress fibers in the cell that may interfere or compete with the formation of actin filaments necessary for invasin-promoted uptake. Removal of stress fibers by inactivation of RhoA may facilitate penetration of the phagosome into the cell by eliminating meshwork that

normally would have to be broken to allow uptake.

It is surprising that there is such a distinct dependence on Rac1 for invasin-promoted uptake, especially considering the fact that engagement of β_1 integrins is believed to result in activation of multiple Rho family members. It is possible that the Rac-dependent signal from the bound bacterium activates the cytoskeletal processes to the extent that there is no consequence from the introduction of interfering forms of CDC42 from the system. By this formulation, other Rho family members may be involved in uptake, but there is simply no interference by GDP-bound forms of these proteins. Alternatively, the actin-associated events necessary to facilitate integrin-promoted internalization of a bound bacterium may only be coordinated by the action of Rac1. If this is the case, then a unique combination of downstream effectors that only Rac1 can activate may be necessary for efficient uptake.

Determination of the effectors downstream of Rac1 will probably allow an explanation for the specific dependence on this GTPase. Targets of Rac1 are likely to promote divergent events associated with actin nucleation and scission. Recently, it has become clear that Rho family members have the ability to activate Arp 2/3-dependent actin nucleation by interacting either directly or indirectly with members of the WASP/Scar1 family [47]. The best characterized of this group in regards to actin interactions is the protein N-WASP, which appears to directly bind CDC42 and activate Arp 2/3 nucleation of actin. As dominant interfering forms of CDC42 have no affect on invasin-promoted uptake (M. Alrutz and R. Isberg, unpublished), it does not seem likely that N-WASP plays a direct role in this process, since the interfering form of CDC42 probably would take N-WASP out of action in such cells. On the other hand, other

nucleators are potential candidates for initiating actin rearrangements, since at least one WASP/Scar1 family member does not directly bind any known Rho family members, and this protein may be indirectly activated by Rac1 [48]. It is interesting to note that this protein, WAVE (also called Scar1), is insensitive to inhibition by dominant interfering forms of CDC42 [48], suggesting that not all actin nucleation events initiated by members of WASP/Scar1 proteins require the activity of this particular Rho family member.

Actin filament scission events, on the other hand, are probably controlled by a set of effectors that are working in parallel to actin nucleation after activation by Rac1. There is good reason to believe that scission of actin may be required for invasin-promoted uptake. First, actin presents a relatively tight mesh in the cell that may antagonize phagocytic uptake. Signaling events that initiate localized breakdown of actin may be required to move a nascent phagosome into the cell. Second, integrin receptors that are immobilized in focal contacts may need to be liberated from these sites to be recruited to the phagocytic cup to complete the uptake process. Breakdown of actin may facilitate this movement. Third, if too much polymerized actin is recruited to the site of bacterial binding, then this may result in a phagosome that is paralyzed, with a stable focal adhesion forming at this site that prevents completion of the phagocytic process. Finally, scission events themselves may generate foci for further actin polymerization. The details of how Rho family members regulate filament breakage is poorly understood, but recent results indicate that members of the PAK65 family of CDC42/Rac1-interacting proteins, which are ser/thr kinases, may be important initiators of actin filament breakage.

Future approaches

In regards to invasin-promoted uptake, the details of what must occur at the extracellular surface of the mammalian cell to promote uptake are well characterized. To analyze these events, a series of experiments were performed that not only identified conditions that inhibited uptake, but also uncovered factors that lead to enhanced uptake, such as increasing the affinity of binding as well as multimerization of invasin. In contrast, the known mammalian cytoplasmic proteins required for uptake have been almost entirely identified by the use of inhibitors, or the use of dominant-inhibitory forms of signaling molecules. As these treatments result in global changes to the cell that may affect many processes other than bacterial uptake, it is often difficult to determine if these experiments cause direct effects on the phagocytic cup or are changing the general phagocytic capability of the cell. Future work will require that assays be devised that allow identification of the minimal requirements for phagocytic uptake, as well as manipulating the amounts of proteins believed to be required for uptake. Such experiments will require the use of cell lines from mouse mutants specifically defective for single cytoskeletal factors, transfected with complementing clones that allow fine regulation of expression of the appropriate factors. As the cytoskeleton has many apparently redundant factors, these experiments will certainly generate their own complications. By analyzing both the clear results and the difficulties that arise from such approaches, a detailed understanding of how integrin receptors can collaborate with cytoskeletal factors to promote bacterial uptake will certainly be achieved.

Acknowledgments

Work described in the laboratory of R. I. was supported by grant RO1-AI23538 from NIAID, and a Howard Hughes Medical Institute Investigator award. We thank Dr. Kaspar Locher for help in constructing Figure 1, and Dr. Guillaume Dumènil for review of the text.

Figure Legends

Fig. 1. Crystal structure of invasin modeled on the surface of the bacterial outer membrane. Shown is the determined crystal structure of *Y. pseudotuberculosis* invasin carboxyl terminal residues 503-986 [23] placed on the bacterial cell surface with an outer membrane spanning region modeled on the crystal structure of FhuA. Shown in blue is a model for the outer membrane spanning region using the published structure of the β-barrel region of FhuA [49, 50]. The invasin structure consists of five domains in tandem (D1-D5, with D1 being at the amino terminal end), of which D1-D4 have the folding patterns of IgSF members. D5 has a C-type lectin-like fold, and the large interface between D4 and D5 gives the appearance of a single superdomain (D5+D4). D5+D4 is the cell adhesion module of the protein.

Fig. 2. Cell adhesion domains of invasin and Fn show similar alignments of aspartates and arginine known to be involved in Fn recognition of receptor. Shown are the crystal structures of just the two domain integrin recognition modules of each molecule that have mutations affecting receptor recognition. Displayed in green are aromatic residues found in the region between Asp911 and Arg883 of invasin that may provide an interface with integrin receptors.

Fig. 3. Test of the Zipper Model for invasin promoted uptake. Bacteria are allowed to bind to mammalian cells on ice. To determine whether this binding event triggers uptake, or whether there is a requirement for further binding of receptors about

the surface of the bacterium, the infection mix is incubated with a function blocking mAB that recognizes the D4+D5 region of invasin. The incubation temperature is then raised, and uptake is determined. If bacteria are internalized in the presence of the mAb, then this is taken to indicate uptake occurs as a result of an initial "triggering" event that signals completion of the process. If the mAb inhibits uptake, then this implies an initial signaling event due to bacterial binding is not sufficient to promote uptake, and that circumferential binding of the mammalian cell surface about the bacterium ("zippering") is necessary for particle internalization. When this experiment is performed, the bacteria are unable to be internalized, indicating that binding about one face of the bacterium is insufficient to trigger uptake.

Fig. 4. Models for the role of FAK in invasin-promoted uptake. Model 1: Clustering of integrin receptor allows recruitment of FAK to the site of bacterial adhesion. The presence of FAK at this site, or its kinase activity, results in the ability of downstream signals to be sent to activate cytoskeletal events leading to uptake of the bacterium. Shown are Paxillin and α-actinin, two potential downstream cytoskeleton-associated proteins known to localize at site of integrin adhesion. Model 2: Action of FAK antagonizes the formation of stable focal adhesion at site of bacterial binding. FAK prevents an immobile structure from forming that prevents the phagocytic cup from moving about the bacterial cell surface. By breaking up a cytoskeletal complex that forms at the site of bacterial binding, uptake can proceed after the initial binding event. Model 3: FAK promotes turnover of focal adhesions at the site of integrin binding to extracellular matrix. The presence of FAK allows the integrin receptor to be liberated

from the focal adhesion, so that the receptor may move freely in the membrane. This allows the eventual recruitment of the integrin to the site of bacterial binding.

References

1. Finlay, B. B., and Falkow, S. Common themes in microbial pathogenicity revisited. Microbiol. Mol. Biol. Rev. **61** (1997) 136-69.
2. Sansonetti, P. J., Tran Van Nhieu, G., and Egile, C. Rupture of the intestinal epithelial barrier and mucosal invasion by *Shigella flexneri*. Clin. Infect. Dis. **28** (1999) 466-75.
3. Cirillo, D. M., Valdivia, R. H., Monack, D. M., and Falkow, S. Macrophage-dependent induction of the *Salmonella* pathogenicity island 2 type III secretion system and its role in intracellular survival. Mol. Microbiol. **30** (1998) 175-88.
4. Marra, A., and Isberg, R. R. Invasin-dependent and invasin-independent pathways for translocation of *Yersinia pseudotuberculosis* across the Peyer's patch intestinal epithelium. Infect. Immun. **65** (1997) 3412-21.
5. Pepe, J. C., and Miller, V. L. *Yersinia enterocolitica* invasin: a primary role in the initiation of infection. Proc. Nat. Acad. Sci. USA **90** (1993) 6473-7.
6. Pepe, J. C., Wachtel, M. R., Wagar, E., and Miller, V. L. Pathogenesis of defined invasion mutants of *Yersinia enterocolitica* in a BALB/c mouse model of infection. Infect. Immun. **63** (1995) 4837-48.
7. Bottone, E. J. *Yersinia enterocolitica*: the charisma continues. Clin. Microbiol. Rev. **10** (1997) 257-76.
8. Santoro, M. J., Chen, Y. K., Seid, N. S., Abdulian, J. D., and Collen, M. J. *Yersinia enterocolitica* liver abscesses unmasking idiopathic hemochromatosis. J. Clin. Gastroenterol. **18** (1994) 253-4.
9. Isberg, R. R., and Leong, J. M. Multiple beta 1 chain integrins are receptors for

invasin, a protein that promotes bacterial penetration into mammalian cells. Cell. **60** (1990) 861-71.

10. Hynes, R. O. Integrins: Versatility, Modulation, and Signaling in Cell Adhesion. Cell. **69** (1992) 11-25.

11. Clark, M. A., Hirst, B. H., and Jepson, M. A. M-cell surface beta 1 integrin expression and invasin-mediated targeting of *Yersinia pseudotuberculosis* to mouse Peyer's patch M cells. Infect. Immun. **66** (1998) 1237-43.

12. Cornelis, G. R., Boland, A., Boyd, A. P., Geuijen, C., Iriarte, M., Neyt, C., Sory, M. P., and Stainier, I. The virulence plasmid of *Yersinia*, an antihost genome. Microbiol. Mol. Biol. Rev. **62** (1998) 1315-52.

13. Persson, C., Carballeira, N., Wolf-Watz, H., and Fallman, M. The 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** (1997) 2307-18.

14. Black, D. S., and Bliska, J. B. Identification of p130Cas as a substrate of *Yersinia* YopH (Yop51), a bacterial protein tyrosine phosphatase that translocates into mammalian cells and targets focal adhesions. EMBO J. **16** (1997) 2730-44.

15. Rosqvist, R., Persson, C., Hakansson, S., Nordfeldt, R., and Wolf-Watz, H. Translocation of the *Yersinia* YopE and YopH virulence proteins into target cells is mediated by YopB and YopD. Contrib. Microbiol. Immunol. **13** (1995) 230-4.

16. Agin, T. S., and Wolf, M. K. Identification of a family of intimins common to *Escherichia coli* causing attaching-effacing lesions in rabbits, humans, and swine. Infect. Immun. **65** (1997) 320-6.

17. Kenny, B., and Finlay, B. B. Intimin-dependent binding of enteropathogenic *Escherichia coli* to host cells triggers novel signaling events, including tyrosine phosphorylation of phospholipase C-gamma1. Infect. Immun. **65** (1997) 2528-36.

18. Leong, J. M., Fournier, R. S., and Isberg, R. R. Identification of the integrin binding domain of the *Yersinia pseudotuberculosis* invasin protein. EMBO J. **9** (1990) 1979-89.

19. Leong, J. M., Morrissey, P. E., Marra, A., and Isberg, R. R. An aspartate residue of the *Yersinia pseudotuberculosis* invasin protein that is critical for integrin binding. EMBO J. **14** (1995) 422-31.

20. Leahy, D. J., Aukhil, I., and Erickson, H. P. 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. Cell. **84** (1996) 155-64.

21. Saltman, L. H., Lu, Y., Zaharias, E. M., and Isberg, R. R. A region of the *Yersinia pseudotuberculosis* invasin protein that contributes to high affinity binding to integrin receptors. J. Biol. Chem. **271** (1996) 23438-44.

22. Aota, S., Nomizu, M., and Yamada, K. M. The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function. J. Biol. Chem. **269** (1994) 24756-61.

23. Hamburger, Z. A., Brown, M. S., Isberg, R. R., and Bjorkman, P. J. Crystal structure of invasin: a bacterial integrin-binding protein. Science. **286** (1999) 291-295.

24. Dersch, P., and Isberg, R. R. A region of the *Yersinia pseudotuberculosis* invasin protein enhances integrin-mediated uptake into mammalian cells and promotes

self-association. EMBO J. **18** (1999) 1199-213.

25. Weis, W. I., Taylor, M. E., and Drickamer, K. The C-type lectin superfamily in the immune system. Immunol. Rev. **163** (1998) 19-34.

26. Tran Van Nhieu, G., and Isberg, R. R. The *Yersinia pseudotuberculosis* invasin protein and human fibronectin bind to mutually exclusive sites on the alpha 5 beta 1 integrin receptor. J. Biol. Chem. **266** (1991) 24367-75.

27. Rankin, S., Isberg, R. R., and Leong, J. M. The integrin-binding domain of invasin is sufficient to allow bacterial entry into mammalian cells. Infect. Immun. **60** (1992) 3909-12.

28. Tran Van Nhieu, G., and Isberg, R. R. Bacterial internalization mediated by beta 1 chain integrins is determined by ligand affinity and receptor density. EMBO J. **12** (1993) 1887-95.

29. DeStrooper, B., Van Leuven, F., Carmeliet, G., Van Den Berghe, H., and Cassiman, J.-J. Cultured human fibroblasts contain a large pool of precursor beta 1 integrin but lack an intracellular pool of mature subunit. Eur. J. Biochem. **199** (1991) 25-33.

30. Reszka, A. A., Hayashi, Y., and Horwitz, A. F. Identification of amino acid sequences in the integrin beta 1 cytoplasmic domain implicated in cytoskeletal association. J. Cell. Biol. (1992) 1321-30.

31. Otey, C. A., Vasquez, G. B., Burridge, K., and Erickson, B. W. Mapping of the alpha-actinin binding site within the beta 1 integrin cytoplasmic domain. J. Biol. Chem. **268** (1993) 21193-7.

32. Schaller, M. D., Otey, C. A., Hildebrand, J. D., and Parsons, J. T. Focal adhesion

kinase and paxillin bind to peptides mimicking beta integrin cytoplasmic domains. J. Cell. Biol. **130** (1995) 1181-7.

33. Zhang, X. A., and Hemler, M. E. Interaction of the integrin beta 1 cytoplasmic domain with ICAP-1 protein. J. Biol. Chem. **274** (1999) 11-9.

34. Tran Van Nhieu, G., Krukonis, E. S., Reszka, A. A., Horwitz, A. F., and Isberg, R. R. Mutations in the cytoplasmic domain of the integrin beta 1 chain indicate a role for endocytosis factors in bacterial internalization. J. Biol. Chem. **271** (1996) 7665-72.

35. Rosenshine, I., and Finlay, B. B. Exploitation of host signal transduction pathways and cytoskeletal functions by invasive bacteria. Bioessays. **15** (1993) 17-24.

36. Bliska, J. B., Guan, K. L., Dixon, J. E., and Falkow, S. Tyrosine phosphate hydrolysis of host proteins by an essential *Yersinia* virulence determinant. Proc. Nat. Acad. Sci. USA. **88** (1991) 1187-91.

37. Fu, Y., and Galan, J. E. The *Salmonella typhimurium* tyrosine phosphatase SptP is translocated into host cells and disrupts the actin cytoskeleton. Mol. Microbiol. **27** (1998) 359-68.

38. Fu, Y., and Galan, J. E. A *Salmonella* protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. Nature. **401** (1999) 293-7.

39. Rosqvist, R., Magnusson, K. E., and Wolf, W. H. Target cell contact triggers expression and polarized transfer of *Yersinia* YopE cytotoxin into mammalian cells. EMBO J. **13** (1994) 964-72.

40. Schlaepfer, D. D., Broome, M. A., and Hunter, T. Fibronectin-stimulated

signaling from a focal adhesion kinase-c-Src complex: involvement of the Grb2, p130cas, and Nck adaptor proteins. Mol. Cell. Biol. **17** (1997) 1702-13.

41. Schaller, M. D., Hildebrand, J. D., Shannon, J. D., Fox, J. W., Vines, R. R., and Parsons, J. T. Autophosphorylation of the focal adhesion kinase, pp125FAK, directs SH2-dependent binding of pp60src. Mol. Cell. Biol. **14** (1994) 1680-8.

42. Ilic, D., Damsky, C. H., and Yamamoto, T. Focal adhesion kinase: at the crossroads of signal transduction. J. Cell. Sci. **110** (1997) 401-7.

43. Ilic, D., Furuta, Y., Kanazawa, S., Takeda, N., Sobue, K., Nakatsuji, N., Nomura, S., Fujimoto, J., Okada, M., and Yamamoto, T. Reduced cell motility and enhanced focal adhesion contact formation in cells from FAK-deficient mice. Nature. **377** (1995) 539-44.

44. Alrutz, M. A., and Isberg, R. R. Involvement of focal adhesion kinase in invasin-mediated uptake. Proc. Natl. Acad. Sci. USA. **95** (1998) 13658-63.

45. Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M., and Aktories, K. Glucosylation of Rho proteins by *Clostridium difficile* toxin B. Nature. **375** (1995) 500-3.

46. Tapon, N., and Hall, A. Rho, Rac and Cdc42 GTPases regulate the organization of the actin cytoskeleton. Curr. Opin. Cell Biol. **9** (1997) 86-92.

47. Welch, M. D. The world according to Arp: regulation of actin nucleation by the Arp2/3 complex. Trends Cell. Biol. **9** (1999) 423-427.

48. Miki, H., Suetsugu, S., and Takenawa, T. WAVE, a novel WASP-family protein involved in actin reorganization induced by Rac. EMBO J. **17** (1998) 6932-41.

49. Locher, K. P., Rees, B., Koebnik, R., Mitschler, A., Mouliner, L., Rosenbusch, J.

P., and Moras, D. Transmembrane signaling across the ligand-gated FhuA receptor: Crystal structures of free and ferrichrome-bound states reveal allosteric changes. Cell. **95** (1998) 771-778.

50. Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K., and Welte, W. Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science. **82** (1998) 2215-20.

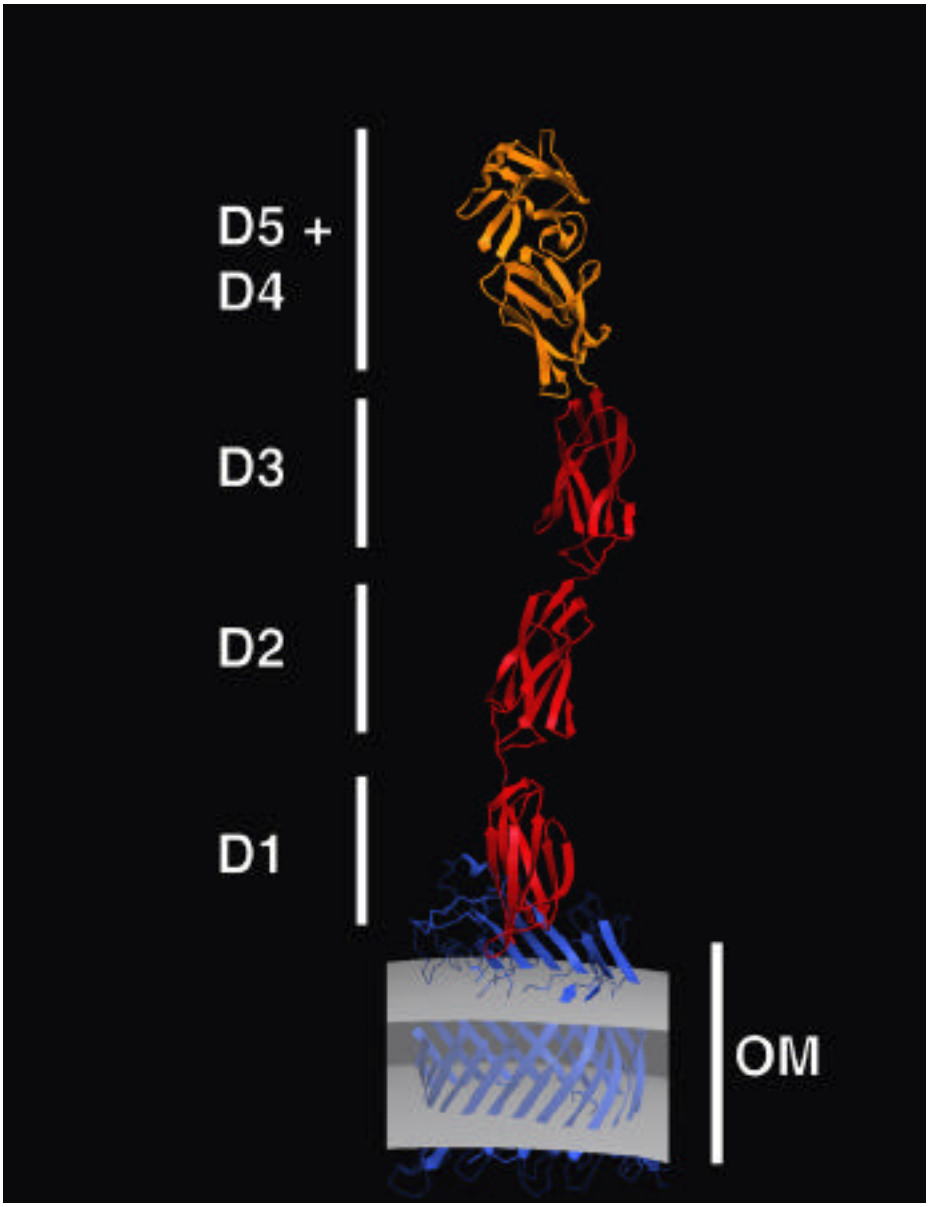


Figure 1. Isberg et al.

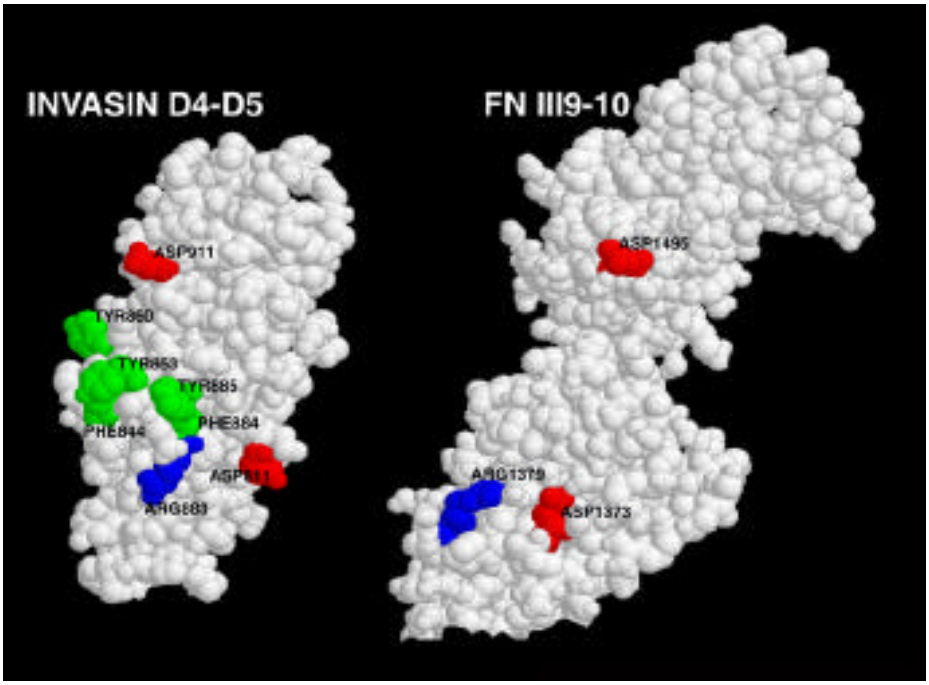


Figure 2. Isberg et al.

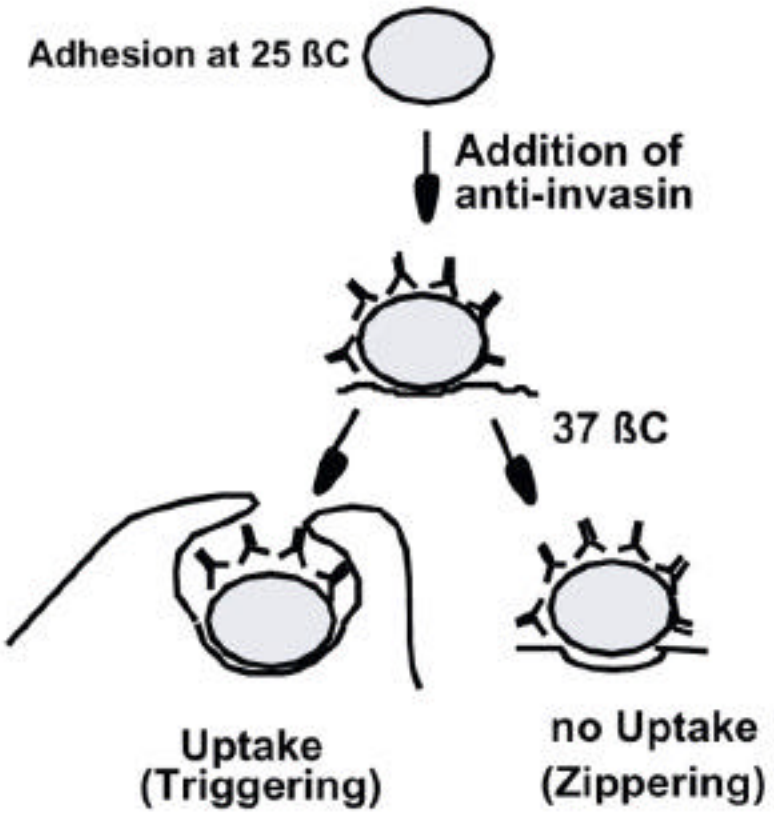


Figure 3. Isberg et al.

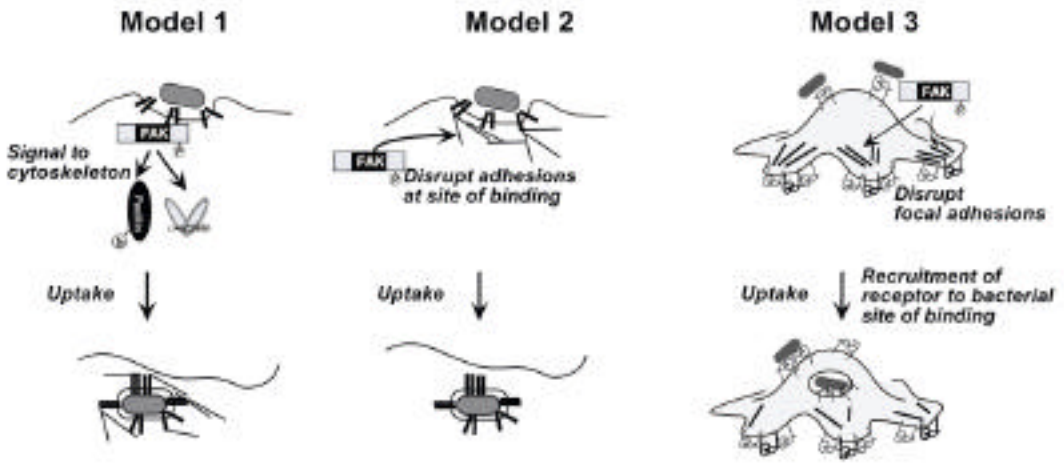


Figure 4. Isberg et al.