Structural basis for selective AMPylation of Rac-subfamily GTPases by Bartonella effector protein 1 (Bep1)

Nikolaus Dietz,1,1, Markus Huber,1,1, Isabel Sorg, Arnaud Goepfert,2,2, Alexander Harms, Tilman Schirmer,3,3, and Christoph Dehio*3,3

1Biozentrum, University of Basel, 4056 Basel, Switzerland

Edited by Ralph R. Isberg, Tufts University School of Medicine, Boston, MA, and approved February 8, 2021 (received for review November 8, 2020)

Small GTPases of the Ras-homology (Rho) family are conserved molecular switches that control fundamental cellular functions in eukaryotic cells. As such, they are targeted by numerous bacterial toxins and effector proteins, which have been intensively investigated regarding their biochemical activities and discrete target spectra; however, the molecular mechanism of target selectivity has remained largely elusive. Here we report a bacterial effector protein that selectively targets members of the Rac subfamily in the Rho family of small GTPases but none in the closely related Cdc42 or RhoA subfamilies. This exquisite target selectivity of the FIC domain family of small GTPases but none in the closely related Cdc42 or RhoA subfamilies. This exquisite target selectivity of the FIC domain AMP-transferase Bep1 from Bartonella rochalimae is based on electrostatic interactions with a subfamily-specific pair of residues in the nucleotide-binding G4 motif and the Rho insert helix. Residue substitutions at the identified positions in Cdc42 enable modification by Bep1, while corresponding Cdc42-like substitutions in Rac1 greatly diminish modification. Our study establishes a structural understanding of target selectivity toward Rac-subfamily GTPases and provides a highly selective tool for their functional analysis.

AMPylation | structure function | FIC domain | RhoGTPases | Bartonella effector protein

Small GTPases of the Ras-protein superfamily are molecular switches that control fundamental cellular functions in eukaryotes by cycling between GTP-bound “on” and GDP-bound “off” conformational states of their switch regions 1 (Sw1) and 2 (Sw2) (1, 2). Members of the Ras-homology (Rho) protein family function as signaling hubs and regulate cytoskeletal rearrangements, cell motility, and the production of reactive oxygen species (3, 4). The defining element in Rho-family GTPases is the presence of a Rho insert, a highly variable, 13-residue-long, α-helical insert close to the C terminus. The Rho insert has previously been implicated in the wiring of Rho-family GTPases to their specific biological functions (5, 6). Six members of the Rho-protein family closely related to Cdc42 share an altered amino acid sequence in the G4 nucleotide binding motif with a glutamine residue instead of lysine in the second position.

Due to their central role in eukaryotic cell signaling, especially in the immune response, Rho-family GTPases are targeted by a plethora of bacterial virulence factors, including secreted bacterial toxins that autonomously enter host cells and effector proteins that are directly translocated from bacteria into host cells via dedicated secretion systems (7, 8). By means of these virulence factors, pathogens established ways to stimulate, attenuate, or destroy the intrinsic GTPase activity of Rho-family GTPases, either directly through covalent modification of residues in the Sw1 or Sw2 regions (8) or indirectly by mimicking guanine nucleotide exchange factor (GEF) or GTPase-activating protein (GAP) function. However, the structural basis for selective targeting of Rho-family GTPase subfamilies has remained unknown (7).

The bacterial genus Bartonella comprises a rapidly expanding number of virtually omnipresent pathogens adapted to mammals, many of which have been recognized to cause disease in humans (9). The stealth infection strategy of Bartonella spp. (10) rely to a large extent on translocation of multiple Bartonella effector proteins (Beps) via a dedicated type 4 secretion system. Strikingly, the majority of the currently known several dozens of Beps contains enzymatic FIC domains (9, 11), indicating that Bartonella spp. successfully utilize this effector type in their lifestyle. In order to gain more insights into the function of FIC domain-containing Beps we have here investigated Bep1 of Bartonella rochalimae originally described by Harms et al. (11).

Filamentation induced by cyclic AMP (FIC) domain-containing effector proteins belong to the ubiquitous FIC protein family with a conserved molecular mechanism for posttranslational modification of target proteins. FIC domains consist of six helices with a common HxFxD/E/GNGRxR motif between the central helices 4 and 5 (12). Some of the FIC domain-containing effector proteins have been recognized to modify Rho-family GTPases by catalyzing transfer of the AMP moiety from the ATP substrate to specific target hydroxyl side chains (12, 13). Prototypical examples are the effector proteins IbpA from Histophilus somnii and VopS from Vibrio parahaemolyticus, which both target a wide range of Rho-family GTPases and AMPylate (adenylylate) a conserved tyrosine or

Significance

Mammalian cells regulate diverse cellular processes in response to extracellular cues. Small GTPases of the Rho family act as molecular switches to rapidly regulate discrete cellular activities, such as cytoskeletal dynamics, cell movement, and innate immune responses. Numerous bacterial virulence factors modulate the function of Rho-family GTPases and thereby manipulate intracellular signaling. For many of these virulence factors we have gained detailed understanding how they covalently modify individual Rho-family GTPases to reprogram their activities; however, their mechanisms of selective targeting of distinct subsets of Rho-family GTPases remained elusive. Using a combination of structural biology and biochemistry, we demonstrate for the effector protein Bep1 exclusive specificity for Rac-subfamily GTPases and propose the underlying mechanism of target selectivity.

Author contributions: N.D., M.H., I.S., A.H., T.S., and C.D. designed research; N.D., M.H., I.S., A.H., A.H., performed research; N.D., M.H., A.G., A.T., and C.D. analyzed data; and N.D., M.H., T.S., and C.D. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1N.D. and M.H. contributed equally to this work.

2Present address: Ichnos Sciences Biotherapeutics SA, 1066 Epalinges, Switzerland.

3To whom correspondence may be addressed. Email: tilman.schirmer@unibas.ch or christoph.dehio@unibas.ch.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2023245118/-/DCSupplemental.

Published March 15, 2021.
Bep1 selectively AMPylates Rac-subfamily GTPases. Bep1 is composed of a canonical FIC domain followed by an oligosaccharide binding (OB) fold and a C-terminal BID domain (11). The latter domain is implicated in recognition and translocation by the type 4 secretion system VirB/VirD4 of Bartonella (18, 19).

In search for Bep1 targets we performed AMPylation assays by incubating lysates of Escherichia coli expressing Bep1 with eukaryotic cell lysates and α-P32-labeled ATP and observed a radioactive band migrating with an apparent molecular weight of 20 kDa (SI Appendix, Fig. SA), consistent with modification of Rho-family GTPases as previously described for Ibpa and VopS (15, 16). To investigate further, we explored the target spectrum of Bep1 and compared it to those of the FIC domains of Ibpa (IbpAFIC) or VopS (VopSFIC) by selecting 19 members of the Ras superfamily (Fig. 1A) with an emphasis on members of the Rho family. While AMPylation activity of all three enzymes was strictly confined to Rho-family GTPases, their target selectivity spectra differed markedly: while Bep1 modified exclusively members of the Rac subfamily (IbpAFIC2 or VopSFIC2) by selecting 19 members of the Ras superfamily (Fig. 1A) with an emphasis on members of the Rho family. While AMPylation activity of all three enzymes was strictly confined to Rho-family GTPases, their target selectivity spectra differed markedly: while Bep1 modified exclusively members of the Rac subfamily (IbpAFIC2 or VopSFIC2) by selecting 19 members of the Ras superfamily (Fig. 1A) with an emphasis on members of the Rho family.

Next, we designed a minimal Bep1FIC construct (residues 13 to 229) that proved sufficient for selective target modification. Bep1 belongs to the class I of FIC proteins that are regulated by a small regulatory protein, here BiaA, that inhibits FIC activity by inserting a glutamate residue (E33) into the ATP binding pocket (20). In order to improve expression level and stability, we coexpressed Bep1FIC with an inhibition relieved mutant (E33G) of BiaA, yielding the stabilized minimal AMPylation-competent Bep1FIC/IBaA(E33G) complex, in short, Bep1FIC*

Bep1FIC* efficiently AMPylates its targets, and the activity depends on the presence of the catalytic histidine (H170) of the signature motif (Fig. 1B), consistent with the canonical AMPylation mechanism (20). Bep1FIC*, in contrast to VopSFIC, does not AMPylate Rac1Y32F (Fig. 1C), indicating that Bep1FIC* modifies Y32 of the Rac1 Sw1 as confirmed by mass spectrometry (SI Appendix, Fig. S1C). Thus, Bep1FIC* catalyzes the equivalent modification as IbpaFIC2 (15, 21), whereas VopS modifies T35 (16).

In contrast to the GDP form, GTP-loaded GTPases may not be amenable to FIC-mediated modification of Y32 since this residue is known to be involved in GTP binding via interaction with the γ-phosphohate group (22) (SI Appendix, Fig. S2D). Indeed, exchanging GDP against GTP efficiently protected the GTP hydrolysis deficient mutant Rac1Q61L from modification, and the same effect was observed when replacing GDP bound to wild-type Rac1 with nonhydrolyzable GTPγS (SI Appendix, Fig. S2C). Thus, we conclude that GDP-loaded GTPases are the physiological targets of Bep1-mediated AMPylation.

The Crystal Structure of Bep1FIC Reveals an Extended Target Recognition Flap. To reveal the structural basis of target selectivity, we solved the crystal structure of Bep1FIC to 1.6 A resolution. The structure (Fig. 2) closely resembles those of other FIC domains with AMPylation activity such as VbhT (20), Ibpa (21), and VopS (23), featuring the active site defined by the conserved signature motif encompassing the α4–α5 loop and the N-terminal part of α5. Comparison with the apo crystal structure of the close Bep1 homolog from Bartonella clarridgeiae (Protein Data Bank [PDB] ID 4nps) shows that the presence of the small regulatory protein mutant BiaA (E33G) in Bep1FIC* does not affect the structure of the FIC domain (SI Appendix, Fig. S2B).

**Fig. 1.** Bep1 selectively targets Rac-subfamily GTPases. (A) 32P-autoradiograms of in vitro AMPylation reactions using the indicated purified and GDP-loaded Rho-family GTPases display exquisite selectivity of full-length Bep1 for Rac-subfamily GTPases in contrast to the broader target spectrum of IbpaFIC2 and VopSFIC2. (B) The FIC domain of Bep1 in complex with the regulatory protein BiaA (Bep1FIC*) is sufficient for the recognition of Rac-subfamily GTPases and the catalytic H170 is required for AMPylation. (C) Bep1FIC* AMPylates residue Y32 of Rac1 and RhoG since the respective Y32F mutants are not modified. AMPylation by the T35-specific VopSFIC indicates structural integrity of the analyzed GTPases and their Y32F mutants. (D) Venn diagram showing AMPylation target selectivity of tested FIC domains, overlaid to the phylogenetic relation of Rho-family GTPases (4).
The active site is partly covered by a β-hairpin flap (Fig. 2A) that serves to register the segment carrying the modifiable side chain (here Sw1) to the active site via β-sheet augmentation, as has been inferred from bound peptides (16, 24), observed directly in the IbpAFIC2:Cdc42 complex (21), and discussed elsewhere (17). Strikingly, the flap of Bep1 and its orthologs in other Bartonella species (SI Appendix, Fig. S2A) is considerably longer than in other FIC structures (e.g., of IbpAFIC2) and features a well-defined bulge at its tip (Fig. 2B and C).

Bep1FIC:Target Model Suggests That Charged Residues of the Flap Determine Target Selectivity. The complex structure of an FIC enzyme with a small GTPase target and the mechanism of FIC catalyzed AMPylation reaction has been elucidated for IbpAFIC2 in complex with GDP-loaded Cdc42 (21) (Fig. 3B). The detailed view in Fig. 3D shows that the Sw1 segment of Cdc42 exhibits an extended conformation and forms antiparallel, largely sequence-independent, β-sheet interactions with the flap of the FIC enzyme, thereby aligning the modifiable Y32 with the active site. Considering the close structural homology of the catalytic core of Bep1FIC with IbpAFIC2 (rmsd = 1.0 Å for 32 Ca atoms in the active site helices) and of Rac-subfamily GTPases with Cdc42 (rmsd = 0.44 Å for 175 Ca positions), we reasoned that computational assembly of a Bep1FIC:Rac complex could provide a structural basis for an understanding of Bep1 target selectivity.

Fig. 3A shows the assembled Bep1FIC:Rac2 complex that was obtained by individual superposition of 1) the Bep1FIC active site helices and the flap with the corresponding elements in IbpAFIC2 and 2) the Sw1 loop of Rac2 with that of Cdc42. Thereby, we assumed implicitly that the interaction between these central segments should be very similar since both FIC enzymes utilize a homologous set of active residues to catalyze AMP transfer to a homologous residue (Y32) on Sw1.

The local structural alignment resulted in a virtually identical relative arrangement of the FIC core to the GTPase as in the template structure (compare Fig. 3A and B) and caused no steric clashes. Conspicuously, the extended Bep1FIC flap is accommodated in a groove formed by Sw1 (residues 31 to 40), the GDP-loaded nucleotide binding G4 motif [T(K/Q)xD, residues 115 to 118] (25), and the following Rho-insert helix (Rac2 residues 121 to 133) (Fig. 3C and SI Appendix, Fig. S2E).

The manually created complex model was used as input for an adapted Rosetta modeling protocol to allow for sampling of backbone and side chain torsion angles in the interface of the complex, as described in Materials and Methods (26, 27). Consistent with the low affinity of the complex in vitro (see below), the models confirm the relatively small interface area of ~800 Å². Common to all top scoring models we find that the modifiable residue Y32 is pointing toward the active site of Bep1, where it is held in place by a main chain-mediated interaction between the base of the flap and the Sw1 loop of the GTPase (SI Appendix, Fig. S3A), indicating that the configuration of active site residues and the modifiable tyrosine side chain is, indeed, most likely the same as in the template complex.

However, in the IbpAFIC2:Cdc42 complex, the aforementioned GTPase groove on the nucleotide binding face is not utilized for AMPylation of Rac-subfamily GTPases by Bartonella effector protein 1 (Bep1)
Instead, the so-called arm domain of IbpAFIC2 (Fig. 3B) constitutes a major part of the interface and contacts the highly conserved Sw2 loop of Cdc42. This rationalizes the broad target spectrum of arm domain-containing FIC AMP transferases like IbpA and VopS (12, 23). In turn, residues of the groove predicted to get recognized exclusively by Bep1FIC are likely to be important for the limited target range of Bep1. Conspicuously, the top scoring models revealed two potential salt bridges between the Bep1 flap and the Rac2 groove, namely, D119(Bep1)–K116(Rac2) and K117(Bep1)–D124(Rac2) (Fig. 3C and SI Appendix, Fig. S3A). Since the combination of K116 and D124 is exclusively found in the Rac subfamily as revealed by sequence alignment of Rho-family GTPases (Fig. 3E), we reasoned that these residues may contribute significantly to the specific recognition of Rac GTPases by Bep1 (Fig. 1A).

**Two Salt Bridges between Flap and Target Are Crucial for Selective Interaction of Bep1FIC with Rac-Subfamily GTPases.** The relevance of the two identified salt bridges in the Bep1FIC*:Rac2 complex (Fig. 3C) for affinity and selectivity was tested by single and double replacements of the constituting residues 116 and 124 in a Bep1 target and a nontarget GTPase. For Rac1, we tested if substitutions at these residues with corresponding amino acids of Cdc42—a nontarget of Bep1 with the highest conservation in regions flanking the proposed interaction sites (Fig. 3E)—influence target recognition (loss-of-function approach; see interaction...
schemes in Fig. 4(A)). In addition, we tested whether Cdc42 can be converted to a Bep1 target by reciprocal substitution(s) of these sites with the corresponding Rac1 residues (gain-of-function approach; Fig. 4B).

First, we applied, as for Fig. 1A, the autoradiography endpoint assay with $^{32}$P-α-ATP as substrate. Compared to wild-type Rac1, mutant D124S showed no significant difference in the amount of AMPylated target, whereas AMPylation of mutant K116Q and, even more, of the double mutant was found drastically reduced (Fig. 4C and SI Appendix, Fig. S4A). Conversely, in the gain-of-function approach, Cdc42 mutant S124D did not convert the GTPase to a Bep1 target, while mutant Q116K and the double mutant showed low but significant AMPylation (Fig. 4D and SI Appendix, Fig. S4B). In a fairly undiscriminating way, IbpAFIC2 modified all investigated GTPase variants (SI Appendix, Fig. S4C and D) indicating their proper folding. Together, the semiquantitative radioactive endpoint assay demonstrated a major role of K116 in target recognition by Bep1FIC*, while a contribution of D124 could not be demonstrated.

To overcome the limitations of the radioactive endpoint assay and to characterize target AMPylation quantitatively, we developed an online ion exchange chromatography (oIEC) assay (Materials and Methods) which allows separation of reaction components (Fig. 4E) and efficient acquisition of enzymatic progress curves to determine initial velocities, $v_{init}$ (see, for instance, SI Appendix, Fig. S4F, Inset). For AMPylation of Rac1 by Bep1FIC*, titration experiments yielded $K_M$ values of 0.52 and 1.4 mM for the substrates ATP and Rac1, respectively, and a $k_{cat}$ of 1.9 s$^{-1}$. The comparison with published values on other Fic AMP transferases (SI Appendix, Table S1) shows that the $K_M$ values are comparable to IbpA but that $k_{cat}$ is smaller by about two orders of magnitude.

Considering the physiological conditions in the cell with an ATP concentration above $K_M$, Bep1 can be expected to be saturated with ATP and only partially loaded with the target (target concentration << $K_M$ target). In such a regime, the AMPylation rate will be given by

$$v = \frac{k_{cat}}{K_M \text{target}} \times [E_0] \times [\text{target}]$$

(28), i.e., will depend solely on the second order rate constant $k_{cat}/K_M \text{target}$ (efficiency constant), which is, thus, the relevant parameter for enzyme comparison.

Next, we determined the efficiency constants for all GTPase variants. In the loss-of-function series, the single mutants reduced the efficiency constant by 2- and 6-fold, and the double mutant reduced the efficiency constant by about 30-fold (Fig. 4E and SI Appendix, Table S1).

Under the assumptions that 1) $k_{cat}$ is not changed upon the mutations, since they affect sites on the target that are distant from the catalytic center, and 2) $K_M$ is equal to the $K_D$ of the enzyme–target complex, as is warranted for a slow enzyme, the difference in the measured efficiency constants can be attributed to an altered stability of the Michaelis–Menten complex.

---

**Fig. 4.** Two salt bridges are crucial for Rac-subfamily selective AMPylation. (A) Schematic view of the two intramolecular Bep1FIC–Rac1 salt bridges (Left) and their partial disruption upon site-directed Rac1 mutagenesis, yielding Rac1 loss-of-function mutants (Right). (B) Absence of ionic interactions in the predicted Bep1FIC–Cdc42 interface (Left) and partial establishment of salt bridges in Cdc42 gain-of-function mutants (Right). (C and D) AMPylation of the variants given in A and B as measured by autoradiography. Note that due to the employed higher Bep1FIC* concentration (Material and Methods), the experiments in D also revealed auto-AMPylation of Bep1FIC*. (E and F) Enzymatic efficiency constants, $k_{cat}/K_M$ for Bep1FIC* catalyzed AMPylation of the GTPase variants shown in A and B as derived from the oIEC measurements shown in SI Appendix, Fig. S4. b.d., below detection limit. Error bars indicate standard deviation of reaction efficiencies.
Furthermore, the change of the free energy of binding upon mutation (ΔΔG) can be derived from the measured efficiency constants of wild-type and mutant target under these assumptions. The calculations given in SI Appendix, Table S2, show that the ΔΔG of the double mutant is larger by about 25% compared to the ΔΔG sum of the single mutants, suggesting that the contributions of the two salt bridges are largely independent. In the gain-of-function series, wild-type Cdc42 showed no and mutant S124D only marginal modification, while mutant Q116K showed a significant (about 30-fold larger than that of S124D) effect. Again, as in the previous series, the double mutant showed the largest effect (Fig. 4F and SI Appendix, Fig. S1).

Summarizing, the quantitative oIEC assay confirmed the prominent dependence of Bep1FIC* catalyzed target modification on the type of residue in target position 116 that had already been revealed by the radioactive endpoint assay and predicted by modeling (SI Appendix, Fig. S3A) but also demonstrated a significant influence of the residue in position 124, such that both salt bridges appear to be crucial for efficient Bep1-mediated AMPylation of Rac-subfamily GTases.

Discussion

Single residue alterations in the effector loop (switch I region) of Rac-family GTases can alter the specificity for interaction with downstream effectors in cellular signaling cascades (29). Several protein interaction modes have been described for Rho-family GTases (30, 31), even though the basis of discrimination between these structurally conserved but functionally diverse GTases remained elusive. The highly divergent Rho insert has been linked to a number of biological effects, such as membrane ruffling, Rho kinase activation by RhoA (32, 33), or the interaction of Rac with the NADPH oxidase complex (34). However, these studies relied on deletion of the Rho insert, and it is unclear if respective mutant proteins were properly folded. More recent structural work on complexes between Formins (mDia and FMNL2) and RhoC (35) or Cdc42 (36, 37) show the direct involvement of the C-terminal residues of the Rho insert in complex formation. While the Rho insert contributes only marginally to RhoC-mDia complex formation (35), it is crucial for interaction specificity in the FMNL2:Cdc42 complex (36). Our structure–function analysis substantially augments this body of work and demonstrates that target selectivity of Bep1 for GTases is encoded by interaction with a different set of Rho-family specific structural elements: Bep1 interacts with N-terminal residues of the Rho insert helix as well as the G4 motif residues. The observation that Cdc42 cannot be converted fully to a Rac1-like Bep1 target by the respective residue substitutions suggests additional, yet unknown, structural or dynamic features that contribute to efficient AMPylation.

Remarkably, Bep1’s selectivity is based by and large on a short insert of six residues in the conserved lid loop of the FIC domain (Fig. 2C). This simple, yet elegant, evolutionary treat equips Bartonella with a precise molecular tool to interfere specifically with host signaling. As such, Bep1 is the first bacterial effector to selectively target Rac-subfamily GTases without affecting the Rho or Cdc42 GTase subfamilies. Insertions of few amino acids in loop regions as exemplified by Bep1 are found in other FIC proteins; however, their functional consequences are hard to predict based on sequences alone. However, it is conceivable that they contribute to the specificity for different target spectra. Targeting a broad range of Rho GTases seems to require a more complex addition to the FIC domain as exemplified by the arm domain found in IbpA or VopS (Fig. 3A and B).

We speculate that in the infection process of Bartonella, the selective inactivation of Rac-subfamily GTases plays a critical role for the evasion of the innate immune response, without causing the collateral damage and activation of the immune system associated with effectors that target a broad-spectrum of Rho GTases, such as VopS or IbpA. In fact, Rac-subfamily selective AMPylation does not trigger a response of the innate immune system via activation of the pyrin inflammasome, which has been shown to accompany RhoA inactivation by covalent modification in the Sw1 region (38). Thus, avoiding RhoA inactivation may provide a substantial benefit for Bartonella to establish a largely asymptomatic chronic infection in their host.

Patients with impaired signaling of Rac-subfamily GTases cannot clear bacterial infections due to diminished ability for ROS production in immune cells, as seen in patients suffering from chronic granulomatosis disease or case studies from patients with dysfunctional Rac2 genes resulting in neutrophil immunodeficiency syndrome (39, 40). Along these lines, we speculate that selective targeting of GDP-complexed Rac-subfamily GTases provides the additional benefit that protein levels of GDP-bound Rac are not down-regulated via proteasomal degradation (41), resulting in a stable pool of inactive Rac subfamily GTases that would subdue Rac-mediated immune responses effectively.

Beyond providing a molecular understanding for target selectivity among Rho-family GTases, the narrow target spectrum of Bep1 for Rac-subfamily GTases also provides a unique tool for dissecting their specific functions in cellular processes, such as cytoskeletal rearrangements related to the Rac1-dependent formation of membrane ruffles, the Rac2/RhoG-dependent production of reactive oxygen in immune cells, or the role of Rac1 in carcinogenesis.

Considering the simple topology and small size of the FIC domain, we find a surprisingly modular division of functions. While the conserved catalytic core allows efficient AMPylation of a target hydroxyl residue located in an extended loop that registers to the active site via β-strand augmentation, target affinity and thereby selectivity is encoded separately in a short loop insertion. The modular nature and amenable size of this structural framework appears well suited for the rational design of synthetic Rho-subfamily selective FIC domain AMP transferases with novel physiological activities and beyond.

Materials and Methods

Protein Expression and Purification

The FIC domain of Bep1 was cloned, expressed and purified in complex with the inhibition-relieved regulatory protein BiaA2161G as described for the crystallization construct and is subsequently referred to as Bep1FIC*. For the generation of cleared bacterial lysate, the bacterial pellet was resuspended in reaction buffer (50 mM Tris HCl, pH 8.0, 150 mM NaCl, 5 mM MgCl2) supplemented with protease inhibitor mixture (complete EDTA-free mini, Roche) and lysed by sonication. After clearing the lysates by centrifugation (120,000 × g for 30 min at 4 °C), the supernatant was directly used as assay buffer if stored at −20 °C. Protein expression and purification of GST- or HIS-tagged GTases and GST-tagged FIC domains of VopS and IbpA followed standard GST- or HIS-fusion-tag protocols. In short, E. coli BL21 or BL21 AI (Invitrogen) were transformed with expression plasmids and used for protein expression. Bacteria were grown in LB medium supplemented with appropriate antibiotic on a shaker until absorbance at 600 nm = 0.6 to 0.8 at 30 °C. Protein expression was induced by addition of 0.2 mM isopropyl-β-thiogalactopyranoside (IPTG) (AppliChem GmbH) or 0.1% w/v arabinose (Sigma-Aldrich) for 4 to 5 h at 22 °C.

Bacteria were harvested by centrifugation at 6,000 × g for 6 min at 4 °C, resuspended in lysis buffer (20 mM Tris HCl, pH 7.5, 10 mM NaCl, 5 mM MgCl2, 1% Triton X-100, 5 mM DTT and protease inhibitor mixture [protease Mini EDTA-free, Roche]), and lysed using a French press (Thermo Fisher). After ultracentrifugation at 120,000 × g for 20 min at 4 °C the cleared lysate of GST-tagged GTases was added to equilibrated glutathione-Sepharose resin (GeneScript) and incubated for 1 h at 4 °C on a turning wheel. After four washing steps with wash buffer (20 mM Tris HCl, pH 7.5, 10 mM NaCl, 5 mM MgCl2) the bound protein was eluted with wash buffer supplemented with 10 mM reduced glutathione (Sigma-Aldrich).

Cleared lysate of HIS-tagged GTases was injected on Hitrap HP columns (GE Healthcare) after equilibration with binding buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 5 mM MgCl2) and subsequent volumes of binding buffer was followed by elution with 5 column volumes of elution buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 5 mM MgCl2, 500 mM imidazole). HIS-tagged GTases were incubated with 50 mM EDTA and further
purified by size exclusion chromatography (HiLoad 16/600 Superdex 75 pg, GE Healthcare) with SEC buffer (50 mM Hepes, pH 7.5, 150 mM NaCl, 5 mM MgCl2). EDTA was removed by buffer exchange (50 mM Hepes, pH 7.5, 150 mM NaCl, 5 mM MgCl2) and the protein was used for quantitative AMPylation assays.

**Nucleotide Loading of GTPases.** To preoad purified GTPases with the respective nucleotide, 50 µM protein was incubated with 3 mM nucleotide (GDP, GTP, GTP-γS, or GMP-PNP) and 8 mM EDTA in reaction buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM MgCl2) for 20 min at room temperature. Then 16 mM MgCl2 was added to stop the nucleotide exchange. The protein was then used for both in vitro AMPylation assays.

**Radioactive AMPylation Assay.** The in vitro AMPylation activity was assayed using either cleared bacterial lysates expressing full-length Bep1 or purified FIC domains of Bep1, VopS, and Ibpa.

To analyze the AMPylation activity of Bep1, Bep1αc*, VopSαc, and Ibpaαc*, 10 µM purified GTPase, preloaded with respective nucleotide, was incubated in presence of the respective AMPylator with 10 µCi [α-32P]-ATP (Hartmann Analytic) in reaction buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM MgCl2) for 60 min at room temperature. The reaction was stopped by addition of SDS-sample buffer and heating to 95 °C for 5 min. Samples were separated by SDS-PAGE and subjected to autoradiography.

For AMPylation of Rac1, Cdc42, and their mutant variants, 5 µM of purified His-tagged GTPases, preloaded with GDP, were incubated with Bep1αc*, (1 and 5 µM in Rac1 and Cdc42 variants, respectively) in the presence of [α-32P]-ATP (Hartmann Analytic) for 40 min in reaction buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM MgCl2) at 20 °C.

**Quantitative AMPylation Assay.** We employed an oIEC assay, monitoring the UV absorption of GTPase targets at 260 nm. The observed increase in absorbance due to AMPylation could be readily quantified and resulted in UV absorption of GTPase targets at 260 nm.

**To analyze and quantify the AMPylation efficiency (kcat/KM).** A 1-ml Resource Q column (GE Healthcare) was equilibrated with loading buffer (20 mM Tris-HCl, pH 8.5 or 6.5 for Rac1 or Cdc42, respectively). The purified AMPylated GTPases were then injected into the column buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM MgCl2) in a large volume (200 µL) and the reaction was started at t = 0 by addition of 3.2 mM ATP (final concentration, supplemented with 6.4 mM MgCl2). A small fraction (20 µL) of the reaction mixture was injected automatically on the column at intervals of 6 min. After washing with loading buffer, a gradient of elution buffer [1 M (NH4)2SO4 in loading buffer] was applied, yielding a chromatogram for each injection.

Reaction progress was monitored by quantification of GTPase peak area measured at 260 nm from each chromatogram by numerical peak integration. Note that this peak comprised both native and AMPylated GTPase. A heuristic quadratic function was fitted to the progress curves to yield the initial velocity. Calibration with ATP samples of known concentrations allowed absolute AMPylation velocities. Enzymatic kcat and KM parameters were derived from Vmax (S) type Michaelis–Menten plots (SI Appendix, Fig. S4 F and G). Depending on the activity, Bep1αc* concentrations were chosen such that the enzyme velocities were kept within a similar range (SI Appendix, Fig. S4 H and I). Nominal GTPase concentrations were corrected based on the back-extrapolated peak absorbance at t = 0. Fitting of single-substrate kinetic measurements by the Michaelis–Menten equation was developed in python 3 with standard modules provided in the Anaconda distribution.

**Crystalization and Structure Determination.** The full-length biaA gene that codes for the small ORF directly upstream of bep1 gene and part of the bep1 gene from B. rochalimae encoding the FIC domain (amino acid residues 13 to 229) were PCR amplified from genomic DNA. The PCR products for biaA and the fragment of bep1 were cloned into the vector pRSD-Duet1, pRSD-Duet1 containing biaA or bep1 were introduced into E. coli BL21 (DE3) by transformation. The constructs were expressed and purified as described for Vbhs/VbhT(FIC) (20) with the difference that 5 mM DTT was additionally used throughout the purification procedure. Fractions were pooled and concentrated to 13.6 mg mL⁻¹ for crystallization.

Crystals were obtained at 4 °C using the hanging-drop vapor diffusion method using mixing 1 µL protein solution with 1 µL reservoir solution. The reservoir solution was composed of 0.2 M Hepes (pH 7.5), 2.3 M ammonium sulfate, and 2% vol/vol PEG 400. For data collection, crystal was frozen in liquid nitrogen without additional cryoprotectant. Diffraction data were collected on-beam-line X06SA (PXII) of the Swiss Light Source (λ = 1.0 Å) at 100 K on a MAR CCD detector. Data were processed with XDS and the structure solved by molecular replacement with Phaser (42) using the VbhT/FIC structure (PDB 35HG) as search model. Several rounds of iterative model building and refinement were performed using Coot (43) and BUSTER (44), respectively. The final structure shows high similarity to the Vbhs/VbhT(FIC) structure (rmsd 1.4 Å for 183 Cx positions). Structural data are given in SI Appendix, Tables S3. Figs. 2 A–C and 3 A–D and SI Appendix, Figs. S2 B and D–F and S3A have been generated using Pymol (45).

**Homology Modeling of the Bep1:Target Complex and Generation of Structure-Based Sequence Alignments.** The input structure for homology modeling was chosen from all available Rac-subfamily structures (i.e., Rac1-3 and RhoG). In total, 43 PDB entries were analyzed (SI Appendix, Table S4), Cdc42 (chain D) of the Ibpa–Cdc42 complex served as reference for all superimpositions. The superimposition was carried out in two steps: a global superimposition over all Cx atom positions and a second, local superimposition using all atom positions of residues 27 to 37 (Sw1) of Cdc42. Both steps used the align-algorithm implemented in Pymol (version 1.8) with standard settings.

We observed high structural agreement between Rac-subfamily GTPase structures in the PDB and the reference chain with an average Cα rmsd below 0.5 Å. In contrast, we noticed large variations in the all-atom rmsds of residues in the Sw1 region that correlate with the nucleotide state of the GTPase. In order to find the most suitable PDB for homology modeling we selected the smallest root mean square deviations to the Cdc42 reference chain: three GDP-loaded GTPase structures display an rmsd of coordinates to the below 1 Å (SI Appendix, Table S4). Two of these structures are complexes of the Rho-GDP-dissociation inhibitor (RhoGDI) with either Rac1 (PDB ID 1hh4) or Rac2 (PDB ID 1ds6) representing the cytosolic storage form of the GTPases. The third structure is the Zn²⁺-bound trimeric form of Rac1 (PDB: 2P2L), in which Sw1 is involved in the Zn²⁺-mediated trimer interface. From these candidate PDBs, we selected the 1ds6 as the most appropriate for homology modeling since it represents a physiologic state of a Rac-GTPase (in contrast to 2P2L). Further, 1ds6 features a fully resolved Sw1 region and a higher resolution compared to entry 1hh4. To correspond closely to the reference structure, we built an alternative standard rotamer for the solvent-exposed Y32 of Rac2 in the PDB 1ds6 (Fig. 3C). The FIC domains of Bep1 and Ibpa were superimposed using the Cα atom positions of flap residues that adopt β-sheet-like conformation of the VbhA/Cdc42 complex. Superimposing Ibpaαc* residues 3,667 to 3,670 and 3,673 to 3,677, corresponding to Bep1 residues 110 to 113 and 122 to 126, respectively, yields an rms error of 0.87 Å for 9 CA pairs.

Modeling of the complex structure was carried out using the manually selected, superimposed, and curated model described above as starting structure for an adapted flexxDG protocol (26) implemented in the Rosetta package. For short, ligand-free (GDP- and hydrated Mg²⁺)- and ordered water molecules (as found in PDB entry 1ds6, as well as one water molecule in the center of the Bepl1 flap, shown in Fig. 2B) that are part of the protein interface were parameterized for the use in Rosetta and included in the modeling process to increase precision and validity of the resulting models. The selected small molecules had been refined with 8 factors that are comparable to neighboring main chain atoms in the respective PDB entries (1ds6 and Seu0). Next, the curated input model is subjected to a global minimization of backbone and side chain torsions in Rosetta (Minimize step) followed by local sampling of backbone and side chain degrees of freedom for all residues with Cx-atoms within 10 Å distance of Rac2 residue D124 (Backrub step). The side chains of the resulting models are optimized globally (Packing step), and backbone and side chain torsion energies are minimized globally (Minimize step 2). Finally, models are scored on the all-atom level using the suggested 2014 Rosetta function (20) with best scoring models were analyzed visually. The recommended total of 35 independent simulations is calculated for the complex with a maximum number of 5,000 minimization iterations (convergence limit score 1.0) and 35,000 backrub trials each step.

Structure guided multiple sequence alignments (MSA) were generated by manual adjustment of MSA generated using the ClustalW algorithm as implemented in the GENEIOUS software package (46) version 7.1.7.

**Quantification and Statistical Analysis.** Statistical parameters are given in SI Appendix, Tables S1 and S2. Error bars in quantitative AMPylation assays show the SD of reaction efficiencies (kcat/KM) derived from the least-square minimization of the fitting routine.
ACKNOWLEDGMENTS. We thank the laboratories of Klaus Aktories, Kim Orth, and Seema Mattoo for providing expression constructs of Rho-family GTases. VoP5, and Ibpa, respectively. We thank the beamline staff at the Swiss Light Source in Villigen for expert help in data acquisition. Rosetta modelling calculations were performed at sciCORE (https://scicore.unibas.ch/) scientific computing core facility at University of Basel. This work was supported by grant 310030A_173119 from the Swiss National Science Foundation (http://www.snf.ch/en/Pages/default.aspx) and advanced grant 340330 (FichelModFun) from the European Research Council, both to C.D.

1. J. Didsbury, R. F. Weber, G. M. Bokoch, T. Evans, R. Snyderman, rac, a novel ras-related family of proteins that are butylumon toxin substrates. J. Biol. Chem. 264, 16378–16382 (1990).

2. K. Wennerberg, K. L. Rossman, C. J. Der, The Ras superfamily at a glance. J. Cell Biol. 118, 843–856 (2002).

3. A. M. Yarbrough et al., Role of distinct type-IV-secretion systems and secreted effector sets in host adaptation by pathogenic Bartonella species. Cell. Microbiol. 21, e13004 (2019).

4. S. J. Heasman, A. J. Ridley, Mammalian Rho GTPases: New insights into their functions and downstream signaling. Nat. Rev. Mol. Cell Biol. 9, 493–507 (2008).

5. M. G. Bokoch, B. A. Diebold, Current molecular models for NADPH oxidase regulation by Rac GTPase. Blood 100, 2692–2696 (2002).

6. A. E. Karnoub, M. Symons, S. L. Campbell, C. J. Der, Molecular basis for Rho GTPase signaling specificity. Breast Cancer Res. Treat. 84, 61–71 (2004).

7. K. Aktories, Bacterial protein toxins that modify host regulatory GTPases. Nat. Rev. Microbiol. 9, 487–498 (2011).

8. K. Aktories, L. A. Quilliam, The insert region of RhoA is essential for Rho activation by Rac2 mutation. J. Biol. Chem. 283, 26946–26956 (2008).

9. A. Wagner, C. Dehio, Role of distinct type-IV-secretion systems and secreted effector sets in host adaptation by pathogenic Bartonella species. Cell. Microbiol. 21, e13004 (2019).

10. A. Harms, C. Dehio, Intruders below the radar: Molecular pathogenesis of Bartonella spp. Clin. Microbiol. Rev. 25, 42–78 (2012).

11. A. Harms et al., Evolutionary dynamics of pathoadaptation revealed by three independent acquisitions of the VirB/D4 type IV secretion system in Bartonella. Genome Biol. Evol. 9, 761–776 (2017).

12. A. Harms, F. V. Stanger, C. Dehio, Biological diversity and molecular plasticity of FIC domain proteins. Annu. Rev. Microbiol. 70, 341–360 (2016).

13. C. Hedberg, A. Itzen, Molecular perspectives on protein adenylylation. ACS Chem. Biol. 10, 212–219 (2015).

14. S. Mattoo et al., Comparative analysis of Histophilus somni immunoglobulin-binding protein A (IbpA) with other fic domain-containing enzymes reveals differences in substrate and nucleotide specificities. J. Biol. Chem. 286, 32834–32842 (2011).

15. C. A. Worby et al., The fic domain: Regulation of cell signalling by adenylylation. Mol. Cell 34, 93–103 (2009).

16. M. L. Yarbrough et al., AMPylation of Rho GTPases by Vibrio VoP5 disrupts effector binding and downstream signaling. Science 323, 269–272 (2009).

17. C. R. Roy, J. Cherni, Structure and function of Fic proteins. Nat. Rev. Microbiol. 13, 631–640 (2015).

18. R. Schulein et al., A bipartite signal mediates the transfer of type IV secretion substrates of Bartonella henselae into human cells. Proc. Natl. Acad. Sci. U.S.A. 102, 856–861 (2005).

19. A. Wagner, C. Tisches, C. Dehio, Versatility of the BID domain: Conserved function as type-IV-secretion-signal and secondarily evolved effector functions within Bartonella-infected host cells. Front. Microbiol. 10, 921 (2019).

20. P. Engel et al., Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. Nature 482, 107–110 (2012).

21. J. Xiao, C. A. Worby, S. Mattoo, B. Sankaran, J. E. Dixon, Structural basis of Fic-mediated adenylylation. Nat. Struct. Mol. Biol. 17, 1004–1010 (2010).

22. K. Lapouge et al., Structural basis of the TRP domain of p67phox in complex with Rac-GTP. Mol. Cell 6, 899–907 (2000).

23. P. Luong et al., Kinetic and structural insights into the mechanism of AMPylation by VoP5 Fic domain. J. Biol. Chem. 285, 20155–20163 (2010).

24. A. Goeppert, F. V. Stanger, C. Dehio, T. Schirmer, Conserved inhibitory mechanism and competent ATP binding mode for adenylyltransferases with Fic fold. PLoS One 8, e64901 (2013).

25. A. Wittinghofer, J. R. Vetter, Structure-function relationships of the G domain, a canonical switch motif. Annu. Rev. Biochem. 80, 943–971 (2011).

26. A. Wagner, C. Dehio, Role of distinct type-IV-secretion systems and secreted effector sets in host adaptation by pathogenic Bartonella species. Cell. Microbiol. 21, e13004 (2019).

27. A. Harms, C. Dehio, Intruders below the radar: Molecular pathogenesis of Bartonella spp. Clin. Microbiol. Rev. 25, 42–78 (2012).

28. A. G. Marangoni, Enzyme Kinetics: A Modern Approach (Wiley, Hoboken, NJ, 2002).

29. A. Harms et al., Evolutionary dynamics of pathoadaptation revealed by three independent acquisitions of the VirB/D4 type IV secretion system in Bartonella. Genome Biol. Evol. 9, 761–776 (2017).

30. I. R. Vetter, A. Wittinghofer, The guanine nucleotide-binding switch in three dimensions. Science 294, 1299–1304 (2001).

31. R. Donsky, M. R. Ahmadian, Always look on the bright site of Rho: Structural implications for a conserved intermolecular interface. EMBO Rep. 5, 1130–1136 (2004).

32. A. E. Karnoub, C. J. Der, S. L. Campbell, Theinsert region of Rac1 is essential for membrane ruffling but not cellular transformation. Mol. Cell. Biol. 21, 2847–2857 (2001).

33. H. Zong, K. Kaibuchi, L. A. Quilliam, The insert region of RhoA is essential for Rho kinase activation and cellular transformation. Mol. Cell. Biol. 21, 5287–5298 (2001).

34. J. L. Freeman, J. D. Lambeth, NADPH oxidase activity is independent of p47phox in vitro. J. Biol. Chem. 271, 22578–22582 (1996).

35. R. Rose et al., Structural and mechanistic insights into the interaction between Rho and mammalian Dia. Nature 439, 513–518 (2005).

36. S. Kühn et al., The structure of FNM1-L2-Cdc42 yields insights into the mechanism of lamellipodia and filopodia formation. Nat. Commun. 6, 7088 (2015).

37. M. Lammers, S. Meyer, D. Kühlmann, A. Wittinghofer, Specificity of interactions between mDia isoforms and Rho proteins. J. Biol. Chem. 283, 35236–35246 (2008).

38. H. Xue et al., Innate immune sensing of bacterial modifications of Rho-GTPases by the Pyrin inflammasome. Nature 513, 237–241 (2014).

39. D. R. Ambroz et al., Human neutrophil immunodeficiency syndrome is associated with an inhibitory Rac2 mutation. Proc. Natl. Acad. Sci. U.S.A. 97, 4654–4659 (2000).

40. A. G. Kurkchubasche, J. A. Panepinto, T. F. Tracy, Jr., G. W. Thurman, D. R. Ambroz, Clinical features of a human Rac2 mutation: A complex neutrophil dysfunction disease. J. Pediatr. 139, 141–147 (2001).

41. E. A. Lynch, J. Stahl, G. Schmidt, P. Chavrier, C. D’Souza-Schorey, Proteasome-mediated degradation of Rac1-GTP during epithelial cell scattering. Mol. Biol. Cell 17, 2236–2242 (2006).

42. A. J. McCoy et al., Phaser crystallographic software. J. Appl. Cryst. 40, 658–674 (2007).

43. P. Emley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).

44. G. Bricogne et al., BUSTER version X.Y.Z Cambridge (Global Phasing Ltd., United Kingdom, 2016).

45. J. Siebourg-Polster et al., NEMix: Single-cell nested effects models for probabilistic pathway stimulation. PLoS Comput. Biol. 11, e1004078 (2015).

46. M. Kearse et al., Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649 (2012).