Engineered variants of InIB with an additional leucine-rich repeat discriminate between physiologically relevant and packing contacts in crystal structures of the InIB:MET complex

Hartmut H. Niemann,^{1,2*} Ermanno Gherardi,³ Willem M. Bleymüller,¹ and Dirk W. Heinz⁴

¹Department of Chemistry, Bielefeld University, 33501 Bielefeld, Germany

²BioChemTech, Center for Biotechnology (CeBiTec), Bielefeld University, 33501 Bielefeld, Germany

³Division of Immunology and General Pathology, Department of Molecular Medicine, University of Pavia, 27100 Pavia, Italy

⁴Department of Molecular Structural Biology, Helmholtz Center for Infection Research, 38124 Braunschweig, Germany

Received 31 May 2012; Revised 1 August 2012; Accepted 6 August 2012

DOI: 10.1002/pro.2142

Published online 10 August 2012 proteinscience.org

Abstract: The physiological relevance of contacts in crystal lattices often remains elusive. This was also the case for the complex between the invasion protein internalin B (InIB) from *Listeria monocytogenes* and its host cell receptor, the human receptor tyrosine kinase (RTK) MET. InIB is a MET agonist and induces bacterial host cell invasion. Activation of RTKs generally involves ligand-induced dimerization of the receptor ectodomain. The two currently available crystal structures of the InIB:MET complex show the same arrangement of InIB and MET in a 1:1 complex, but different dimeric 2:2 assemblies. Only one of these 2:2 assemblies is predicted to be stable by a computational procedure. This assembly is mainly stabilized by a contact between the Cap domain of InIB from one and the Sema domain of MET from another 1:1 complex. Here, we probe the physiological relevance of this interaction. We generated variants of the leucine-rich repeat (LRR) protein InIB by inserting an additional repeat between the first and the second LRR. This should allow formation of the 1:1 complex but disrupt the potential 2:2 complex involving the Cap-Sema contact due to steric distortions. A crystal structure of one of the engineered proteins showed that it folded properly. Binding affinity to MET was comparable to that of wild-type InIB. The InIB variant induced MET phosphorylation and cell scatter like wild-type InIB. These results suggest that the Cap-Sema interaction is not physiologically relevant and support the previously proposed assembly, in which a 2:2 InIB:MET complex is built around a ligand dimer.

Keywords: c-MET; crystal packing; crystal structure; dimerization; internalin B; leucine-rich repeat; protein engineering; protein-protein interaction; pseudo-merohedral twinning; receptor tyrosine kinase

Grant sponsor: Deutsche Forschungsgemeinschaft; Grant number: NI 694/3-1; Grant sponsor: Priority Program 1150 (SPP1150 "Signal Pathways to the Cytoskeleton and Bacterial Pathogenesis").

*Correspondence to: Hartmut H. Niemann, Department of Chemistry, Bielefeld University, PO Box 10 01 31, 33501 Bielefeld, Germany. E-mail: Hartmut.Niemann@uni-bielefeld.de

Introduction

Intermolecular contacts in crystal structures of proteins that form unstable, low-affinity oligomers may or may not represent the physiologically relevant contacts. This ambiguity can lead to different conclusions with respect to the biologically relevant assembly as exemplified by the controversy about

the correct quaternary structure of the fibroblast growth factor receptor (FGFR) ectodomain in complex with its ligand FGF.^{1–4} FGFR belongs to the superfamily of receptor tyrosine kinases (RTKs), single-pass transmembrane proteins with an extracellular ligand-binding and an intracellular catalytic domain.⁵ Early on it was suggested that RTKs are activated by ligand induced dimerization of their ectodomains, facilitating trans-phosphorylation of their tyrosine kinase domains^{6–9} and a number of crystal structures of ligand-receptor complexes have revealed 2:2 complexes for several RTK families. These structures, in concert with biochemical evidence, show that different ligands cause receptor dimerization by diverse mechanisms.^{10,11}

The proteins MET and RON form a distinct RTK family. MET is the receptor for hepatocyte growth factor/scatter factor (HGF/SF) and transduces signals essential during vertebrate development and wound healing.¹² MET activation results in complex cellular responses comprising, among others, cell division, cell migration, and morphogenesis and aberrant activation of MET can contribute to tumor progression.¹³ Furthermore MET signaling is involved in infection by the Gram-positive pathogen *Listeria monocytogenes* resulting in bacterial invasion of a wide variety of normally non-phagocytic cells.^{14,15} The bacterial ligand responsible for MET activation is the surface protein InlB that binds MET with high affinity. The cellular response to stimulation by soluble InlB resembles that to treatment with HGF/SF in most aspects.^{16–18}

InlB belongs to the internalin family, a group of proteins from *L. monocytogenes* that share an internalin domain consisting of a variable number of leucine-rich repeats (LRRs) flanked N-terminally by a helical Cap domain and C-terminally by an immunoglobulin-like inter-repeat (IR) region.¹⁹ Leucine-rich repeats form typical protein–protein interaction domains. In the case of InlB a fragment consisting of the Cap and the LRR region (InlB₂₄₁) is sufficient to bind the MET receptor with high affinity.^{14,17} Additional C-terminal domains of InlB, the B-repeat and three GW modules, do not interact with MET but enhance MET signaling.^{14,20–22}

Cellular assays have shown that dimerization of the receptor ectodomain is sufficient for MET activation (reviewed in Ref. ¹¹). The structural basis for MET dimerization by HGF/SF has so far largely remained elusive and different models were proposed.^{23,24} The minimal InlB fragment that can activate MET as a soluble, monomeric protein is the internalin domain (Cap, LRR, IR) of InlB, also called InlB₃₂₁.²¹ This stable and structurally well characterized fragment has been crystallized with MET in two different crystal forms.¹⁷ Both crystal forms contain two 1:1 complexes in the asymmetric unit. The structure of the high affinity 1:1 InlB:MET complex

is essentially identical in both crystal forms. However, these 1:1 complexes assemble into different 2:2 complexes. Here, we experimentally scrutinize the significance of a potential 2:2 complex that was predicted by computational methods to form a stable assembly stabilized mainly by interactions involving the InlB Cap domain. We used protein engineering to generate an InlB variant that allowed formation of the 1:1, but not the 2:2 complex by inserting an additional LRR between the first and the second LRR of wild-type InlB. The resulting protein folded properly and supported MET activation, suggesting that the 2:2 InlB:MET complex stabilized by the Cap-Sema interaction is not the signaling competent InlB:MET complex present on cells. This finding is consistent with recent data showing that the signaling-active 2:2 complex is stabilized by a contact on the convex side of the InlB LRR.²⁵

Results

Different dimeric InlB:MET assemblies in two crystal forms

There are two published crystal forms (crystal form I, PDB-ID 2uzx; crystal form II, 2uzy) of InlB₃₂₁ [Fig. 1(A)] in complex with a large part of the MET extracellular domain (Sema+PSI+Ig1+Ig2 = MET₇₄₁; [Fig. 1(B)]). Each crystal form contains two InlB:MET complexes per asymmetric unit, which are related by two-fold non-crystallographic symmetry. The four crystallographically independent 1:1 complexes are structurally very similar, but the arrangement of potential 2:2 complexes is radically different and mediated by different interactions in the two structures [Fig. 1(C–E)]. To address, which of these different 2:2 assemblies may bear physiological relevance, we evaluated the results of the Protein Interfaces, Surfaces and Assemblies (PISA) server.²⁶ The PISA server predicted no stable 2:2 complex for crystal form II. For crystal form I, however, it suggested that the tetrameric 2:2 complex in the asymmetric unit formed a stable assembly. In this potential 2:2 complex, each InlB molecule contacts two MET molecules, and consequently each receptor is in contact with both ligands, forming a circular arrangement [Fig. 1(C)]. Transition from a 1:1 to this 2:2 complex would bury about 3700 Å² of surface area. The main contacts (some 2900 Å² of buried surface area) contributing to this assembly are formed between the N-terminal region of InlB from one 1:1 complex and the Sema domain of MET from the other 1:1 complex. InlB contacts the inner opening of the β-propeller toroid on the “bottom face” that harbors the loops connecting strands a–b and c–d of each propeller blade (Fig. 2). The main contact is formed between the tip of the InlB Cap domain (residues Val40, Pro41, Pro43, Gln46, Asp50, Lys58, Lys64, Ser65, Thr67, Asp68, Ala69) and

Figure 1. Crystal structures of InIB₃₂₁, MET₇₄₁ and arrangement of their potential 2:2 complexes. A: InIB₃₂₁ with the Cap shown in green, the LRR region in cyan and the IR region in dark blue. B: MET₇₄₁ with the Sema domain in yellow, the PSI domain in orange, the Ig1 domain in red and the Ig2 domain in gray. C–E: Twofold symmetric 2:2 InIB₃₂₁/MET₇₄₁ assemblies present in crystal form I and II viewed along the non-crystallographic twofold axis (top view) and rotated by 90° (side view), so that the twofold axis is vertical and the membrane plane lies horizontally below the complexes. The position of the C-termini of both MET molecules (residue 741) is indicated by a pink sphere. C: Asymmetric unit of crystal form I (2uzx). This assembly is predicted to be stable by the PISA server. Dimerization of two 1:1 InIB:MET complexes is mainly stabilized by contacts between the N-terminus of InIB₃₂₁ of one and the MET Sema domain of the other 1:1 complex. D: Asymmetric unit of crystal form II (2uzy) as deposited in the PDB. Contacts stabilizing the 2:2 assembly are formed between the Sema domain of one and the Ig2 domain of the second MET molecule. E: Twofold symmetric 2:2 InIB:MET complex from crystal form II (2uzy) that could form the content of the asymmetric unit, if another origin was chosen. Contacts stabilizing this 2:2 assembly are formed between the convex side of two InIB molecules and between the Ig2 domains of two MET molecules.

Figure 2. Details of the Cap-Sema contact whose importance is addressed in this work. A: The bottom face of the seven-bladed β -propeller formed by the MET Sema domain is shown with each blade in a different color. Blades are numbered, and strands in blade 1 labeled with lower case letters. Residues involved in the contact with the InlB Cap domain on blades 1, 2, 3, and 5 are shown as sticks. B: Contact formed between the InlB Cap and the MET Sema domain from one 1:1 InlB:MET complex and the MET Sema domain from the other 1:1 InlB:MET complex. Hydrogen bonds and salt bridges are shown as dashed purple lines.

blade1 (residues His58, Glu59), blade 2 (Val122, Asp123, Thr124, Tyr125, Asp128) blade 3 (Ser187, Val188, Lys189), and blade 5 (Lys324, Pro325, Leu337, Asn338) of the MET propeller. Further minor contributions arise from contacts between the loops connecting the β -strand to the helical region of LRR1 and LRR2 in InlB and blades 6 and 7 of the MET propeller. Direct ligand–ligand contacts do not occur and direct receptor–receptor contacts are negligible.

Rationale for generation of InlB variants

To test whether this 2:2 assembly from crystal form I is physiologically relevant, we aimed at preventing its formation. To this end, we generated an InlB variant that should still be able to form the high affinity 1:1 complex leaving the contact between the InlB LRR and MET Ig1 largely unchanged, while formation of the 2:2 assembly would be impaired. The interface between the N-terminal region of InlB and the MET Sema domain involves several hydrogen bonds to the InlB backbone and is relatively open [Fig. 2(B)]. This suggested that single amino acid substitutions in the InlB Cap would not lead to disruption of this contact with certainty. A deletion of the Cap would have been ideal, but was not experimentally feasible. We cloned a Cap-less version of InlB₃₂₁ containing InlB residues 77–321 and expressed it as GST-fusion in *E. coli*. Under the same expression conditions as those for wild-type GST-InlB₃₂₁, the GST-fusion of Cap-less InlB₃₂₁ was insoluble and formed inclusion bodies (data not shown). Therefore, we created an InlB variant with a more pronounced difference to wild-type InlB by inserting an additional LRR between the first and the second LRR of InlB (InlB +1LRR; Fig. 3). This does not change the Cap structure itself, but it shifts the Cap relative to the high affinity MET binding site in the LRR of InlB [Figs. 3(A,B) and 4(C)]. As a consequence formation of the circular 2:2 complex from two 1:1 InlB:MET complexes will no longer be possible due to distortion of its geometry [Fig. 3(B)]. When, for example, two 1:1 MET:InlB+1LRR complexes are joined together so that one Cap-Sema contact retains its original geometry [interface 2 in Fig. 3(B) on the right], the second Cap-Sema contact [interface 4 in Fig. 3(B) right] cannot form. Using the crystal structure of InlB+1LRR (described below) to model this showed that residues of the Cap domain of the second InlB+1LRR molecule would be shifted up to 8 Å relative to the real geometry of the Cap-Sema contact. This would result in severe overlap involving side-chain/side-chain, side-chain/main-chain, main-chain/side-chain and even main-chain/main-chain clashes between the Sema domain of the first MET and the Cap domain of the second InlB+1LRR molecule.

Strategy of engineering InlB variants

We generated two versions in which the inserted LRR differed slightly in sequence [Fig. 3(C)]: The first version (InlB+1LRRa) was designed based on a preliminary low-resolution structure of crystal form I and the sequence of the inserted LRR was based on a consensus of residues in LRRs from *Listeria* internalins. After collection of higher resolution data and final refinement of crystal form I, it became apparent, that three surface residues of LRR1 exposed on the concave side (Gln80, Asn84, and

Figure 4. Crystal structure of InlB₃₂₁+1LRRa. A: Structure of InlB₃₂₁+1LRRa. Color coding is as in Figure 1(A), the inserted LRR is shown in red. B: Electron density (2Fo-Fc map contoured at 1.0 r.m.s.d.) and stick model of the inserted LRR with carbon atoms of the conserved consensus residues colored yellow. C: Overlay of InlB₃₂₁+1LRRa and complete InlB₃₂₁wt (1h6t; shown in light brown) performed on the C-terminal segment (residues 99–321 of wild-type). The shift of the Cap domain along the long axis of the molecule becomes apparent. D: InlB₃₂₁+1LRRa overlaid separately with both the N-terminal (residues 36–98) and the C-terminal segment of wild-type InlB₃₂₁ (both shown in light brown).

compared with a total curvature of 77.3° for seven repeats in wild-type as calculated with the ANGULATOR server.²⁷ The Cap is shifted by some 6 Å relative to the LRR and IR regions with the shift parallel to the long axis of the molecule accounting for some 4 Å [Fig. 4(C)]. The structure N-terminal and C-terminal of the inserted LRR shows hardly any deviations from respective regions of wild-type InlB₃₂₁ [Fig. 4(D)]. The root mean square deviation (r.m.s.d.) in the position of backbone atoms between InlB₃₂₁-wt (PDB ID 1h6t²⁸) and any of the three chains of InlB₃₂₁+1LRRa is some 0.4 Å for Cap and LRR1 (amino acids 36–98) and some 0.8 Å for the C-terminal part comprising the remaining LRRs and the IR region (amino acids 99–321 of InlB₃₂₁-wt and 121–343 of InlB₃₂₁+1LRRa). The largest deviations

localize the tip of the IR region (residues 243–255 and 283–295 of wild-type InlB; [Fig. 4(C,D)]) and can be attributed to inherent flexibility rather than to the insertion of the additional LRR.

MET-binding properties

Formation of the complex between the MET ectodomain and InlB₃₂₁+1LRRb was assayed by analytical gel filtration [Fig. 5(A)]. The elution profile of the MET:InlB₃₂₁+1LRRb complex was compared to that of the MET ectodomain on its own and in complex with wild-type InlB₃₂₁. The behavior of MET in complex with InlB₃₂₁+1LRRb was virtually identical to the complex with wild-type InlB₃₂₁. In both cases a larger peak area and a shift to smaller elution volumes (higher molecular weight) were observed. The small magnitude of these shifts upon binding of InlB₃₂₁ or InlB₃₂₁+1LRRb to MET is compatible with formation of 1:1 but not 2:2 complexes.²⁹ The MET-binding ability of the two InlB variants was further tested in an enzyme linked immunosorbent assay (ELISA) using immobilized MET ectodomain and soluble glutathione-S-transferase (GST)-fusion protein of the InlB₃₂₁ variants [Fig. 5(B)]. Binding affinity of both InlB₃₂₁+1LRRa and InlB₃₂₁+1LRRb to MET was very similar to that of wild-type InlB₃₂₁, confirming that MET binding was not significantly altered by insertion of an additional LRR between LRR1 and LRR2 of InlB.

The InlB internalin domain with an inserted LRR stimulates MET phosphorylation

Next we tested, whether insertion of the additional LRR impeded the ability of the engineered protein to activate the MET receptor on cells. The isolated monomeric internalin domain of InlB (i.e., the construct InlB₃₂₁) is sufficient to stimulate phosphorylation of MET and its downstream target ERK. Here, we used ERK phosphorylation as readout for MET activation. As negative controls, we used the F104S variant of InlB₃₂₁ that has reduced affinity to MET³⁰ and the shorter construct InlB₂₄₁ that has the same affinity as InlB₃₂₁ but no biological activity.¹⁷ Like wild-type InlB₃₂₁, InlB₃₂₁+1LRRb induced ERK phosphorylation in Vero cells at a concentration of 10^{−7} M, whereas the negative control remained inactive also at a ten-fold higher concentration [Fig. 5(C)].

Full-length InlB with an inserted LRR induces cell scattering

Despite its ability to stimulate MET and ERK phosphorylation, InlB₃₂₁ does not induce a cellular phenotype like cell motility that is induced by full-length InlB.^{14,17,25} To test whether insertion of an additional LRR has an influence on the stimulation of cell motility, we produced InlB-fl+1LRRb. Stimulation of cell motility was tested in a colony scatter

assay with Madin-Darby canine kidney (MDCK) cells. Just like wild-type full-length InlB, InlB-fl+1LRRb was able to induce the dispersal of MDCK cell colonies down to a concentration of 10^{-10} M (Fig. 6). Thus, the ability of the full-length protein to evoke a cellular phenotype is not altered by insertion of the additional LRR in InlB.

Discussion

Our aim was to assess whether a crystal contact involving mainly the Cap domain of InlB and the Sema domain of the MET receptor may stabilize a putative signaling-active 2:2 ligand:receptor com-

plex, as suggested by the quaternary structure analysis of the PISA server. As the contact is loosely packed, simple substitutions in the InlB cap domain were unlikely to disrupt the interaction with MET with certainty. Therefore, more drastic alterations of InlB seemed necessary in order to scrutinize the functional importance of the crystal contact between Cap and Sema domain. Deletion of the Cap domain was not possible, because the resulting Cap-less InlB₃₂₁ was insoluble and formed inclusion bodies. This is not surprising, as LRR proteins generally have specialized capping structures that are thought to serve a structural role by shielding the hydrophobic core of the LRRs from solvent.³¹ In InlB the Cap domain was additionally shown to be important for folding. InlB folds along a polarized pathway from the N- to the C-terminus with the helical Cap serving as fast-folding nucleus onto which the LRRs are added consecutively.³² Therefore, we devised an alternative strategy to assess whether the Cap-Sema contact in the 2:2 complex of crystal form I is relevant for InlB-mediated MET receptor activation. The insertion of an additional LRR retains the N-terminal Cap as such but changes the orientation of the Cap relative to the two InlB surfaces that interact with MET. The resulting steric distortions will prevent formation of the potential 2:2 complex found in crystal form I.

The modular architecture of InlB facilitates protein design

Because of their modular design, LRR proteins are amenable to targeted manipulation and can be tailored to meet specific needs. Examples include the construction of synthetic libraries of designed LRR proteins as binding molecules intended to replace

Figure 5. Biochemical and biological activity of InlB₃₂₁+1LRR. **A:** Elution profiles from gel filtration are shown for the complete MET ectodomain and its complexes with wild-type InlB₃₂₁ or InlB₃₂₁+1LRRb. The behavior of the MET:InlB₃₂₁+1LRRb complex is virtually identical to that of the complex with wild-type InlB₃₂₁. InlB₃₂₁ and InlB₃₂₁+1LRRb were used in molar excess giving rise to the additional peaks at around 15.2 mL. **B:** InlB₃₂₁+1LRRa and InlB₃₂₁+1LRRb bind to the MET ectodomain with virtually the same affinity as InlB₃₂₁-wt. The complete MET ectodomain (MET₉₂₈) was immobilized on ELISA plates and incubated with increasing concentrations of GST-tagged InlB₃₂₁ variants. Binding was detected with an HRP-coupled anti-GST antibody. Three independent experiments, in which each protein was tested in duplicate or triplicate, were normalized and then averaged. **C:** The ability of the engineered InlB₃₂₁ variants to induce phosphorylation of ERK, which is a downstream signaling target of MET, was assayed in Vero cells. HGF/SF was used as positive control. InlB₂₄₁ and the MET-binding deficient mutant F104S of InlB₃₂₁ were used as negative control.

Figure 6. Biological activity of engineered full-length InlB with an additional LRR. The ability of full-length InlB+1LRRb to induce cell motility was tested in an MDCK cell colony scatter assay. HGF/SF was used as positive control, medium without ligand as negative control. Like wild-type full-length InlB, InlB-fl+1LRRb stimulated colony dispersal at a concentration of 100 pM but showed no activity at 10 pM.

antibodies,³³ or the use of hybrid LRR proteins for crystallization of toll like receptors.³⁴ In very recent work published during preparation of this manuscript, the InlB Cap domain was grafted onto LRRs of variable lymphocyte receptors in order to generate a binding scaffold with favorable properties.³⁵ We did not place the insertions between Cap and LRR1 because the hydrophobic core of the first LRR is packed more loosely than that of the other LRRs due to the bulky side chain of Phe53 from the Cap pointing towards LRR1. Positions 5 and 7 of LRR1 are occupied by isoleucine and alanine, respectively, instead of leucine as in all other LRRs, setting LRR1 apart from the other LRRs. Thus, the Cap and LRR1 had to be treated as a structural unit in InlB. The crystal structure of InlB₃₂₁+1LRRa confirmed that InlB tolerates the insertion of an additional

consensus LRR between LRR1 and LRR2 leaving the protein fold intact. The ELISA showed that formation of the 1:1 high affinity complex with MET is not impeded, confirming the validity of the design strategy. This work is thus another example that LRR proteins lend themselves to a rational approach of protein engineering and protein design.

The 2:2 InlB:MET complex stabilized by Cap-Sema contacts is not physiologically relevant

It is often difficult to distinguish fortuitous packing contacts from physiologically relevant contacts based only on a crystal structure without further functional data.³⁶ PISA, which is used for annotation by the PDB and is also available as a web server, identifies macromolecular complexes not solely based on the buried surface area, but as chemically stable associations, that is, those with a positive free energy of dissociation, which additionally takes into account entropic contributions and other factors.³⁶ For InlB and MET, one would expect the signaling active complex to have a 2:2 stoichiometry, as it was found for other RTKs in complex with their ligands, for example, KIT in complex with stem cell factor (SCF),^{37,38} the epidermal growth factor receptor (EGFR) in complex with EGF,^{39,40} or Axl in complex with its ligand Gas6.⁴¹ The fact that two different crystal forms showed two different 2:2 assemblies for the InlB:MET complex suggested that only one or neither of the two crystal forms captured the physiological signaling complex. This is in line with the observation that weak interactions with $K_d > 100 \mu\text{M}$ have a considerable chance of being lost during the course of crystallization, in which case crystal packing will misrepresent macromolecular complexes and interactions.^{36,42} In solution, the complex consisting the InlB internalin domain and the MET ectodomain clearly has 1:1 stoichiometry as shown by different biophysical methods like analytical ultracentrifugation, dynamic light scattering or small angle X-ray and neutron scattering.^{17,29} It was thus unclear, whether any of the 2:2 InlB:MET complexes from crystal forms I or II would represent a functional complex. The reported high success rate of predictions by PISA,³⁶ warranted experimental scrutiny of the assembly suggested to be stable by this successful algorithm. The data presented here, however, clearly establish that the main contact stabilizing this 2:2 assembly, namely the one formed between the InlB Cap and the MET Sema domain, is not involved in MET activation by InlB. Therefore, the 2:2 assembly from crystal form I does not bear any physiological relevance but merely represents crystal packing, emphasizing the pivotal role of mutagenesis experiments in combination with cell-based assays in the analysis of complex crystallographic assemblies.

The C-termini of the two receptor molecules seen in this arrangement point into opposite directions [Fig. 1(C)] and the necessary attachment of both receptors in the same membrane would require a sharp kink in the MET stalk region between Ig-like domain 2 (the last one present in the crystallized construct) and Ig-like domain 3. Such flexibility is not impossible given the structural variability of the MET ectodomain visualized by electron microscopy.²⁴ Crystal form II contains a more plausible 2:2 InlB:MET complex, in which the MET C-termini point into the same direction.²⁵ However, dimerization in this complex is mediated by only a relatively small contact (some 1400 Å² buried surface area) on the convex back-side of InlB [Fig. 1(E)] and computational methods did not identify this as a potentially stable complex. The negative findings presented here corroborate our recent results showing that this 2:2 InlB:MET assembly from crystal form II likely represents the signaling active complex.²⁵ Thus a coherent picture is emerging, in which the mutagenesis data presented here show that the crystallographic InlB:MET complex predicted to be most stable [Fig. 1(A)] is not the biological assembly [Figs. 5(C) and 6] and additional mutagenesis published elsewhere establish that the biological assembly [Fig. 1(E)] is one that the computational methods failed to highlight.²⁵ Interestingly, the signaling active complex of MET and the HGF/SF splice variant NK1 is most likely also built around a twofold symmetric ligand dimer.⁴³ Further studies to address the quaternary structure of the 2:2 InlB:MET complex on cell membranes using site-specific fluorescent labels and distance measurements with Förster resonance energy transfer are under way in our laboratory.

Materials and Methods

Cloning of Cap-less InlB, InlB+1LRRa, and InlB+1LRRb

Cap-less InlB was generated by a PCR-based approach using the vector pETM-30-InlB₃₂₁ (HN04-15) as template. The Cap domain was deleted from this vector so that the tobacco etch virus (TEV) protease recognition site and the first LRR of InlB (starting at residue 77) are separated by a single glycine. InlB₃₂₁+1LRRa was generated by placing an additional consensus LRR (LRRa) with the sequence “NLTSLNLSNNQITDISPIQYLP” between LRR1 and LRR2 of InlB. The insertion was achieved by opening the vector pETM-30-InlB₃₂₁ at the unique EcoRI site in the InlB gene and inserting a double-stranded oligonucleotide with EcoRI-compatible 5' overhangs coding for the sequence “IQYLPNLTSLNLSNNQITDISP” to yield vector HN06-02 (sequence for the sense-strand: “aattcag-tacctgccgaacctgaccagcctgaacctgagcaataaccagattaccgatatcagccc”). InlB₃₂₁+1LRRb (vector HN06-07) was

created in the same way resulting in the additional amino acids “IEYLPNLDQLILNNSIASIVG” (sequence for the sense-strand oligonucleotide “aatt-gaatacctgccgaacctgaccagcctgattctgaacaataacagcattgc-tagcattgtggg”). The equivalent version of full-length InlB+1LRRb was created in the same way with pETM-30-InlB-fl (HN06-04) as target vector resulting in vector HN06-09.

Protein expression and purification

Wild-type full length InlB and wild-type InlB₃₂₁ were purified as described.¹⁷ All variants of InlB were expressed as GST-fusions and purified essentially as described.¹⁷ Briefly, *E. coli* BL21 (DE3) CodonPlus RIL cells were grown with shaking to an OD₆₀₀ of 0.6 in LB medium at 37°C, shifted to 20°C, induced with 1 mM isopropyl-β-D-thiogalactopyranoside and further incubated overnight. Harvested cells were resuspended in phosphate buffered saline (PBS) with benzonase and complete protease inhibitor cocktail and lysed in a French Press. After centrifugation the cleared lysate was applied to glutathione sepharose (GE Healthcare), followed by thorough washing with PBS. Fusion proteins used for ELISAs were eluted with reduced glutathione and further purified by anion exchange (MonoQ, GE Healthcare) chromatography or size exclusion chromatography (Superdex 200, GE Healthcare) if necessary. Proteins used for assays with cells and for crystallization were cleaved from the GST-tag using TEV protease and further purified using anion exchange chromatography (InlB₃₂₁+1LRR) or anion exchange and size exclusion chromatography (full-length InlB+1LRRb).

Crystallization and data collection

InlB₃₂₁+1LRRa was crystallized at 17 mg mL⁻¹ in 25 mM Tris, pH 8, 20 mM NaCl by sitting drop vapor diffusion in 96-well plates at 20°C. Drops were mixed from 2 μL of protein and 2 μL of reservoir solution (2M ammonium sulfate). Needle-shaped crystals were cryoprotected with reservoir solution additionally containing glycerol and flash frozen in liquid nitrogen. Data were collected on a MAR-CCD 225 detector on the EMBL X12 beamline at the DORIS storage ring, DESY, Hamburg.

Data processing, structure determination, and refinement

Diffraction data were processed with the XDS package.⁴⁴ The structure was solved by molecular replacement using the program Phaser⁴⁵ with the appropriate fragments of InlB²⁸ as search models. Errors were corrected and missing residues were added manually in the program Coot.⁴⁶ Diffraction data of InlB₃₂₁+1LRR could be indexed and the structure could be solved and refined reasonably well both in space group P321 with one molecule in

Table I. Data Collection and Refinement Statistics

Data collection statistics	
Space group	C2
Unit cell dimensions [Å], [°]	$a = 179.08$ $b = 102.15$ $c = 69.81$ $\alpha = \gamma = 90.0$ $\beta = 90.234$
Resolution [Å]	20–1.95 (2.00–1.95)
Completeness [%]	97.4 (96.6)
Multiplicity	5.7 (5.7)
Unique reflections	89,100
I/ σ (I)	15.5 (2.85)
R_{meas}	9.3 (84.3)
$R_{\text{mrgd-F}}$	12.3 (68.3)
Molecules per asymm. unit	3
Solvent content (%)	60%
Refinement statistics	
Resolution	19.7–1.93 (1.978–1.928)
Observations work set/test set	87083 (3701)/2016 (70)
R/R_{free} [%]	16.69 (22.2)/20.54 (25.8)
Number of atoms	
Total	8080
Protein/solvent/others	7441/463/176
R.m.s. deviation	
Bonds [Å]	0.015
Angles [°]	1.715
Ramachandran plot	
Residues in favored regions [%]	96.1
Residues in allowed regions [%]	99.8
Residues in disallowed regions [%]	0.22
Twin fractions for operator [%]	
H, K, L	66.4
1/2H – 3/2K, –1/2H – 1/2K, –L	19.2
1/2H + 3/2K, 1/2H – 1/2K, –L	14.4

Numbers in parentheses refer to the highest resolution shell.

the asymmetric unit or in C2 with a 3-fold non-crystallographic rotation axis parallel to the c-axis. Scaling of diffraction data, however, produced significantly better results in C2 than in P321 (C2: $R_{\text{meas}} = 9.3\%$, $I/\sigma(I) = 15.5$; P321: $R_{\text{meas}} = 18.1\%$, $I/\sigma(I) = 10.8$). In C2 R_{free} -flags were chosen using phenix-reflection_file_converter taking the higher symmetry of the trigonal lattice into account.⁴⁷ The structure was refined using Refmac5 version 5.2⁴⁸ in C2 without twinning to values of 20.4% and 25.2% for R and R_{free} , respectively. R versus R statistics⁴⁹ as implemented in phenix.xtriage⁴⁷ of this structure gave for twin law (0.5 h – 1.5 k, –0.5 h – 0.5 k, –l) $R_{\text{twin}}^{\text{obs}} = 15.7\%$ and $R_{\text{twin}}^{\text{calc}} = 26.4\%$ suggesting a maximum likelihood twin fraction of 13.7% and for twin law (0.5 h + 1.5 k, 0.5 h – 0.5 k, –l) $R_{\text{twin}}^{\text{obs}} = 15.5\%$ and $R_{\text{twin}}^{\text{calc}} = 24.6\%$ suggesting a maximum likelihood twin fraction of 13.9%. Automatic twin refinement was carried out in Refmac5 version 5.5 yielding twin fractions of 19.3% for twin law (0.5 h – 1.5 k, –0.5 h – 0.5 k, –l) and 15.3% for twin law (0.5 h + 1.5 k, 0.5 h – 0.5 k, –l) and resulting in a decrease of R from 20.9% to 19.3% and R_{free} from 25.7% to 22.2%,

respectively, without any rebuilding. Repeating the R versus R calculation in phenix.xtriage now resulted in values of $R_{\text{twin}}^{\text{calc}} = 32.3\%$ and $R_{\text{twin}}^{\text{calc}} = 30.1\%$ for twin laws (0.5 h – 1.5 k, –0.5 h – 0.5 k, –l) and (0.5 h + 1.5 k, 0.5 h – 0.5 k, –l), respectively. Refinement was finished with Refmac5 version 5.6 using amplitude based twin refinement, 18 TLS groups as determined by the TLS motion determination server⁵⁰ and automatically generated local NCS restraints resulting in final twin fractions of 19.2% for twin law (0.5 h – 1.5 k, –0.5 h – 0.5 k, –l) and 14.4% for twin law (0.5 h + 1.5 k, 0.5 h – 0.5 k, –l). The structure was checked with Molprobit.⁵¹ Data collection and refinement statistics are given in Table I. Coordinates and structure factors were deposited in the PDB under accession code 4aw4.

Analytical gel filtration

Analytical gel filtration was carried out using a Superdex200 10/300 GL column (GE Healthcare) equilibrated in PBS. 30 μg of the complete MET ectodomain (MET₉₂₈) were run alone or after mixing with a molar excess of InlB₃₂₁ or InlB₃₂₁+1LRRb.

Binding, phosphorylation, and cell scattering assays

Binding, phosphorylation and cell scattering assays were carried out essentially as described.¹⁷

Acknowledgment

The authors are grateful for access to the EMBL X12 beamline at the DORIS storage ring, DESY, Hamburg. They thank Christina Geerds for technical assistance.

References

- Plotnikov AN, Schlessinger J, Hubbard SR, Mohammadi M (1999) Structural basis for FGF receptor dimerization and activation. *Cell* 98:641–650.
- Pellegrini L, Burke DF, von Delft F, Mulloy B, Blundell TL (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature* 407:1029–1034.
- Harmer NJ, Ilag LL, Mulloy B, Pellegrini L, Robinson CV, Blundell TL (2004) Towards a resolution of the stoichiometry of the fibroblast growth factor (FGF)-FGF receptor-heparin complex. *J Mol Biol* 339: 821–834.
- Mohammadi M, Olsen SK, Ibrahimi OA (2005) Structural basis for fibroblast growth factor receptor activation. *Cytokine Growth Factor Rev* 16:107–137.
- Schlessinger J (2000) Cell signaling by receptor tyrosine kinases. *Cell* 103:211–225.
- Yarden Y, Schlessinger J (1987) Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor. *Biochemistry* 26: 1443–1451.
- Yarden Y, Schlessinger J (1987) Self-phosphorylation of epidermal growth factor receptor: evidence for a model of intermolecular allosteric activation. *Biochemistry* 26: 1434–1442.

8. Schlessinger J (1988) Signal transduction by allosteric receptor oligomerization. *Trends Biochem Sci* 13: 443–447.
9. Ullrich A, Schlessinger J (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell* 61: 203–212.
10. Lemmon MA, Schlessinger J (2010) Cell signaling by receptor tyrosine kinases. *Cell* 141:1117–1134.
11. Niemann HH (2011) Structural insights into Met receptor activation. *Eur J Cell Biol* 90:972–981.
12. Birchmeier C, Birchmeier W, Gherardi E, VandeWoude GF (2003) Met, metastasis, motility and more. *Nature Rev* 4:915–925.
13. Boccaccio C, Comoglio PM (2006) Invasive growth: a MET-driven genetic programme for cancer and stem cells. *Nat Rev Cancer* 6:637–645.
14. Shen Y, Naujokas M, Park M, Ireton K (2000) InIB-dependent internalization of *Listeria* is mediated by the Met receptor tyrosine kinase. *Cell* 103:501–510.
15. Seveau S, Pizarro-Cerda J, Cossart P (2007) Molecular mechanisms exploited by *Listeria monocytogenes* during host cell invasion. *Microbes Infect* 9:1167–1175.
16. Li N, Xiang GS, Dokainish H, Ireton K, Elferink LA (2005) The *Listeria* protein internalin B mimics hepatocyte growth factor-induced receptor trafficking. *Traffic* 6:459–473.
17. Niemann HH, Jager V, Butler PJ, van den Heuvel J, Schmidt S, Ferraris D, Gherardi E, Heinz DW (2007) Structure of the human receptor tyrosine kinase met in complex with the *Listeria* invasion protein InIB. *Cell* 130:235–246.
18. Gao X, Lorinczi M, Hill KS, Brooks NC, Dokainish H, Ireton K, Elferink LA (2009) Met receptor tyrosine kinase degradation is altered in response to the leucine-rich repeat of the *Listeria* invasion protein internalin B. *J Biol Chem* 284:774–783.
19. Niemann HH, Schubert WD, Heinz DW (2004) Adhesins and invasins of pathogenic bacteria: a structural view. *Microbes Infect* 6:101–112.
20. Jonquieres R, Pizarro-Cerda J, Cossart P (2001) Synergy between the N- and C-terminal domains of InIB for efficient invasion of non-phagocytic cells by *Listeria monocytogenes*. *Mol Microbiol* 42:955–965.
21. Banerjee M, Copp J, Vuga D, Marino M, Chapman T, van der Geer P, Ghosh P (2004) GW domains of the *Listeria monocytogenes* invasion protein InIB are required for potentiation of Met activation. *Mol Microbiol* 52:257–271.
22. Ebbes M, Bleymuller WM, Cernescu M, Nolker R, Brutschy B, Niemann HH (2011) Fold and function of the InIB B-repeat. *J Biol Chem* 286:15496–15506.
23. Stamos J, Lazarus RA, Yao X, Kirchhofer D, Wiesmann C (2004) Crystal structure of the HGF beta-chain in complex with the Sema domain of the Met receptor. *EMBO J* 23:2325–2335.
24. Gherardi E, Sandin S, Petoukhov MV, Finch J, Youles ME, Ofverstedt LG, Miguel RN, Blundell TL, VandeWoude GF, Skoglund U, Svergun DI (2006) Structural basis of hepatocyte growth factor/scatter factor and MET signaling. *Proc Natl Acad Sci USA* 103: 4046–4051.
25. Ferraris DM, Gherardi E, Di Y, Heinz DW, Niemann HH (2010) Ligand-mediated dimerization of the Met receptor tyrosine kinase by the bacterial invasion protein InIB. *J Mol Biol* 395:522–532.
26. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372: 774–797.
27. Bublitz M, Holland C, Sabet C, Reichelt J, Cossart P, Heinz DW, Bierne H, Schubert WD (2008) Crystal structure and standardized geometric analysis of InIJ, a *Listeria* virulence factor and leucine-rich repeat protein with a novel cysteine ladder. *J Mol Biol* 378:87–96.
28. Schubert WD, Gobel G, Diepholz M, Darji A, Kloer D, Hain T, Chakraborty T, Wehland J, Domann E, Heinz DW (2001) Internalins from the human pathogen *Listeria monocytogenes* combine three distinct folds into a contiguous internalin domain. *J Mol Biol* 312:783–794.
29. Niemann HH, Petoukhov MV, Hartlein M, Moulin M, Gherardi E, Timmins P, Heinz DW, Svergun DI (2008) X-ray and neutron small-angle scattering analysis of the complex formed by the Met receptor and the *Listeria monocytogenes* invasion protein InIB. *J Mol Biol* 377:489–500.
30. Machner MP, Frese S, Schubert WD, Orian-Rousseau V, Gherardi E, Wehland J, Niemann HH, Heinz DW (2003) Aromatic amino acids at the surface of InIB are essential for host cell invasion by *Listeria monocytogenes*. *Mol Microbiol* 48:1525–1536.
31. Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* 11: 725–732.
32. Courtemanche N, Barrick D (2008) The leucine-rich repeat domain of Internalin B folds along a polarized N-terminal pathway. *Structure* 16:705–714.
33. Stumpp MT, Forrer P, Binz HK, Pluckthun A (2003) Designing repeat proteins: modular leucine-rich repeat protein libraries based on the mammalian ribonuclease inhibitor family. *J Mol Biol* 332:471–487.
34. Jin MS, Lee JO (2008) Application of hybrid LRR technique to protein crystallization. *BMB Rep* 41:353–357.
35. Lee SC, Park K, Han J, Lee JJ, Kim HJ, Hong S, Heu W, Kim YJ, Ha JS, Lee SG, Cheong HK, Jeon YH, Kim D, Kim HS (2012) Design of a binding scaffold based on variable lymphocyte receptors of jawless vertebrates by module engineering. *Proc Natl Acad Sci USA* 109: 3299–3304.
36. Krissinel E (2011) Macromolecular complexes in crystals and solutions. *Acta Cryst D* 67:376–385.
37. Liu H, Chen X, Focia PJ, He X (2007) Structural basis for stem cell factor-KIT signaling and activation of class III receptor tyrosine kinases. *EMBO J* 26: 891–901.
38. Yuzawa S, Opatowsky Y, Zhang Z, Mandiyan V, Lax I, Schlessinger J (2007) Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. *Cell* 130:323–334.
39. Garrett TP, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Zhu HJ, Walker F, Frenkel MJ, Hoyne PA, Jorissen RN, Nice EC, Burgess AW, Ward CW (2002) Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* 110:763–773.
40. Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, Kim JH, Saito K, Sakamoto A, Inoue M, Shirouzu M, Yokoyama S (2002) Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* 110:775–787.
41. Sasaki T, Knyazev PG, Clout NJ, Cheburkin Y, Gohring W, Ullrich A, Timpl R, Hohenester E (2006) Structural basis for Gas6-Axl signalling. *EMBO J* 25:80–87.
42. Krissinel E (2010) Crystal contacts as nature's docking solutions. *J Comp Chem* 31:133–143.
43. Tolbert WD, Daugherty J, Gao C, Xie Q, Miranti C, Gherardi E, Woude GV, Xu HE (2007) A mechanistic basis for converting a receptor tyrosine kinase agonist

- to an antagonist. *Proc Natl Acad Sci USA* 104: 14592–14597.
44. Kabsch W (2010) Xds. *Acta Cryst allogr D* 66:125–132.
 45. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Cryst* 40:658–674.
 46. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Cryst allogr D* 66: 486–501.
 47. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst allogr D* 66:213–221.
 48. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst allogr D* 67:355–367.
 49. Lebedev AA, Vagin AA, Murshudov GN (2006) Intensity statistics in twinned crystals with examples from the PDB. *Acta Cryst allogr D* 62:83–95.
 50. Painter J, Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Cryst* 39:109–111.
 51. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst allogr D* 66:12–21.