Thermodynamically reengineering the listerial invasion complex InlA/E-cadherin

Thomas Wollert*, Dirk W. Heinz†, and Wolf-Dieter Schubert*‡

*Molecular Host–Pathogen Interactions, †Division of Structural Biology, Helmholtz Centre for Infection Research, Inhoffenstrasse 7, D-38124 Braunschweig, Germany

Edited by Stephen L. Mayo, California Institute of Technology, Pasadena, CA, and approved July 19, 2007 (received for review March 9, 2007)

Biological processes essentially all depend on the specific recognition between macromolecules and their interaction partners. Although many such interactions have been characterized both structurally and biophysically, the thermodynamic effects of small atomic changes remain poorly understood. Based on the crystal structure of the bacterial invasion protein internalin (InlA) of *Listeria monocytogenes* in complex with its human receptor E-cadherin (hEC1), we analyzed the interface to identify single amino acid substitutions in InlA that would potentially improve the overall quality of interaction and hence increase the weak binding affinity of the complex. Dissociation constants of InlA-variant/hEC1 complexes, as well as enthalpy and entropy of binding, were quantified by isothermal titration calorimetry. All single substitutions indeed significantly increase binding affinity. Structural changes were verified crystallographically at ≤2.0-Å resolution, allowing thermodynamic characteristics of single substitutions to be rationalized structurally and providing unique insights into atomic contributions to binding enthalpy and entropy. Structural and thermodynamic data of all combinations of individual substitutions result in a thermodynamic network, allowing the source of cooperativity between distant recognition sites to be identified. One such pair of single substitutions improves affinity 5,000-fold. We thus demonstrate that rational reengineering of protein complexes is possible by making use of physically distant hot spots of recognition.

long-distance synergy | protein–protein interaction | structure-based reengineering | *Listeria monocytogenes* | internalin

The Gram-positive bacterium *Listeria monocytogenes* causes severe infections in immunocompromised individuals and unborn fetuses (1). As part of its invasion strategy, *L. monocytogenes* is able to breach the intestinal barrier by inducing its own uptake into normally nonphagocytic cells using the invasion protein internalin (InlA) (2). Structurally, InlA consists of an N-terminal cap, a leucine-rich repeat domain (LRR), and an Ig-like interrepeat domain (3, 4), followed by three spacer domains to allow presentation on the cell surface (5). E-cadherin, the most abundant protein in epithelial–cell adherens junctions, is crucial in embryogenesis (6) and in maintaining epithelial integrity (7). It consists of five extracellular Ig-like domains (EC1–5), a transmembrane α-helix, and an intracellular domain linked to the actin cytoskeleton (8). The N-terminal domain of human E-cadherin (hEC1) is responsible for cell–cell contacts (9) and is also the receptor of InlA (10, 11). The crystal structure of the functional domain of InlA in complex with hEC1 revealed that InlA binds hEC1 through the concave face of its LRR domain (5). Despite burying 2,400 Å^2 of solvent-accessible surface area on complex formation, the binding affinity ($K_d = 8 \pm 4$ μM) is rather weak. However, weak affinity does not correlate with low binding specificity, as indicated by a narrow range of EC1 domains of other species recognized by InlA (11).

Manipulating protein–protein interaction surfaces to increase binding affinity or to change binding specificity still represents a major challenge. Essentially two competing approaches exist, based either on selection of efficient binders from large randomly created libraries (12) or on computational design (13). The latter, however, requires comprehensive knowledge of protein recognition and underlying physical mechanisms, which have not yet been fully analyzed or quantified (14). Unresolved questions include (*i*) enthalpic and entropic contributions to binding affinity, (*ii*) the source of enthalpy–entropy compensation phenomena (15), (*iii*) the precise role of water molecules within protein–protein interaction surfaces (16), and (*iv*) cooperativity between adjacent and distant interaction sites (17).

To contribute to a better understanding of protein–protein interactions, we have applied a rational protein design approach without relying on computational methods. Instead, the high-resolution crystal structure of the InlA/hEC1 complex allowed single amino acid substitutions in InlA to be identified that potentially increase its binding affinity for hEC1. We characterized the binding affinity of InlA variants carrying single or double substitutions by isothermal titration calorimetry (ITC) deriving precise enthalpic and entropic contributions to complex formation. We have similarly solved the crystal structures of InlA variants in complex with hEC1, yielding precise structural data on changes introduced through the substitutions. Thermodynamic properties may thus be directly correlated with structural changes, providing unique insights into enthalpic and entropic contribution of single amino acid side chains to macromolecular complex formation and to the cooperative behavior of combinations of single.

Results

Structure-Based Variant Design. Although the crystal structure of the InlA/hEC1 complex (5) (Fig. 1*A*) indicates a comparatively large (18) interaction interface (2,400 Å^2), the binding affinity of InlA for hEC1 at $K_d = 8 \pm 4$ μM is surprisingly weak, presumably because of low surface complementarity and a large number of bridging water molecules. Using the structural information, we designed the following InlA variants to increase the binding affinity of the complex:

Tyr369Ala (Y369A). In uncomplexed InlA, Tyr-369 forms a well ordered stacking interaction with Tyr-347 (Fig. 2*A*). In the complex InlA/hEC1, Asn-27$_{hEC1}$ displaces Tyr-369, causing it to swing around its χ_1 angle away from its stacking interaction with Tyr-347 and displacing Asn-370 and His-392 from a similar stacking interaction with Phe-348. Replacing Tyr-369 by alanine would eliminate this rearrangement of surface residues.

Author contributions: T.W. and W.-D.S. designed research; T.W. performed research; T.W., D.W.H., and W.-D.S. analyzed data; and T.W., D.W.H., and W.-D.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: ITC, isothermal titration calorimetry; hEC1, N-terminal domain of human E-cadherin; LRR, leucine-rich repeat.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2OMT, 2OMU, 2OMX, and 2OMZ).

‡To whom correspondence should be addressed. E-mail: wolf-dieter.schubert@helmholtz-hzi.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0702199104/DC1.

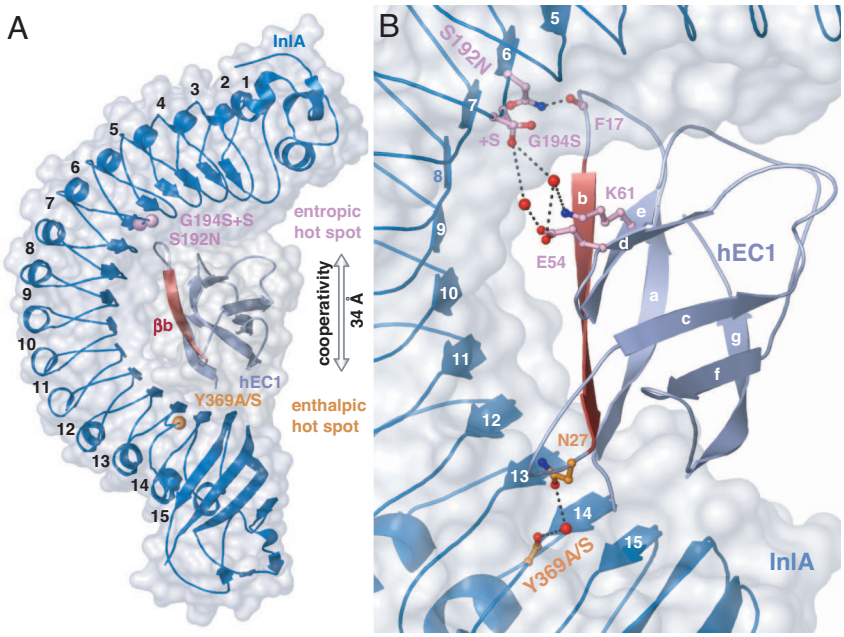

Fig. 1. Overall structure of InlA/hEC1 (dark/light blue). (*A*) Reengineered residues are indicated by spheres. Although separated by 34 Å, combinations of substitutions from entropically (pink) and enthalpically (orange) dominated ''hot spots'' act synergistically by stabilizing β-strand **b** of hEC1 (red). (*B*) Closeup view of the interaction interface. S192N and Y369A/S (ball and stick) stabilize opposite ends of β**b**. G194S+S shortens the distance to residues Glu-54_{hEC1} and Lys-61_{hEC1} (ball and stick) in β**d** and β**e**, respectively. Stabilization is transmitted through β-sheet **bde** to the N terminus of β**b**. All structural figures were prepared by using Pymol.

Tyr369Ser (Y369S). Although an alanine at position 369 eliminates the unfavorable conformation of Tyr-369 (above), its small size and lack of hydrogen bond donor or acceptor groups would prevent it from directly interacting with hEC1. Placing serine at this position should allow direct or water-mediated hydrogen bonds to hEC1.

Ser192Asn (S192N). Ser-192_{InlA} forms a water-mediated hydrogen bond to hEC1. Because a direct hydrogen bond between InlA and hEC1 would increase binding affinity, we replaced Ser-192 by slightly longer asparagine, allowing it to potentially bridge the gap to hEC1.

Gly194Ser+Ser (G194S+S). Compared with other LRRs of InlA, repeat 6 consists of 21 residues instead of the canonical 22. This shortens the loop after the LRR β-strand (5), discontinues the asparagine ladder characteristic of LRR proteins (19), and creates a 7.5-Å-wide hydrophobic water-filled cavity on the surface of InlA. To restore the regular LRR architecture, an additional serine (+S) was introduced after Gly-194, whereas Gly-194 itself was replaced by serine, the most common residue at this position in other LRRs of InlA.

Structural Verification of Predicted Atomic-Scale Changes. InlA-hEC1 complexes were analyzed by x-ray crystallography at ≤2.0-Å resolution. For data collection and refinement statistics, see supporting information (SI) Table 1. Superimposing all complexes indicates that single substitutions in InlA do not affect the structure of InlA itself or the geometry of the complex (rmsd ≤ 0.65 Å). This allows atomic changes in the immediate vicinity of the mutation to be analyzed, especially as regards hydrogen-bond networks and water-mediated interactions.

Y369A and Y369S. As postulated, the substitution of Tyr-369 with alanine or serine allows Asn-370 and His-392 to maintain their stacking interaction with Phe-348 as in uncomplexed InlA (Fig. 2*B*). Water molecules near Tyr-369 in InlA/hEC1 are largely conserved in Y369A/hEC1 and Y369S/hEC1 (black spheres). One such water molecule, hydrogen bonded to Asn-27-$N_{\delta 2}$ in all complexes, additionally forms a second hydrogen bond to Ser-369-O_{γ} in Y369S/hEC1, bridging $InlA_{Y369S}$ and hEC1. Two additional water molecules (red) bound by Asn-370 and His-392 are present only in Y369A/hEC1 and Y369S/hEC1. They replace a water molecule from the second solvation shell of InlA/hEC1 (blue).

S192N. In $InlA^{wt}$/hEC1, Ser-192 adopts two distinct equally occupied conformations, each involved in a water-mediated hydrogen bond to the main-chain oxygens of Phe-17_{hEC1} or Pro-18_{hEC1} (blue residues, Fig. 2*C*). The first of these water molecules additionally interacts with Ser-172_{InlA} and the second with Asp-213_{InlA}. Engineered Asn-192_{InlA} displaces the first bridging water, introducing a direct hydrogen bond from Asn-192_{InlA}-$N_{\delta 2}$ to Phe-17_{hEC1}-O. A low B factor of the second, now Asn-192_{InlA}-coordinated water (dark red sphere in Fig. 2*C*), indicates interaction of Asn-192_{InlA}, Pro-18_{hEC1}, and Asp-213_{InlA} in S192N/hEC1 to be much tighter than in InlA/hEC1. Asn-192_{InlA}-$O_{\delta 1}$ intramolecularly hydrogen bonds to backbone atoms of neighboring repeats, preventing additional stabilizing contacts to hEC1.

G194S+S. Replacing Gly-194_{InlA} by serine and inserting a second serine after Ser-194 (+S) allows LRR6 to adopt a structure similar to that of all other repeats (Fig. 2*D*). Correspondingly, Asn-195 flips into the hydrophobic core complementing the asparagine ladder. In addition, the hydrophobic water-filled cavity between wild-type InlA and hEC1 (Fig. 2*E*) is eliminated, reducing the distance between the two proteins from ≈10 to ≈4 Å (Fig. 2*F*). At least four water molecules are displaced from the interface (light-blue spheres in Fig. 2*D*). One water molecule, hydrogen-bonded by residue +S, additionally forms a long-range interaction (4.0 Å) to Glu-54_{hEC1}, whereas a second water, hydrogen-bonded to both Glu-54 and Lys-61_{hEC1}, similarly forms a long-range interaction of 4.2 Å to +S (Fig. 2*D*).

Complex Formation of InlA/hEC1 Is Enthalpy- and Entropy-Driven. Using ITC, we have narrowed down the dissociation constant for wild-type InlA/hEC1, found to be 8 ± 4 μM by analytical

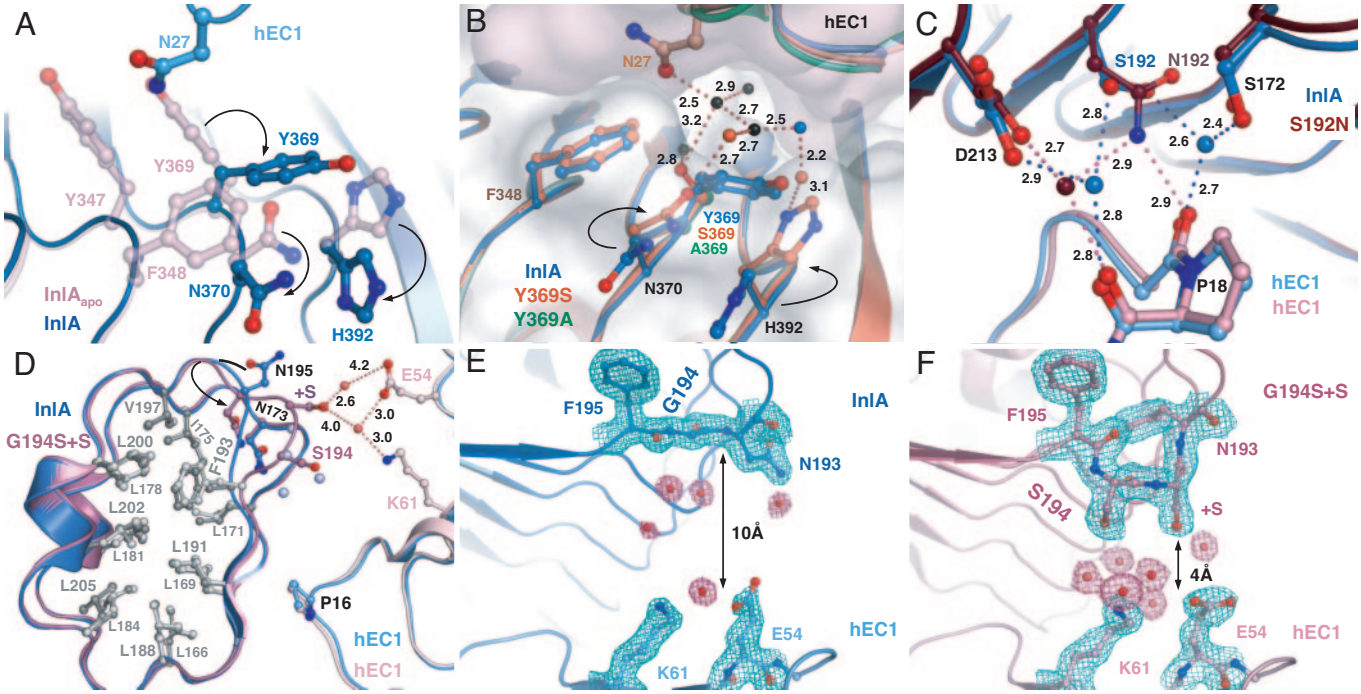

Fig. 2. Structural details of InlA variants. (*A*) Tyr-369$_{InlA}$-induced side-chain rearrangements during complex formation. Superposition of uncomplexed InlA (pink) and InlA/hEC1 (blue). On complex formation, Ans27$_{hEC1}$ (light blue, top) causes Tyr-369$_{InlA}$ (dark blue) to flip to an alternative less-favorable conformation, displacing Asn-370 and His-392 from their stacking interaction with Phe-348 (black arrows). (*B*) The interface near Tyr-369$_{InlA}$ in InlA/hEC1 (blue), Y369A/hEC1 (green), and S192N-Y369S/hEC1 (orange). Molecular surface of Y369A is in gray, and that of hEC1 is in pink. Side chains changing conformation during complex formation are shown for InlA/hEC1 (blue) and S192N-Y369S/hEC1 (dark orange). Spheres represent water molecules: black, conserved in all complexes; orange, present in variant complexes; blue, present only in the wild-type complex. Y369A/S prevents the disruptive reorientation of Tyr-369 during complex formation exposing a water-filled cavity. Ser-369-O$_\gamma$ binds two conserved water molecules, one also coordinated by Asn-27$_{hEC1}$-O$_{\delta1}$. (*C*) Ser-192$_{InlA}$ in InlA/hEC1 (blue) and S192N/hEC1 (dark red/pink). The two conformations of Ser-192 form water-bridged hydrogen bonds to Phe-17$_{hEC1}$-O or Pro-18$_{hEC1}$-O, respectively. The water molecules are additionally hydrogen-bonded by Asp-213$_{InlA}$ and Ser-172$_{InlA}$. Substituting Ser-192 by asparagine displaces one of the water molecules and introduces a direct hydrogen bond to Pro-18-O. The second water molecule maintains the hydrogen-bonding pattern of the wild-type complex. (*D–F*) InlA/hEC1 (blue) and G194S+S/hEC1 (pink). (*D*) The interaction of InlA-LRR5 and -6 with hEC1. The mutation G194S and the insertion of an additional serine (+S) restore the canonical LRR-repeat geometry in LRR6, flipping Asn-195 (arrow) into the hydrophobic core of InlA (dark/light gray, wild-type/variant) and removing a large water-filled cavity. (*E*) Electron density (1σ−contoured $2F_O - F_C$; green, protein; red, water) of LRR6 in InlA/hEC1 (5). The 21-residue LRR6 creates a hydrophobic cavity between Gly-194$_{InlA}$, Glu-54$_{hEC1}$, and Lys-61$_{hEC1}$. Water molecules filling the gap are poorly defined (weak electron density). (*F*) The equivalent view as *E* for G194S+S/hEC1. The gap between interaction partners is narrower, yielding a well defined yet unsaturated water cluster.

ultracentrifugation (5), to 3 ± 1 μM. In addition, ITC allows changes in binding affinity to be separated into enthalpic and entropic contributions.

Analyzing the association of InlA$_{wt}$ and hEC1 in different buffers indicates that the apparent enthalpy (ΔH_{app}) of complex formation depends on the ionization enthalpy (ΔH_{ion}) of the buffer (20), implying that complex formation is associated with an exchange of protons. ΔH_{app} is found to be −6.7 ± 0.3 kJ/mol in cacodylate (ΔH_{ion} = −2.5 kJ/mol), −1.8 ± 0.2 kJ/mol in Hepes (ΔH_{ion} = 23.9 kJ/mol), and 3.5 ± 0.2 kJ/mol in Tris (ΔH_{ion} = 47.7 kJ/mol). Plotting ΔH_{app} against ΔH_{ion} (not shown) indicates that 0.2 ± 0.1 protons (gradient) are taken up during complex formation. Although the side chain involved remains unclear, the binding enthalpy may be corrected for ΔH_{ion}. ΔH_{ion}-independent binding enthalpy is thus −6 kJ/mol (ΔH_{ion} = 0), and entropy ($T\Delta S$) is 25 kJ/mol. Complex formation of InlA and hEC1 is thus both entropy- and enthalpy-driven.

Revealing Atomic Contributions to Binding Enthalpy and Entropy. To compare enthalpic and entropic contributions to complex formation of InlA variants and InlA^{wt}, ITC experiments were performed in Hepes buffer and 20 mM $CaCl_2$. Values were corrected for ΔH_{ion} to place them on an absolute scale (Fig. 3). Differences in thermodynamic quantities remain unaffected, because they are independent of ΔH_{ion}. Surprisingly, the atomic modifications to InlA result in unexpectedly large and divergent changes in the thermodynamics of complex formation:

Y369A and Y369S. Both substitutions contribute enthalpically to complex formation: $\Delta\Delta H$ = −13 kJ/mol for Y369A (green boxes in Fig. 3) and −11 kJ/mol for Y369S (turquoise in Fig. 3; SI Fig. 4). In Y369A, the favorable enthalpic contribution is counteracted by an unfavorable reduction in binding entropy ($\Delta T\Delta S$ = −8 kJ/mol), a case of "enthalpy/entropy compensation" (15). Compared with InlA^{wt}, Y369A therefore increases binding affinity (K_d = 400 ± 100 nM) to hEC1 only 7.5-fold. In Y369S (blue labels in Fig. 3), a much smaller entropic compensation ($\Delta T\Delta S$ = −2 kJ/mol) results in the highest binding affinity for hEC1 (K_d = 90 ± 20 nM) of any single InlA variant investigated in this study.

S192N and G194S+S (magenta and yellow boxes in Fig. 3). Whereas Y369A and Y369S favorably decrease the enthalpy of complex formation compared with wild-type InlA, both S192N and G194S+S unfavorably increase this contribution ($\Delta\Delta H$ = 5 and 11 kJ/mol for S182N and G194S+S). In the case of G194S+S, the increase in enthalpy is sufficient to make complex formation endothermic (ΔH = 5 kJ/mol). The increase in binding enthalpy of both substitutions, however, is more than compensated by a large favorable increase in binding entropy ($\Delta T\Delta S_{S192N}$ = 12 kJ/mol, $\Delta T\Delta S_{G194S+S}$ = 19 kJ/mol), resulting in a significantly higher binding affinity for hEC1 than InlA^{wt}. Both S192N and

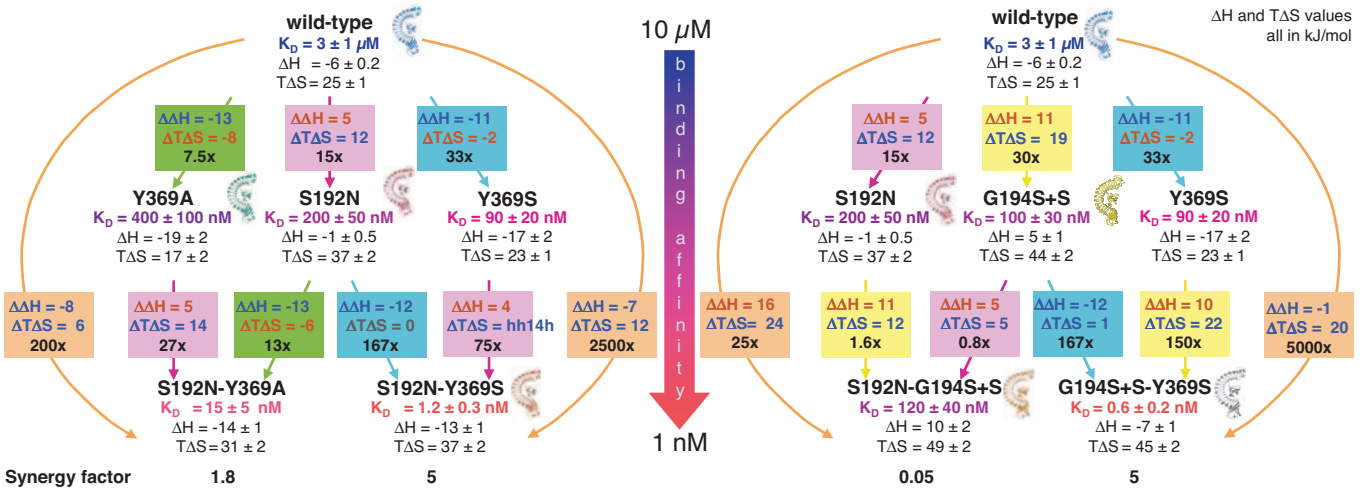

Fig. 3. Thermodynamic network. Binding affinities (K_d), binding enthalpies (ΔH), and entropies (TΔS) are summarized for wild-type and variant complexes of InlA with hEC1 (T = experimental temperature, 293 K). Enthalpies have been corrected for ΔH_{ion} of $InlA^{wt}$/hEC1 to place them on an absolute scale before calculating TΔS values. Substitutions and associated changes in binding enthalpies ($\Delta\Delta H = \Delta H_{var} - \Delta H_{wt}$) and entropies ($\Delta T\Delta S = T\Delta S_{var} - T\Delta S_{wt}$) are listed in color-coded boxes: green, Y369A; magenta, S192N; turquoise, Y369S; yellow, G194S+S. Blue and red fonts denote thermodynamically favorable and unfavorable changes. Note the excellent reproducibility for identical substitutions introduced into different backgrounds. Combining substitutions results in superadditive strengthening of corresponding complexes (orange-colored arrow). Synergy factors (bottom row; see text for definition) indicate positive (>1) and negative (<1) cooperativity between mutations. Ribbon pictograms mark complexes for which crystal structures have been solved. The color-graded arrow indicates binding affinities of InlA variants (blue ≈10 μM and red ≈1 nM K_D).

G194S+S thus improve surface complementarity of InlA/hEC1, allowing the entropically favorable elimination of around one and four constrained water molecules from the interface, respectively.

Thermodynamics of Long-Range Cooperativity Between Combined Substitutions. To achieve higher binding affinity, the described amino acid substitutions were combined to yield the four InlA variants S192N-Y369A, S192N-Y369S, G194S+S-Y369S, and S192N-G194S+S (bottom row, Fig. 3). Our data indicate that the combination of physically distant single substitutions significantly increases the binding affinity for hEC1. Thus, the binding affinities of S192N–Y369A, S192N–Y369S, and G194S+S–Y369S are 200-, 2,500-, and 5,000-fold (orange boxes, Fig. 3) that of $InlA^{wt}$. Although the sites of substitution are separated by ≈34 Å (Fig. 1*A*), binding affinities of individual substitutions are not merely additive but indicate positive cooperativity instead. "Synergy factors" were calculated by dividing the increase in binding affinity of the doubly substituted variant (lower colored box in Fig. 3) by that of the single substitution variant (upper box of identical color). For S192N–Y369A, this amounts to 13/7.5 ≈ 27/15 ≈ 1.8, whereas S192N–Y369S and G194S+S–Y369S both yield a value of ≈5 (Fig. 3). The synergy, although precise in terms of binding affinities, is less well defined with respect to enthalpy or entropy, in common with previous observations (14). Thus, ΔΔH and ΔTΔS for identical substitutions introduced either into $InlA^{wt}$ or combined with an InlA variant, differ by no more than 3 kJ/mol, only slightly larger than the average experimental error of 1–2 kJ/mol. The synergy, however, appears to be linked to an increase in entropy, because ΔTΔS is always 2–3 kJ/mol higher for double-substitution variants than for corresponding single-substitution variants (boxes of identical color in lower and upper rows of Fig. 3), in contrast to differences of 0 to −1 kJ/mol for ΔΔH.

The tightest binding affinity of an InlA variant for hEC1 is that of G194S+S–Y369S, $K_d = 0.6 \pm 0.2$ nM. Only two rationally chosen substitutions in InlA thus suffice to transform the weak binding affinity of the wild-type complex, $K_d = 3 \pm 1$ μM, to a tight fit comparable to that of typical proteinase/proteinaceous-inhibitor complexes (21), one of the highest increases in binding affinity (5,000-fold) reported for any protein–protein interaction (22).

In contrast to the synergy for S192N–Y369A, S192N–Y369S, and G194S+S–Y369S, the substitutions of the fourth double variant S192N–G194S+S are anticooperative. The synergy factor is 0.8/15 ≈ 1.6/30 ≈ 0.05 (Fig. 3), resulting in a binding affinity that is similar to that of the individual substitution variants rather than ≈10-fold stronger, as expected if the effects were additive. An increase of ΔTΔS of ≈7 kJ/mol indicates this to be an entropic effect.

Discussion

Thermodynamics of Complex Formation. Despite complex formation of InlA and hEC1 being both enthalpically and entropically favored and an apparent large interaction surface, the binding affinity of the complex is weak. Thermodynamically, binding entropy (dominated by exclusion of water molecules) outweighs binding enthalpy (hydrophilic interactions). Only two hydrophobic contact areas, centered on $Val\text{-}3_{hEC1}$ and $Pro\text{-}16_{hEC1}$, exist in InlA/hEC1. Nevertheless, by excluding numerous rotationally restrained water molecules during complex formation (5), these hydrophobic interactions would appear to contribute substantially to the favorable increase in binding entropy of $T\Delta S = 25 \pm 1$ kJ/mol. The small enthalpic contribution to complex formation ($\Delta H = -6 \pm 0.2$ kJ/mol) in turn correlates structurally with the paucity of enthalpically favorable direct contacts between the proteins. These include only seven hydrogen bonds, three salt bridges, and eight water-bridged interactions (5). Compared with tighter protein complexes, InlA/hEC1 retains significantly more water molecules within the interface (16).

On the other hand, the low binding affinity and poor surface complementarity of InlA/hEC1 provide us with an optimal system to study the crucial role of water in complex formation. By excluding more water molecules during complex formation, we may entropically stabilize the interaction. Alternatively, water molecules enthalpically contribute to binding affinity if their hydrogen bonding potential is optimized to bridge hydrophilic interfaces (16). The high-resolution structural data (Figs. 1 and 2) for InlA-variant/hEC1 complexes may thus be used to

interpret the observed changes in thermodynamic parameters (Fig. 3).

Note, however, that there are limitations to this approach, in particular because it is clearly impossible to take all contributing factors (including distant water molecules) into account. The tendency of enthalpy or entropy to compensate a change in the other (enthalpy–entropy compensation), especially in weak intermolecular interactions (23), furthermore, affects experimentally determined changes in enthalpy and especially entropy. As a result, small changes, particularly in entropy, are difficult or impossible to interpret in a structural sense. Large changes in thermodynamic contributions are not affected to the same extent, allowing their cause to be discussed qualitatively in terms of structural change, even though their constituent contributions are not precisely resolved. Examples include the rotational entropies of amino acid side chains, shown to be amenable to calculation (24), and the established entropic contribution of excluding water molecules from interfaces (23).

Y369A and Y369S (green and turquoise labels in Fig. 3). Compared with InlA/hEC1, the substitutions Y369A and Y369S improve the enthalpy of binding by $\Delta\Delta H = -13$ or -11 kJ/mol. Structurally, this may be rationalized by the enthalpically favorable stacking (25) of Phe-348, Asn-370, and His-392 being retained in the variant complexes from uncomplexed InlA (transparent pink residues in Fig. 2*A*), rather than being disrupted as in InlA/hEC1 (blue in Fig. 2*A*). This example confirms the general view that preorganized interaction surfaces in the unbound state play a major role in protein recognition (26).

The favorable increase in enthalpy of Y369A and Y369S in complex formation (above) is offset by a reduction in the entropic contribution by $\Delta T\Delta S = -8$ or -2 kJ/mol compared with that of wild-type InlA. These changes in entropy may be substantially affected by solvent entropy compensating the change in enthalpy, thereby limiting their detailed correlation with structural data (27). In a qualitative sense, the loss in entropy in Y369A/hEC1 ($\Delta T\Delta S = -8$ kJ/mol) and Y369S/hEC1 (-2 kJ/mol) appears to correlate with the size of a hydrophobic patch, being exposed as a result of the substitutions.

S192N. Ser-192 adopts two distinct conformations in $InlA_{wt}$/hEC1 (Fig. 2*C* and SI Fig. 5*A*), each of which hydrogen bonds a bridging water molecule (see above). Replacing Ser-192 by asparagine displaces one of these water molecules. Excluding a single water molecule from a protein interface leads to a favorable increase in solvation entropy of ≈6–9 kJ/mol (23). The observed exclusion of a water molecule thus presumably contributes to the 12 kJ/mol increase in binding entropy of S192N/hEC1 (magenta boxes in Fig. 3). The loss of the water-mediated hydrogen bond Ser-172_{InlA}–Phe-17_{hEC1} in turn may explain the corresponding unfavorable increase in binding enthalpy ($\Delta\Delta H =$ 5 kJ/mol).

G194S+S fills a large depression on the surface of InlA (compare Fig. 2 *E* and *F*), excluding four rotationally restrained water molecules from the interface and dramatically increasing binding entropy ($\Delta T\Delta S = 19$ kJ/mol). Why does the binding enthalpy increase to such an extent as to make complex formation endothermic? In the $InlA^{wt}$/hEC1, a distance of 10 Å between InlA and hEC1 is sufficiently large to allow bulk solvent to fill the cavity between the two independently solvated surfaces (Fig. 2*E*). In G194S+S/hEC1, the distance is reduced to ≈4 Å. This distance appears too narrow to allow independent solvation of each surface, yet too wide for a single bridging layer of solvent. Instead, the interprotein hydrogen-bonding network is observed to be discontinuous in the crystal structure, which would enthalpically be unfavorable.

Synergy of Combined Substitutions. Strikingly, our study indicates that changes in binding affinity of single substitutions are not simply additive when the substitutions are combined in a single protein.

For the combination of S192N and G194S+S, we observe anticooperative behavior characterized by a synergy factor of 0.05. Binding affinity of $InlA^{S192N-G194S+S}$/hEC1is thus weaker than the combination of individual substitutions would imply. The effect appears largely entropic, because $\Delta T\Delta S_{S192N\rightarrow S192N\text{-}G194S+S} - \Delta T\Delta S_{InlA\rightarrow G194S+S} \approx -7$ kJ/mol (Fig. 3), implying that overall approximately one water molecule less is displaced by S192N-G194S+S than by S192N and G194S+S combined. The crystal structure of S192N-G194S+S/hEC1 correspondingly indicates that the side chain of Asn-192_{InlA} is locked into a tight intramolecular hydrogen bond to the physically adjacent backbone nitrogen of Ser-194 (SI Fig. 5 *B* and *C*), preventing Asn-192 from displacing a water molecule as described for S192N/hEC1 (compare SI Fig. 5 *B* and *C* with *D*). Note that this explanation, as well as those below, only considers obvious structural changes in the corresponding crystal structures. More subtle contributions to the dynamics of protein association or to the structure and stability of the unbound proteins brought about by substituting individual residues fall outside the scope of this publication.

Potentially more interesting than the negatively cooperative S192N-G194S+S are the double variants S192N-Y369A, S192N-Y369S, and G194S+S-Y369S, all of which are characterized by synergy factors above one, indicating individual substitutions to be cooperative. As indicated above, these observed synergies are unambiguous in terms of binding affinities. The enthalpic (1–2 kJ/mol) and entropic (2–3 kJ/mol) contributions, however, are similar to or only slightly larger than the experimental error (1–2 kJ/mol) and should hence not be overinterpreted. Furthermore, it is interesting to note that S192N-Y369A, S192N-Y369S, and G194S+S-Y369S all combine individual substitutions physically separated by >30 Å (Fig. 1*A*). Previous studies of protein complexes indicated synergy to be limited to clustered residues, whereas spatially distant improvements were strictly additive (26). Only recently has this been challenged by the report of positive cooperativity for substitutions in the T cell receptor variable domain separated by 20 Å (17).

Structurally, the observed long-range synergy in InlA variants may be due to a physical link in the form of a β-strand between two sites of substitution. The variants S192N and Y369A/S increase binding affinity of hEC1 through favorable interactions to Phe-17_{hEC1} and Asn-27_{hEC1}, respectively, located at either end of β-strand **b** (β**b**, residues 19–26) of hEC1 (Fig. 1). In InlA/hEC1, β-strand β**a** (residues 2–10) and loop β**a**-β**b** (residues 10–19) constitute the major part of the interface, whereas interactions of β**b** to InlA are restricted to two water-mediated contacts. By stabilizing either end of β**b**, substitutions S192N and Y369A/S appear to stabilize hEC1 and hence the interface as a whole, resulting in the observed synergistic increase in binding affinities (Figs. 1 and 3).

The reason for a synergy factor of 1.8 for S192N–Y369A, compared with 5 for S192N–Y369S, may structurally be linked to the presence of an additional hydrogen bond to the C-terminal end of β**b** in Y369S (structure of S192N–Y369S/hEC1). This direct interaction is absent in Y369A, potentially indicating a lower stabilization of β**b** of hEC1 and reducing synergy to 1.8.

In the case of G194S+S–Y369S, the insertion of serine (+S) restores the canonical LRR architecture (Fig. 2*D* and SI Fig. 5*C*) and fills a large cavity between LRR6 of InlA and residues Glu-54_{hEC1} and Lys-61_{hEC1}. As the surface of G194S approaches the corresponding hEC1 surface more closely, at least four intervening water molecules are physically excluded. Long-range interactions to Glu-54_{hEC1} and Lys-61_{hEC1} appear to stabilize the complex. Glu-54 and Lys-61 are located in β-strands β**d** and β**e**, respectively, of hEC1 (Fig. 1*B*) that form a β-sheet with strand β**b**. The stabilization of β-strands β**d** and β**e** appears to be transmitted through a β-sheet to β**b** and thereby give rise to a

measurable positive cooperativity. This mechanism of long-range cooperativity may be equivalent to the intramolecular allosteric effects observed in multimeric protein complexes (28).

Methods

Mutagenesis, Expression, and Purification. For structural and biophysical studies, functionally relevant fragments of InlA (residues 36–496) and the first extracellular domain of E-cadherin (hEC1, residues 1–105) were used (5). Site-directed mutations were introduced by using QuikChange mutagenesis (Stratagene, La Jolla, CA). Protein expression and purification were as described (5).

ITC. ITC was performed by using a MCS calorimeter (MicroCal, Northampton, MA). All samples were dialyzed against 50 mM Hepes (pH 7.5) and 20 mM $CaCl_2$. Concentrations were determined spectrophotometrically at 280 nm. Titrations were performed at 20°C by injecting 5- to 10-μl aliquots of wild-type or variant InlA into the ITC cell containing 1.35 ml of hEC1. Data were corrected for heat of dilution (ΔH_{dil}). Binding stoichiometry and enthalpy as well as equilibrium association constants was determined by using the "single set of independent sites" model of molecular association (MicroCal Origin 2.9; MicroCal).

To determine the influence of the ionization enthalpy (ΔH_{ion}) of the buffer on the apparent enthalpy of binding (ΔH_{app}), complex formation of wild-type InlA and hEC1 was recorded in 50 mM cacodylate and in 50 mM Tris/HCl buffers supplemented with 20 mM $CaCl_2$. The relationship between ΔH_{ion} and ΔH_{app} was analyzed by linear regression to evaluate the number of protons exchanged during complex formation. The binding enthalpy was corrected for the ionization enthalpy of the buffer ($\Delta H_{ion} = 0$ intercept of ΔH_{app} plotted against ΔH_{ion}).

Crystallization and Data Collection. Complexes of InlA variants and hEC1 were crystallized by hanging drop vapor diffusion. Total protein concentration was 5 mg/ml, with a stoichiometry of InlA:hEC1 = 1:1. The reservoir solution contained 20–25% PEG 4000, 100 mM Mes/Tris buffer (pH 7.0–7.5), 100 mM Na acetate, and 20–100 mM $CaCl_2$. Twenty percent PEG 400 (vol/vol) was added to the reservoir solution for cryoprotection.

Structure Determination. X-ray data were collected by using synchrotron radiation and MARCCD detectors (Marresearch, Norderstedt, Germany) at beamlines BW6 (Y369A/hEC1 and S192N/hEC1, λ = 1.05 Å) and X13 [European Molecular Biology Laboratory, Deutsches Elektronen Synchrotron, Hamburg, Germany (G194S+S/hEC1, λ = 0.80 Å)] and BL1 [Protein Structure Factory, Berliner Elektronenspeicherring Gesellschaft für Synchrotronstrahlung, Berlin, Germany (S192N-Y369S/hEC1, λ = 0.95 Å)]. Data for complexes G194S+S-Y369S/hEC1 and S192N-G194S+S were collected by using a rotating anode (Rigaku, San Diego, CA; 1.54 Å) and an image plate detector (R-Axis IV). Data were processed by using HKL (29), XDS (30), and the CCP4 suites (31). All structures were solved by molecular replacement by using EPMR (32) and the wild-type InlA/hEC1 complex as search model (5). REFMAC5 (33) was used for refinement and Coot (34) for model building, structural analysis, and structure validation, as well as WHATIF (35). Figures were prepared by using PYMOL (www.pymol.org).

We thank our colleagues Dr. Uwe Kärst and Prof. Jürgen Wehland (Division of Cell Biology, Helmholtz Centre for Infection Research) for the isothermal titration calorimeter. We express our appreciation to the staff of beamlines BW6 and X13 (European Molecular Biology Laboratory, Deutsches Elektronen Synchrotron) and beamline BL1 (Berliner Elektronenspeicherring Gesellschaft für Synchrotronstrahlung) for access to beam time and for assistance during x-ray data collection. Financial support from the Deutsche Forschungsgemeinschaft as part of the Priority Program SPP1150 (Grant SCHU 1560/1-1/1-2, to W.-D.S.) is gratefully acknowledged.

1. Vazquez-Boland JA, Kuhn M, Berche P, Chakraborty T, Dominguez-Bernal G, Goebel W, Gonzalez-Zorn B, Wehland J, Kreft J (2001) *Clin Microbiol Rev* 14:584–640.
2. Gaillard JL, Berche P, Frehel C, Gouin E, Cossart P (1991) *Cell* 65:1127–1141.
3. Schubert WD, Heinz DW (2003) *ChemBioChem* 4:1285–1291.
4. Schubert WD, Göbel G, Diepholz M, Darji A, Kloer D, Hain T, Chakraborty T, Wehland J, Domann E, Heinz DW (2001) *J Mol Biol* 312:783–794.
5. Schubert WD, Urbanke C, Ziehm T, Beier V, Machner MP, Domann E, Wehland J, Chakraborty T, Heinz DW (2002) *Cell* 111:825–836.
6. Gumbiner BM (2005) *Nat Rev Mol Cell Biol* 6:622–634.
7. D'Souza-Schorey C (2005) *Trends Cell Biol* 15:19–26.
8. Gates J, Peifer M (2005) *Cell* 123:769–772.
9. Boggon TJ, Murray J, Chappuis-Flament S, Wong E, Gumbiner BM, Shapiro L (2002) *Science* 296:1308–1313.
10. Lecuit M, Ohayon H, Braun L, Mengaud J, Cossart P (1997) *Infect Immun* 65:5309–5319.
11. Lecuit M, Dramsi S, Gottardi C, Fedor-Chaiken M, Gumbiner B, Cossart P (1999) *EMBO J* 18:3956–3963.
12. Zahnd C, Amstutz P, Plückthun A (2007) *Nat Methods* 4:269–279.
13. Kortemme T, Baker D (2004) *Curr Opin Chem Biol* 8:91–97.
14. Reichmann D, Rahat O, Cohen M, Neuvirth H, Schreiber G (2007) *Curr Opin Struct Biol* 17:67–76.
15. Dunitz JD (1995) *Chem Biol* 2:709–712.
16. Levy Y, Onuchic JN (2006) *Annu Rev Biophys Biomol Struct* 35:389–415.
17. Moza B, Buonpane RA, Zhu P, Herfst CA, Rahman AK, McCormick JK, Kranz DM, Sundberg EJ (2006) *Proc Natl Acad Sci USA* 103:9867–9872.
18. Lo CL, Chothia C, Janin J (1999) *J Mol Biol* 285:2177–2198.
19. Kobe B, Deisenhofer J (1994) *Trends Biochem Sci* 19:415–421.
20. Christensen JJ, Hansen LD, Izatt RM (1976) *Handbook of Proton Ionization Heats and Related Thermodynamic Quantities* (Wiley, New York).
21. Stites WE (1997) *Chem Rev* 97:1233–1250.
22. Buonpane RA, Moza B, Sundberg EJ, Kranz DM (2005) *J Mol Biol* 353:308–321.
23. Dunitz JD (1994) *Science* 264:670.
24. Cole C, Warwicker J (2002) *Protein Sci* 11:2860–2870.
25. Meyer EA, Castellano RK, Diederich F (2003) *Angew Chem Int Ed Engl* 42:1210–1250.
26. Keskin O, Ma B, Nussinov R (2005) *J Mol Biol* 345:1281–1294.
27. Mark AE, van Gunsteren WF (1994) *J Mol Biol* 240:167–176.
28. Changeux JP, Edelstein SJ (2005) *Science* 308:1424–1428.
29. Otwinowski Z, Minor W (1997) *Methods Enzymol* 276:307–321.
30. Kabsch W (1988) *J Appl Crystallogr* 21:916–924.
31. (1994) *Acta Crystallogr D* 50:760–763.
32. Kissinger CR, Gehlhaar DK, Fogel DB (1999) *Acta Crystallogr D* 55:484–491.
33. Murshudov GN, Vagin AA, Dodson EJ (1997) *Acta Crystallogr D* 53:240–255.
34. Emsley P, Cowtan K (2004) *Acta Crystallogr D* 60:2126–2132.
35. Vriend G (1990) *J Mol Graphics* 8:52–56:29.