Fic-mediated adenylylation: catalysis and regulation

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Arnaud Goepfert

von Frankreich

Basel, 2014

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Tilman Schirmer

Prof. Dr. Christoph Dehio

Basel, den 24.04.2012

Prof. Dr. Martin Spiess

Dekan

"Failure is the condiment which gives success its flavor."

Truman Capote

Statement to my Thesis

This work was carried out in collaboration between the group of Prof. Tilman Schirmer in the Core Program Structural Biology and Biophysics and the group of Prof. Christoph Dehio in the Focal Area Infection Biology at the Biozentrum of the University of Basel.

My PhD thesis committee consisted of:

Prof. Tilman Schirmer
Prof. Christoph Dehio

My thesis is written as a cumulative dissertation. It consists of a synopsis covering several aspects related to my work followed by result sections composed of three scientific publications and unpublished results. Finally, I sum up the major findings of my thesis and discuss some facets of this work.

Table of contents

Abstract

Abstract

Adenylylation, also referred to as adenylation or AMPylation, is the process by which adenosine-5'-monophosphate (AMP) is covalently attached to a protein, a nucleic acid, or a small molecule. It is a widespread post-translational modification employed by a large variety of enzymes to regulate multiple cellular functions. This modification, originally discovered in the 1960s, recently re-emerged as a modification used by bacterial effector proteins to regulate key host signaling events upon infection. The Type Three Secretion System effectors VopS and IbpA from *Vibrio parahaemolyticus* and *Histophilus somni*, respectively, harbor a FIC domain, which catalyzes adenylylation of Rho family GTPases. The covalent attachment of the bulky AMP moiety abrogates binding of downstream effectors resulting in actin cytoskeleton collapse and concomitant host cell death.

The FIC domain is not only present in a pathogenic setting but is also found in proteins that play a role in intrinsic signaling processes where its deleterious adenylylation activity needs to be tightly regulated.

In this study, I performed structural and biochemical analyses to unravel the structural determinants and mechanisms governing catalysis and regulation of Fic proteins as widespread signaling proteins.

First, by solving various structures of FIC domain-containing proteins I could show that FIC utilizes conserved active site features to favorably orientate the ATP substrate relative to the target protein, allowing AMP transfer to occur. We suggest that the catalytic mechanism can be generalized and extrapolated to all adenylylation-competent Fic proteins.

Second, I deciphered the structural mechanism controlling Fic-catalyzed adenylylation. An alpha helix characterized by a conserved [S/T]xxxE[G/N] motif tightly associates with the Fic active site and perturbs adenylylation-competent ATP binding. This alpha helix can be part of the Fic protein itself as an N- or a C-terminal extension, or can be provided by a separate antitoxin. Based on structural homology modeling, this regulation mechanism was found to be conserved from bacteria to higher eukaryotes.

The structural insights on the protein-ATP substrate binding specificity accumulated in this study will prove useful for coming efforts on rational drug design. Furtheremore, knowledge of the universal catalytic and inhibitory mechanism of Fic mediated AMP transfer will now pave the way for further studies towards the physiological roles of Fic proteins and particularly the identification of their protein targets.

1   Introduction

1.1 Post-translational modifications in cellular physiology and pathogenesis

1.1.1 Post-translational modifications: expand the structural and functional diversity of proteins

Post-translational modification (PTM) is a widespread mechanism used by prokaryotes and eukaryotes to increase the structural and functional diversity of the proteome in the cell. There are two routes of diversification of the proteome. While changes at the transcriptional and mRNA levels increase the size of the transcriptome relative to the genome[1,2], the combination of different post-translational modifications increases the complexity of the proteome and thus enriches information stored in the genome.

More than 300 PTMs have been discovered to date and they are often classified according to the mechanism used to alter the properties of a protein. Proteins can be modified by, the addition of functional groups (e.g., acetylation and glycosylation), the covalent attachment of other proteins (e.g., ubiquitination ans SUMOylation), changes of the chemical nature of amino acids (e.g., deamidation, eliminylation) or proteolytic cleavage (Fig. 1). This tremendous number of PTMs provides a highly dynamic and largely reversible regulatory network that controls diverse cellular processes such as protein-protein interaction, signaling processes[3], protein degradation[4], subcellular localization[5], and regulation of gene expression[6].
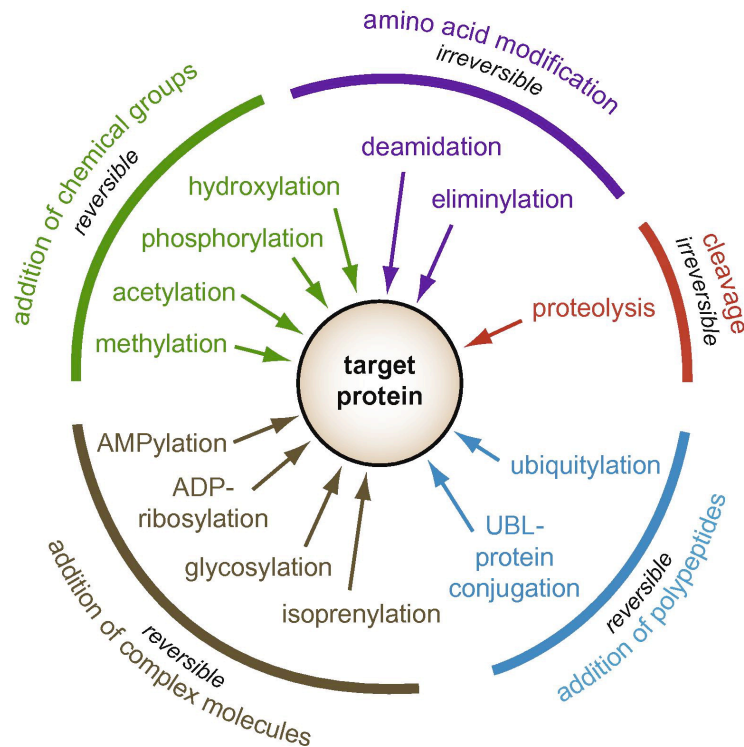
Figure 1: Diversity of post-translational modifications. This graph gives a non-exhaustive list of PTMs classified according to the nature of the modification occurring on the target protein. The reversibility of the modification is indicated. (Taken from Ribet et *al.*[7])

1.1.2   Post-translational modifications and pathogenesis

A hallmark of bacteria pathogenesis is the capacity of bacteria to survive and proliferate within their hosts. To this end, bacteria utilize PTMs as a key strategy to mimic the action of eukaryotic enzymes and interfere with host signaling pathways. PTMs can either be mediated by surface-localized or secreted effectors that cross the host membrane or are directly injected into the host cell via type III or IV secretions systems (T3SS or T4SS, respectively). These bacterial effectors display a large collection of enzymatic activities to modify key signaling molecules or to organize new signaling networks (Fig. 2).

Phosphorylation that consists in the tacking of a phosphate group to an amino acid is one of the most common PTMs utilized by bacteria to modify catalytic properties of host proteins. This reversible post-translational modification is mediated by the highly abundant protein kinases and is counter-balanced by dephosphorylation the removal of phosphate catalyzed by phosphatases. *Yersinia* species translocate several T3SS effectors into host cell that are required for survival and replication of the bacterium. Among these effectors are

YpkA and YopH which possess serine/threonine kinase and phosphatase activities, respectively, which enable tight control of the host actin cytoskeleton and are important to prevent macrophage-mediated phagocytosis[8,9].

Eliminylation is another modification that consists in the irreversible cleavage of a phosphate group. The phosphothreonine lyases OspF from *Shigella flexneri* and its homolog, SpvC of *Salmonella typhimurium,* mediate this unusual β-elimination reaction to dampen host immune response by irreversible inactivation of mitogen-activated protein kinases (MAPKs)[10].

MAPK kinases are also altered by effectors bearing **acetylation** activity. For example, VopA from *Vibrio parahaemolyticus* covalently modify active site lysine residue with an acetate moiety that obstructs the ATP binding site[11].

Glycosylation refers to the enzymatic process that attaches glycans to proteins. Toxins A and B, the major virulence factors of *Clostridium difficile* are the best characterized toxins that harbor glycosyltransferase activity. These toxins transfer a glucose molecule onto a specific threonine within the switch I region of Rho GTPases, locking them in a dominant-negative conformation[12,13].

ADP ribosylation, that is the transfer of ADP ribose from Nicotinamide Adenine Dinucleotide (NAD) is also a widely spread post-translational modification triggered by bacteria to manipulate host cell processes. For instance, clostridial C2 and C3 toxins mediate the attachment of one or more ADP-ribose moieties to actin and Rho-like proteins, respectively, to control actin cytoskeleton dynamics[14].

Some bacterial effectors share sequence and structural similarity with the E3 ubiquitin ligase, the enzyme that terminates the **ubiquitylation** process in eukaryotes. By adding polyubiquitin chains to the plant Fen kinase, the T3SS effector AvrPtoB from *Pseudomonas syringae* targets this host protein to the 26S proteasome for degradation, thereby inhibiting plant immune response[15].

Upon bacterial infection some eukaryotic proteins have also been reported to be deamidated or polyaminated, i.e. covalently modified with polyamines, such as spermidine, spermine and putrescine, upon bacterial infection. The *Bordetella* dermonecrotizing toxin (DNT) bears both functions of **polyamination** and **deamidation** to modify and render host Rho GTPases constitutively active, resulting in massive formation of actin stress fibers and focal adhesions[16].

Several bacterial proteins harbor **proteolytic activities**. YopT, a T3SS effector of *Yersinia enterocolitica* specifically recognizes prenylated Rho GTPases and irreversibly

inactivates them by a proteolytic cleavage near their carboxy-termini, leading to their release from the membrane. This cleavage disrupts the actin cytoskeleton and thus entails an antiphagocytic effect[17].

Phosphocholination, an unusual modification that requires CDP-choline as substrate has newly emerged as a potent mechanism to modulate Rab1/Rab35 GTPase. The *Legionella* pneumophila effector AnkX transfers phosphocholine onto these GTPases to regulate host vesicle transport[18]. The discovery of this unexpected modification highlights the importance of detailed investigations on baterial effectors and their interactions with host cell to further reveal unrecognized enzyme activities and post-translational modifications that govern a variety of distinct cellular functions.

Adenylylation is a recently discovered post-translational modification triggered by bacterial effectors to disrupt Rho GTPase signaling. The two virulence factors VopS from *Vibrio parahaemoliticus* and IbpA from *Histophilus somni* utilize ATP to covalently modify key residues of the Rho GTPases Cdc42, Rac1, and RhoA with an AMP moiety[19,20]. This leads to the collapse of the actin cytoskeleton and to cell death. Adenylylation will be discussed in detail in the following Chapters[19,20].
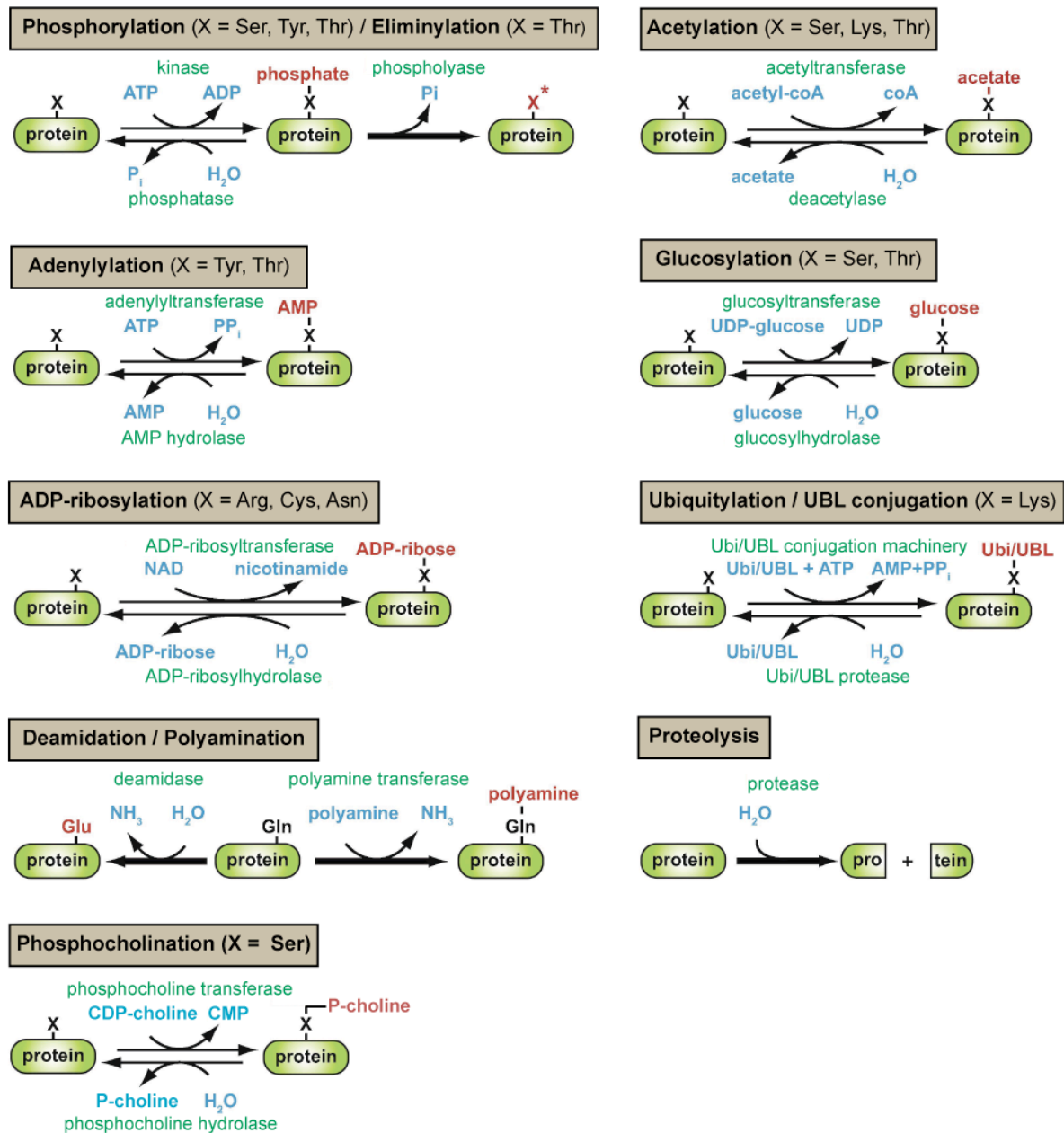
Figure 2: Common post-translational modifications triggered by bacterial pathogens. For each PTM, the enzyme responsible for the modification is in green, the substrate written in blue, and the molecule transferred onto the target protein is given in red. X, the nature of the amino acids modified; X*: irreversibly dephosphorylated amino acid; Pi, inorganic phosphate; PPi, inorganic pyrophosphate; NAD, nicotinamide adenine dinucleotide; coA, coenzyme A; Ubi, ubiquitin; UBL, ubiquitin-like proteins. (Adapted from Ribet et *al.*[7]).

1.2 Adenylylation

Adenylylation, also referred as adenylation or AMPylation, is the process by which adenosine-5'-monophosphate (AMP) is covalently attached to a protein, nucleic acid, or small molecule. This process involves nucleophilic attack of the α-phosphate of ATP by the target molecule leading to AMP transfer (Fig. 3). As phosphorylation, adenylylation results in a covalent modification with phosphorous-containing molecule to regulate multiple cellular functions. Depending on the nature of the attacking nucleophilic group, the AMP moiety can be attached via a stable phosphodiester (P-O) bond or via a transient phosphoramidate (P-N) or carboxylate-phosphate (CO-P) linkage[21] (Fig. 3).

Enzymes mediate adenylylation in order to achieve three main functions. As originally discovered in the 1960s, in the context of the Escherichia coli glutamine synthetase regulation[22], adenylylation is used as a stable post-translational modification to control protein activity.

The second function is to allow thermodynamically unfavorable enzymatic reaction to occur. In this case, transient adenylylation is used to activate an unreactive substrate by adding an efficient leaving group for subsequent nucleophilic attack[23]. In this case, the adenylylation reaction is transient, and involves the generation of less stable and thus more reactive carboxylate-phosphate anhydrides (CO-P) or phosphoramidate (P-N) bonds. This is in contrast to stable phosphodiester (O-P) bonds formed in the case of regulation of protein activity.

Finally, adenylylation occurs as an intermediate step along the biosynthetic pathway of protein cofactors[21].
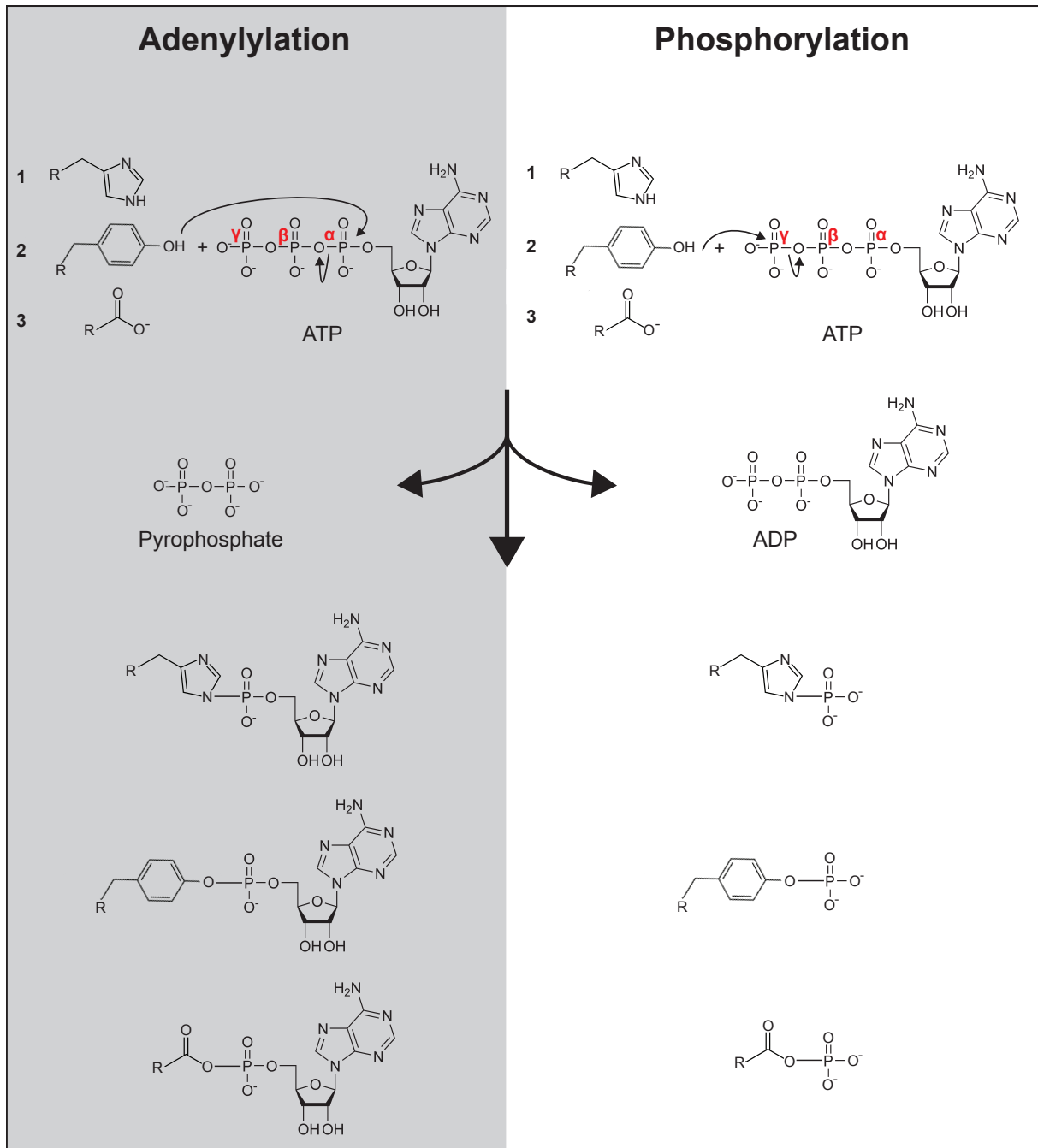
Figure 3: Protein adenylylation and phosphorylation. These two related modifications can occur at various positions on protein, peptide or amino acid and regulate a wide range of cellular processes. In both reactions, a nucleophile (such as the hydroxyl moiety of tyrosine (1), the side-chain amino functions of histidine (2), or the carboxy group of an amino acid (3)) attacks a phosphorous group of ATP. In adenylylation the attack takes place on the α-phosphate, thereby releasing pyrophosphate and transferring AMP. Phosphorylation is mediated by nucleophilic attack on the γ-phosphate, leading to phosphate transfer and release of ADP.

Whereas the phosphodiester bond created in the case of Tyrosine Tyr-AMP or Tyr-P is chemically stable under physiological conditions, His-AMP or His-P and RCOO-AMP or RCOO-P are unstable intermediates in enzymatic reactions.

1.2.1 Adenylylation is a stable post-translational protein modification

1.2.1.1 Glutamine synthetase adenylyltransferase

The ubiquitous glutamine synthetase (GS) is a key enzyme for nitrogen metabolism. It catalyses the ATP-dependent condensation of ammonia with glutamate to form glutamine[22], that is the principal source of nitrogen for protein and nucleic acid synthesis. Because of its critical role in nitrogen metabolism, this enzyme is found in all kingdoms of life from primitive to higher organisms[24]. GS has an elaborate control mechanism including a complex bicyclic cascade that controls its adenylylation state. The regulatory cascade contains the adenylyl transferase (ATase), a signal transduction enzyme called PII, and an uridylyl transferase (UTase).

The master regulator is the bifunctional enzyme, adenylyl transferase (ATase) that allows rapid fine-tuning of GS activity in response to drastically changing environments. ATase consists of two conserved nucleotidyltransferase (NT) domains separated by a central regulatory domain (R). The N-terminal adenylyl removase (AR) domain carries a deadenylylation activity and the C-terminal adenylyl transferase (AT) domain harbors an adenylylation activity. Both domains are able to interact with GS and regulate its function.

When the level of nitrogen is high, glutamine binds to AT and activates its adenylylation activity towards GS. The adenylylation occurs on a conserved tyrosine located below the glutamate active site. Thus, the bulky AMP moiety presumably prevents the binding of the glutamate substrate, resulting in low GS activity. Conversely, under nitrogen-limiting conditions, GS‑AMP is deadenylylated via phosphorolysis by the AR domain of ATase, restoring GS activity. It is noteworthy that the adenylylation and deadenylylation reactions catalyzed by ATase are not the reverse of each other. De-adenylylation occurs via phosphorolysis that requires a nucleophilic attack of inorganic phosphate on the AMP phosphate, thereby producing ADP.

The uridylyl transferase (UTase), a second bifunctional enzyme with nucleotidylylation/denucleotidylylation activity, provides another level of regulation. UTase directly senses the nitrogen level within the cell. When the nitrogen levels are low, UTase covalently attaches an UMP moiety onto the PII enzyme[25,26]. PII-UMP binds to the ATase regulatory domain (R) and stimulates its deadenylylation activity leading to GS-mediated

production of glutamine. Under conditions of nitrogen excess, UTase uses its phosphodiesterase activity to cleave the UMP from PII-UMP. The unmodified PII stimulates the adenylylation activity of ATase, decreasing the rate at which GS converts glutamate to glutamine (Fig. 4).
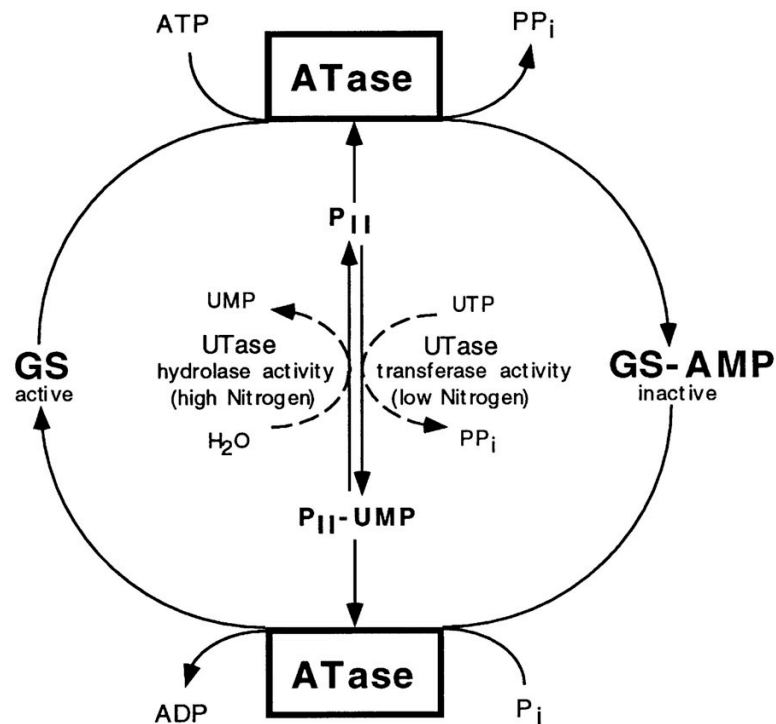
Figure 4: The bicyclic regulatory cascade that controls glutamine synthetase activity. The signal transduction protein PII is interconverted between its unmodified and its uridylylated form by the hydrolase and transferase activities of the UTase. The adenylylation and deadenylylation activities of adenylyl transferase (ATase) are regulated by PII and PII-UMP, respectively. Adenylylation of glutamine synthetase (GS) inhibits its activity, whereas deadenylylation restores its function. (taken from Jaggi et *al.*[27]).

1.2.1.2 Adenylylation of antibiotics

Bacteria have developed different resistance strategies to deal with the long way of antibiotic treatment. The most common route of resistance involves modification of antibiotica by acetylation, phosphorylation or adenylylation. The enzymes mediating these modifications are often plasmid encoded but can also be associated with transposable elements and are thus frequently transferred by conjugation. Adenylyltransferases (ANTs) that trigger adenylylation of antibiotics belong to the nucleotidyltransferase (NT) superfamily, and share structural and functional similarity with the DNA polymerase B and the poly(A) polymerase[28]. ANTs confer to the bacteria resistance against a variety of antibiotic classes, such as aminoglycosides or lincosamides. Based on X-ray crystallography structures, the catalytic mechanism has been proposed to be an in-line nucleophilic attack of the antibiotic hydroxyl group onto the α-phosphate of ATP leading to an AMP-modified antibiotic and pyrophosphate release[28,29] (Fig. 5). These bacterial nucleotidyl-transferases are not specific for ATP and can also mediate guanylylation or uridylylation through GTP or UTP transfer, respectively[30].
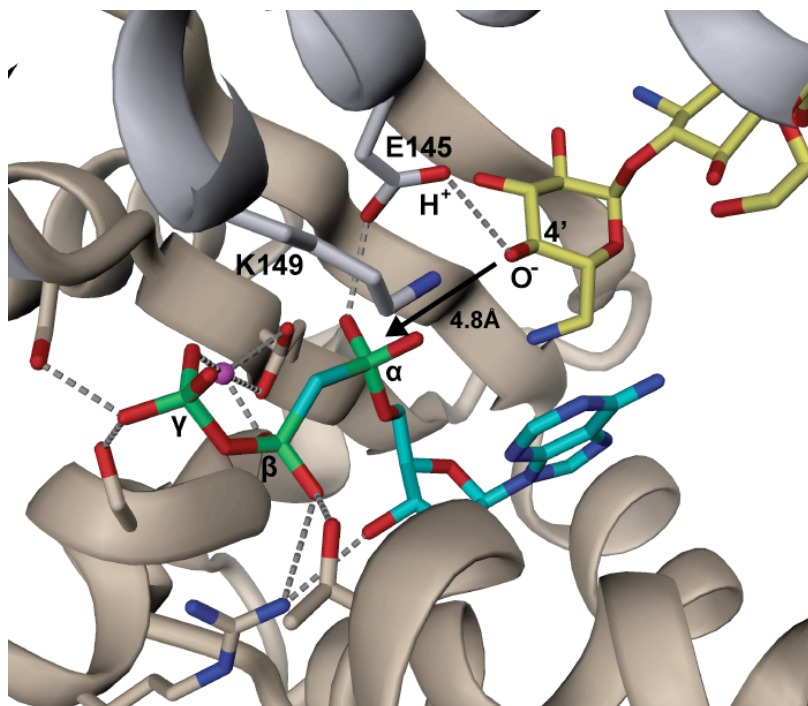
Figure 5: Mechanism of antibiotic adenylylation. Close-up view of the *S. aureus* Kanamycin adenylyl-transferase active site (PDB code 1KNY). The non-hydrolyzable ATP analog AMPCPP and kanamycin are shown in cyan and yellow, respectively. Residues forming interactions with the ligands are shown in full and the magnesium ion with a magenta ball. Hydrogen-bonding interactions are shown as dashed lines. The ligand-binding pocket is created by two monomers shown in grey and beige that wrap one another. The 4' hydroxyl on the kanamycin is located at a distance of 4.8Å from the α-phosphate of the nucleotide and both substrates are in the proper orientation for a single in-line displacement mechanism. The enzyme has to undergo a small conformational change to decrease distance between the two reactants. E145 that is situated within hydrogen-bonding distance to this hydroxyl group acts as a general base and abstracts the proton of the hydroxyl group, leaving kanamycin susceptible for nucleophilic attack on the α-phosphate of ATP. K149, that is close to the α-phosphorous, and Mg^{2+} ion stabilize the pentacovalent transition state and the pyrophosphate leaving group, respectively. ATP is locked in the binding pocket by an extensive hydrogen-bonding network through its phosphate tail and the ribose. The absence of contacts between the adenine ring of ATP and the adenylyl transferase active site is consistent with the lack of specificity for the nucleotide substrate[28].

1.2.2   Activation of unreactive substrate by adenylylation

Adenylylation is often use as a transient modification that increases the reactivity of a molecule[23]. The high energy of AMPylated intermediate will drive subsequent non-spontaneous reactions.

1.2.2.1 Aminoacyl-tRNA synthetases

Aminoacyl-tRNA synthetase uses a transient adenylylation in order to charge tRNA with its cognate amino acid. The synthetase binds simultaneously ATP and the amino acid and facilitates the nucleophilic attack of the amino acid carboxylate group onto the ATP α-phosphorous. As a second step, the appropriate tRNA binds to the synthetase and a high-energy bond between the 2'- or 3'-hydroxyl of the terminal adenosine of tRNA and the adenylylated amino acid is formed, displacing the AMP[31] (Fig. 6). The aminoacyl-tRNA is then released from the enzyme and brings the correct amino acid to the ribosome. The latter will transfer the amino acid to the nascent polypeptide chain whose sequence is specified by the codons of the mRNA.
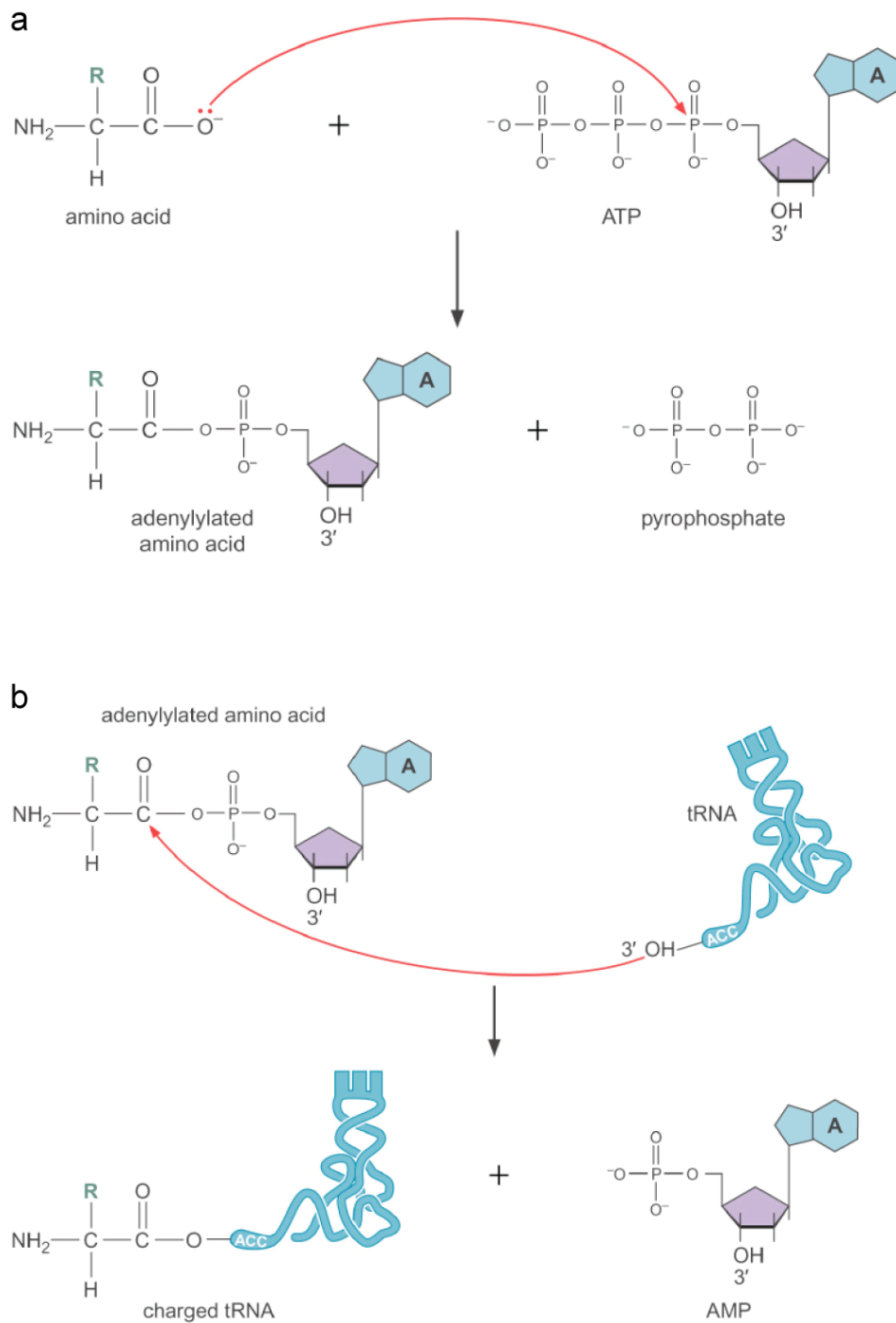
Figure 6: tRNA charging by aminoacyl-tRNA synthetases (aaRSs). aaRSs utilize a two-step mechanism to charge a specific tRNA with its cognate amino acid. (a) The first step involves a nucleophilc attack of the amino acid carboxyl group on the α-phosphate of ATP, generating an adenylylated amino-acid and pyrophosphate.
(b) In step 2, the adenylylated amino acid carbonyl is attacked by the 3' OH of the tRNA CCA ending, displacing AMP.

1.2.2.2 E1 activating enzyme from the ubiquitin-conjugation pathway

Another example of transient adenylylation is the E1 activating enzyme from the ubiquitin-conjugation pathway. Ubiquitylation, i.e. covalent attachment of ubiquitin or ubiquitin-like proteins (UBLs), is one of the major post-translational modifications used to regulate eukaryotic protein functions[32]. The effect of this modification on the fate of the targeted protein depends on the length of the ubiquitin molecules appended to the protein. Polyubiquitylated protein substrates are often targeted to the 26S proteasome for degradation but single ubiquitin monomers are also involved in non-proteolytic regulatory mechanisms, such as endocytosis, immune responses, DNA repair, or histone regulation[33-36].

Attachment of ubiquitin involves a multienzyme cascade with the first step mediated by the E1 enzyme (Fig. 7). E1 binds ATP/Mg^{2+} and the ubiquitin molecule and catalyse adenylylation of the C-terminal carboxylate group of ubiquitin to produce an ubiquitin-adenylylate intermediate[37]. In a second step, a catalytic cysteine sulphydryl group from E1 attacks the acyl-adenylylate to form an activated E1-ubiquitin thioester intermediate. Ubiquitin is then transferred to a cysteine residue in ubiquitin-conjugating enzyme E2 through a trans(thio)esterification reaction. In the final step of the ubiquitylation cascade, ubiquitin molecules become covalently attached to lysine residues in target proteins by the action of the E2 with the assistance of ubiquitin ligase E3 (Fig. 7).

Concomitantly, other UBL–modification pathways, such as SUMOylation (attachment of SUMO molecule) or NEDDylation (attachment of Nedd8 molecule) require initial activation of carboxylate group by adenylylation[38,39].
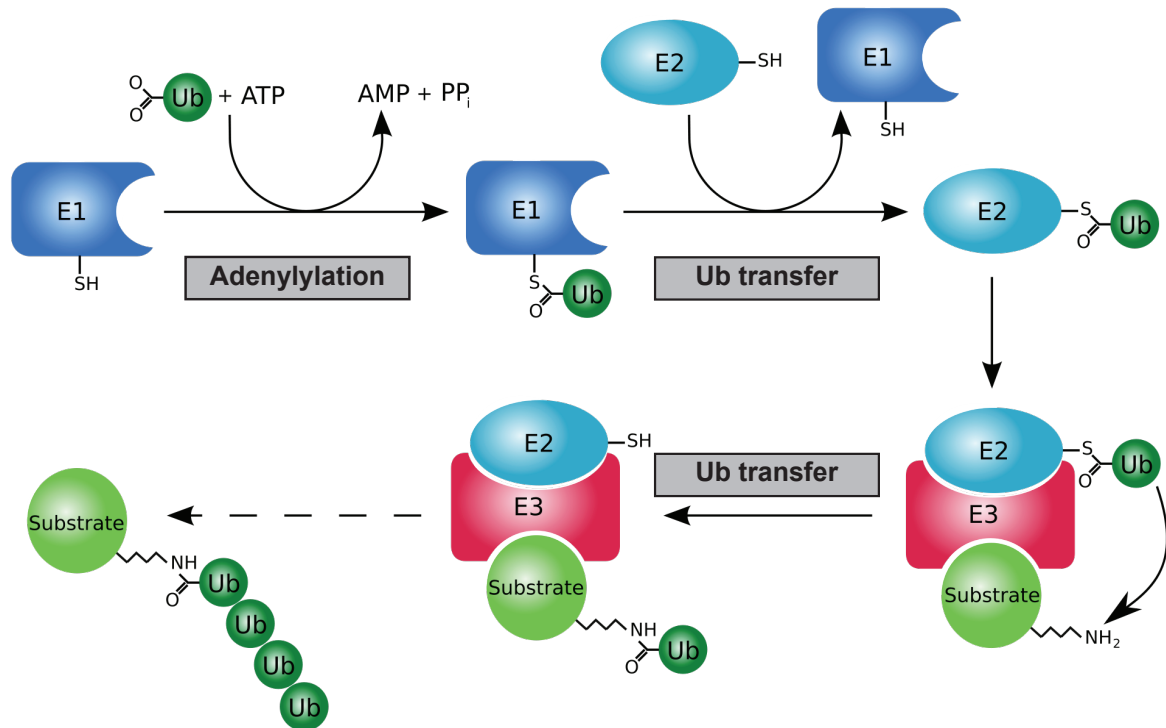
Figure 7: The Ubiquitylation pathway. Ubiquitylation requires a cascade of E1 ubiquitin-activating, E2 ubiquitin-conjugating and E3 ubiquitin ligase enzymes. First, E1 enzyme activates the ubiquitin molecules by adenylylation of the carboxy-terminal carboxylate of ubiquitin (Ub), forming Ub–AMP. The second step transfers ubiquitin to the E1 active site cysteine residue, with release of AMP. This step results in a thioester linkage between the C-terminal carboxyl group of ubiquitin and the E1 cysteine sulfhydryl group (E1-Ub). Ubiquitin is then transferred from E1 to E2 via a trans(thio)-esterification reaction (E2-Ub). Finally, E2, in conjunction with E3 covalently attached ubiquitin to the protein target. Typically, an isopeptide bond is formed between the ubiquitin C-terminal carboxylate and the ε-amino group of a Lys side chain of the substrate protein or the growing ubiquitin chain. (Adapted from Maupin-Furlow et *al.*[33]).

1.2.2.3 Glucose-1-phosphate adenylyltransferase

Many organisms synthesize and store glycogen as energy reserves to cope with starvation conditions. The pathway for synthesis of storage polysaccharides requires initial activation of glucose units for subsequent transfer to the growing α-1,4-glucosidic chain. The first committed step is catalyzed by Glucose-1-phosphate adenylytransferase (also named ADP-glucose pyrophosphorylase) that mediates formation of ADP-glucose, the activated form of glucose, from substrates glucose-1-phosphate and ATP[40] (Fig. 8). The reaction proceeds via attack by a phosphate oxygen of glucose-1-phosphate on the α-phosphorous of ATP, with release of the pyrophosphate moiety[41]. ADP-glucose is then used as a substrate by the glycogen synthase to synthesize α-1,4-polyglucan chains. Whereas, ADP-glucose represents the activated form of glucose in bacteria and plants, mammals and fungi possess glucose-1-phosphate uridylyltransferases that pre-activate glucose-1-phosphate molecules with UMP during carbohydrate metabolism[42].
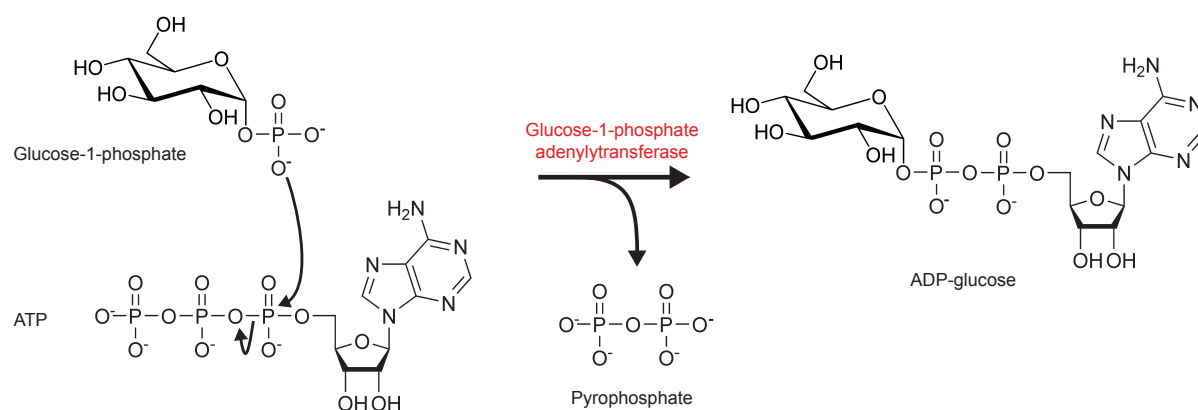
Figure 8: Schematic of glucose activation by the Glucose-1-phosphate adenylyltransferase. Upon binding of the two substrates, namely glucose-1-phosphate and ATP, glucose-1-phosphate adenylyltransferase facilitates the nucleophilic attack of one phosphate oxygen from glucose-1-phosphate onto the α-phosphorous of ATP. This reaction induces formation of ADP-glucose together with release of pyrophosphate.

1.2.3 Adenylylation step is common in the synthesis of organic cofactors

Cofactors are non-protein chemical compounds that are bound to proteins, commonly enzymes, and assist biochemical reactions. Two types of cofactors exist in nature: inorganic cofactors such as metal ions and organic cofactors, which are often derived from vitamins. NAD$^+$, FAD, and Coenzyme A are essential organic cofactors that require adenylyltransferases for their biosynthesis.

Nicotinamide adenine dinucleotide (NAD$^+$) that is derived from the niacin vitamin is a pivotal coenzyme that plays fundamental roles in cellular metabolism. It is involved in redox reactions, DNA repair, stress responses and serves as a substrate in ADP-ribosylation[43,44]. The critical step in NAD generation involves the Nicotinamide/nicotinic acid mononucleotide adenylyltransferase (NMNAT). The enzyme catalyzes the formation of the dinucleotide through condensation of ATP with nicotinamide mononucleotide (NMN)[45] (Fig. 9a).

The flavo-coenzyme flavin adenine dinucleotide (FAD) is another important redox cofactor[46]. Its synthesis requires a two-step mechanism. In eukaryotes, riboflavin, also known as vitamin B2, is first phosphorylated by means of riboflavin kinases (RFK) to yield flavin mononucleotide (FMN). FMN is further adenylylated by Flavin mononucleotide adenylyltransferase (FMNAT) to form FAD (Fig. 9b). In bacteria, synthesis is mediated by the bifunctional enzyme flavin adenine dinucleotide synthetase (FADS), which bears both phosphorylation and AMP transfer activities[47].

AMP transfer also occurs in the course of Coenzyme A (CoA) biosynthesis pathway. The pantothenate (vitamin B5) is converted in CoA in a five-step process. An intermediate formed en route is dephospho-CoA (dPCoA). dPCoA is formed by Phosphopantetheine adenylyltransferase (PPAT), in a adenylylation reaction that requires phosphopantetheine (Ppant) and ATP (Fig. 9c). dPCoA is further utilized by dPCoA kinase that produces the acyl group carrier, CoA[48].
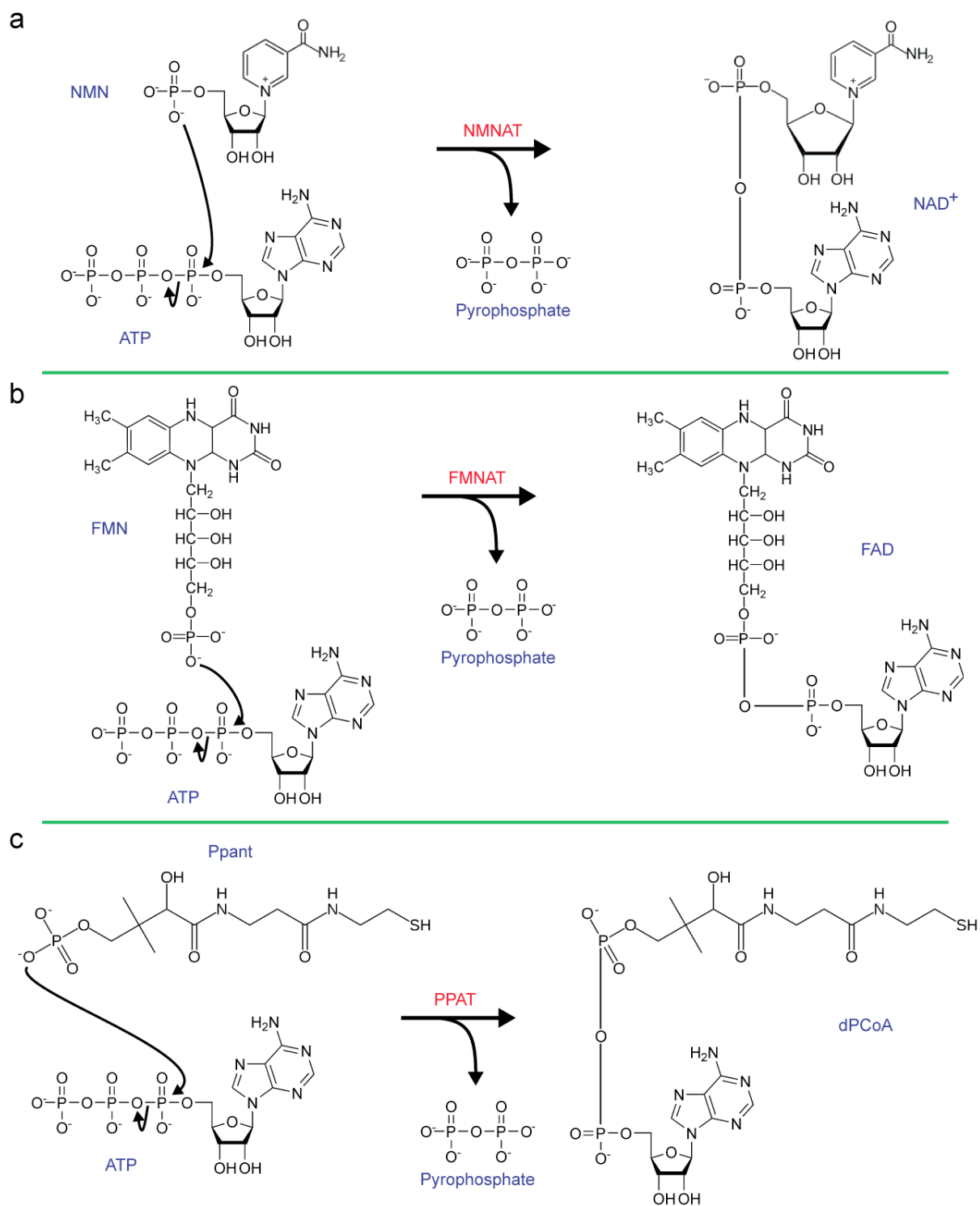
Figure 9: General reaction schemes for the adenylylation process of the vitamin-derived cofactors NAD^+ (**a**), FAD (**b**), and Coenzyme A (**c**). Adenylylation is mediated by Nicotinamide/nicotinic acid mononucleotide adenylyltransferase (NMNAT), Flavin mononucleotide adenylyltransferase (FMNAT), and Phosphopantetheine adenylyltransferase (PPAT), respectively. NMN: Nicotinamide mononucleotide; NAD^+: Nicotinamide adenine dinucleotide; FMN: Flavin mononucleotide; FAD: flavin adenine dinucleotide; Ppant: Phosphopantetheine; dPCoA: Dephospho-Coenzyme A.

These three different adenylyltransferases involved in cofactor synthesis share remarkable structural similarities. They belong to the large nucleotidyltransferase α/β phosphodiesterase superfamily, to which class I aminoacyl tRNA synthetases (aaRS) belong, and adopt the typical nucleotide binding Rosmann fold with a conserved signature (H/T)xGH motif important for ATP recognition and stabilization[49-51]. Superimposition of NMNAT and PPAT also reveals a common binding pocket for the chemically dissimilar substrates, NMN and phosphopantetheine, respectively, suggesting a conserved catalytic mechanism[50] (Fig. 10). The adenylylation process involves nucleophilic attack of a phosphate oxygen from the respective substrates on the α-phosphorous of the adenine nucleotide in an in-line displacement mechanism (Fig. 9 and Fig. 10). Since the phosphate group is a reasonable nucleophile and does not require pre-activation, residues from the enzyme active site are not engaged in the chemistry of the reaction by acid/base catalysis (Fig. 10) (in contrast to antibiotic adenylyl-transferases, see fig. 5). An equivalent mechanism was also reported for tRNA amino-acylation by class I glutaminyl tRNA synthetase (GlnRS)[52] (Fig. 10C).
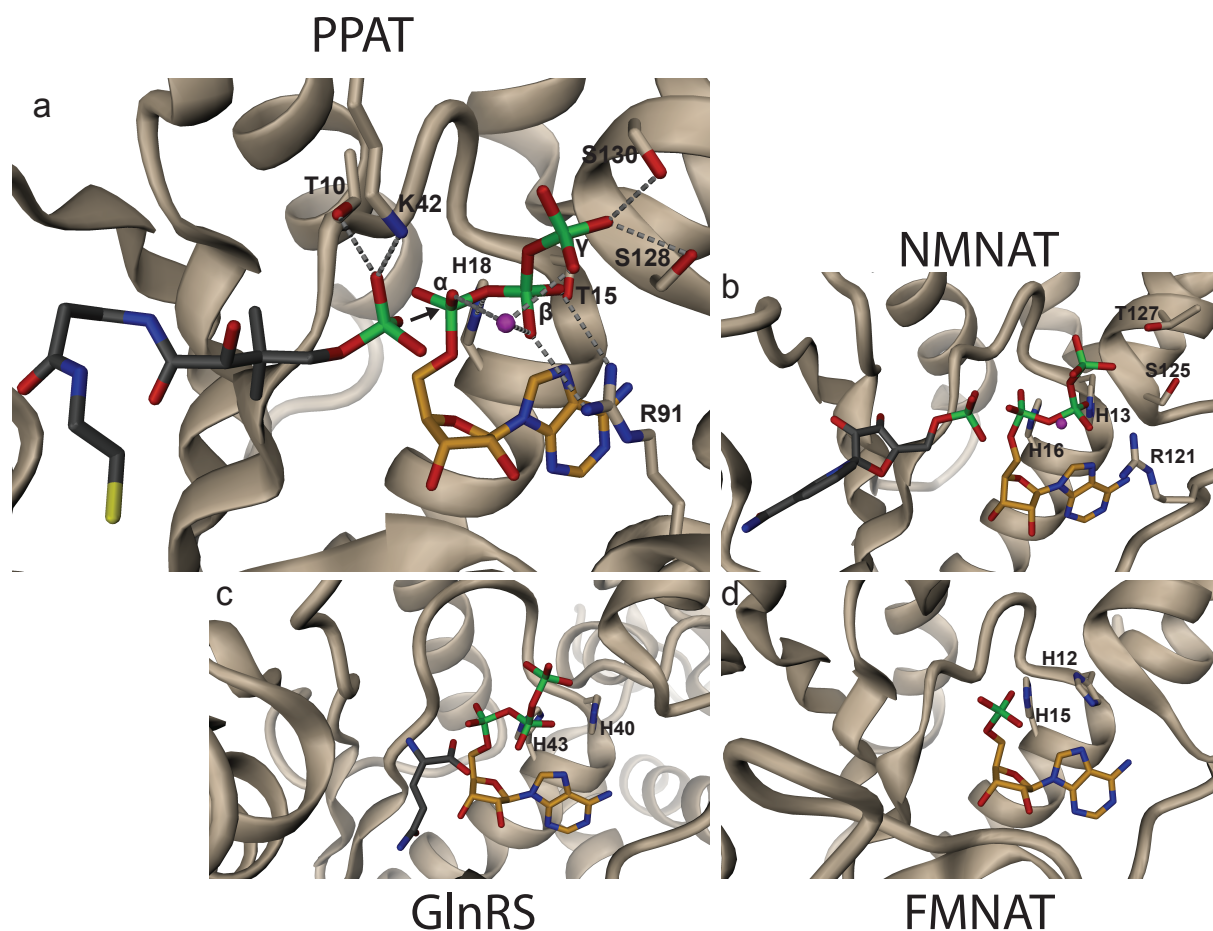
Figure 10: (*legend on next page*)

Figure 10: Close-up view of PPAT, NMNAT, GlnRS and FMNAT active sites. The similar nucleotide binding and the highly conserved H/TxGH sequence motif, involved in nucleotide recognition and stabilization suggest a common catalytic mechanism for enzymes of the nucleotidyltransferase α/β phosphodiesterase superfamily.

(**a**) Zoom-in of the *E. coli* Phosphopantetheine adenylyltransferase (PPAT) active-site with its two substrates obtained after superimposition of structures of PPAT/ATP-Mn^{2+} complex (PDB code 1GN8) with PPAT/4'-phosphopantetheine (Ppant) complex (PDB code 1QJC). ATP and Ppant substrates are shown in full and colored orange and grey, respectively. Mn^{2+} ion is represented with a ball (in magenta). Residues important for substrates recognition are depicted in full with H-bonds marked by stippled lines in grey. Arg91, Ser128, Ser130 and the divalent cation facilitate the adoption of an ideal geometry for the pyrophosphate leaving group. Lys42 and Thr10 are essential to orientate the attacking nucleophile for catalysis. His18 from the conserved T/HxGH motif serves as a proton donor and interacts with the α-phosphate group to stabilize the transition state. Thus, electrostatic support and locking the reaction partners in a productive conformation is all that is required for catalysis to happen. The phosphate of Ppant undergoes nucleophilic attack on ATP α-phosphorous in an in-line displacement mechanism[51].

(**b**) Close-up view of the Nicotinamide mononucleotide adenylyltransferase (NMNAT) active-site with its two substrates obtained after superimposition of structures of *M. Jannaschii* NMNAT/ATP complex (PDB code 1F9A) with human NMNAT/nicotinamide mononucleotide (NMN) complex (PDB code 1GZU). ATP and NMN are colored in orange and grey, respectively. Residues involved in ATP binding (His16, Arg121, Ser125, Thr127) are equivalent in PPAT (His18, Arg91, Ser128, Ser130) and the 2 substrates adopt an orientation similar as seen in (A), supporting the mechanism proposed for PPAT.

(**c**) Zoom-in of the *E. coli* Glutaminyl-tRNA synthetase (GlnRS) active site. The respective positions of the ATP and glutamine substrates have been obtained after superposition of GlnRS/ATP (PDB code 1GTR) with GlnRS/Glutamine (Gln) (PDB code 1O0B) structures. ATP and Gln are shown in orange and grey, respectively. The amino acid activation by adenylylation is mechanistically equivalent as in A and B with the second histidine from the T/HxGH motif stabilizing the pentacovalent transition state. However, in this situation, the carboxyl group of Gln is not ideally positioned in front of the pyrophosphate leaving group and catalysis requires slight conformational changes that is induced by transfer RNA binding[53].

(**d**) Active-site of the flavin mononucleotide adenylyltransferase (FMNAT) domain of *T. maritima* flavin adenine dinucleotide synthetase (FADS) in complex with AMP (PDB code 1T6Y). The AMP conformation, which resembles the one observed for the nucleotides in A, B and C, and the presence of the conserved T/HxGH are reminiscent of a similar catalytic mechanism.

Finally, adenylyltransferases intervene in the formation of the organic, non-vitamin molybdenum (Moco) and 3'-Phosphoadenosine-5'-phosphosulfate (PAPS) cofactors.

Moco-containing enzymes catalyze important redox reactions in the global carbon, sulfur, and nitrogen cycles. It is synthesized by an ancient and conserved biosynthetic pathway that can be divided into five steps, one of which is adenylylation[54]. In *E. coli*, the enzyme mediating adenylylation is named MoeB and activates the C-terminal end of MoaD protein by forming an acyl-adenylate. Subsequently, a sulphurtransferase converts the MoaD acyl-adenylate to a thiocarboxylate that provides sulphur during Moco biosynthesis[55]. MoeB is structurally and functionally similar to E1 enzyme from the ubiquitylation pathway[56,57] (see Chapter 1.2.2.2).

The cofactor PAPS is the sulfate donor in sulfotransferase-catalyzed synthesis and is an important component of the biosynthesis pathway of sulfur-containing amino acids[58]. An intermediate formed during PAPS biosynthesis is adenosine-5′-phosphosulfate (APS), a derivative of adenosine monosphosphate that has a sulfate group attached to the 5' phosphate. APS is formed by sulfate adenylyltransferase, also named ATP sulfurylase, in an adenylylation reaction that requires ATP and sulfate ion. APS is further phosphorylated at the ribose 3' hydroxyl by APS kinases, to yield PAPS[59].

1.2.4 Pathogen-mediated adenylylation

Adenylylation has emerged recently as a post-translational modification hijacked by bacterial pathogens to subvert key host cellular processes upon infection. *Legionella pneumophila* is an opportunistic human pathogen and the causative agent of Legionnaires' disease. A hallmark of *Legionella* pathogenesis is the capacity of the bacteria to form endoplamic reticulum-derived vacuoles, called *Legionella*-containing vacuoles (LCV), to avoid fusion with endocytic vesicles and lysosomes and to allow bacterial survival and proliferation[60]. To this end, *L. pneumophila* manipulates intracellular vesicular trafficking through secretion of effectors via the T4SS-related Dot/Icm (defect in organelle trafficking/intracellular multiplication) translocation machinery[61]. This effectors co-opt host factors, such as Sar1, Sec22b and Rab1 to establish the LCVs[62]. In this context, the versatile *L. pneumophila* factor DrrA (also known as SidM) is involved in the recruitment and manipulation of the small GTPase Rab1 (Fig. 11). DrrA contains three domains. Whereas the C-terminal lipid phosphatidylinositol-4-phosphate binding domain (P4M) mediates attachment to the LCVs membrane, the central guanine nucleotide exchange factor (GEF) and the N-terminal adenylyl transferase activities subvert Rab1 functions (Fig. 11). The N-terminal domain displays high structural similarity with the glutamine synthetase adenylyl transferase, with which it shares also the same catalytic residues[63] (Fig. 11c, see Chapter 1.2.1.1). It efficiently transfers AMP to a conserved tyrosine in the switch II region of Rab1-GTP (Tyr77 in Rab1b), thereby inhibiting GAP-stimulated GTP hydrolysis and trapping Rab1 in an active state[63]. Interestingly, later during infection, L. pneumophila also injects SidD, an effector that possesses AMP hydrolase activity to counteract DrrA function[64,65]. Combination of effectors with opposite activities allows *L.pneumophila* to achieve an accurate spatio-temporal control of the Rab1 signalling pathway.

Likewise, bacterial effectors carrying a FIC domain catalyse adenylylation of host cell proteins. The Fic-mediated adenylylation will be extensively discussed in the following chapters.
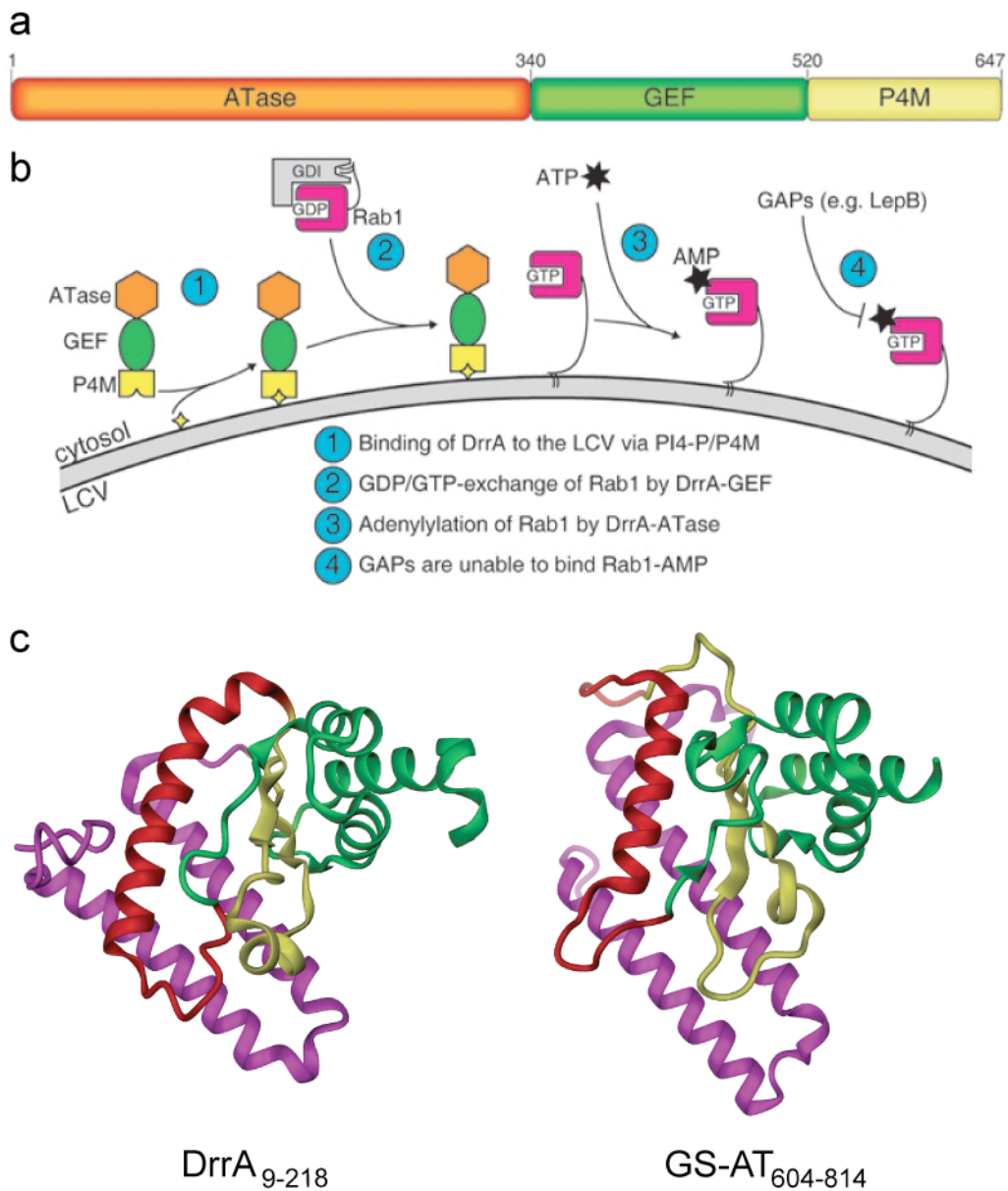
Figure 11: Structure and function of DrrA. **a)** Tripartite domain architecture of DrrA with a N-terminal adenylyltransferase (ATase), a central GTP-GDP exchange factor (GEF) and a C-terminal lipid phosphatidylinositol-4-phosphate binding domain (P4M). **b)** Mechanism of Rab1 modulation by DrrA. DrrA docks on the LCVs through its P4M domain (step 1) where it recruits Rab1, displaces it from GDI and activates it by stimulating GDP/GTP exchange (step 2). Rab1 is then adenylylated by DrrA (step 3) preventing its inactivation by GAPs (step4). (yellow star, phosphatidylinositol-4-phosphate ; grey arc, LCV membrane ; GDI, guanine nucleotide dissociation inhibitor). **c)** Comparison of crystal structures of a DrrA fragment (PDB code 3NKU) and a fragment of the C-terminal adenylytransferase domain of GS-AT (PDB code 3K7D), which highlights high structural similarity. (Adapted from Goody et al.[66]).

1.3 FIC domain: a new paradigm for adenylylation

1.3.1 General introduction

FIC stands for filamentation induced by cAMP and was initially identified as a regulatory factor of cell division. Utsumi et *al*, discovered that a specific mutation (G55R) in the *fic* gene perturbed normal cell division and led to a filamentation phenotype when bacteria were grown at 43°C in presence of 1.5mM[67-73] (Fig. 12). Additional investigation on this *fic* mutant also suggested that the regulatory mechanism might involve folate metabolism by the synthesis of *p*-aminobenzoic acid (PABA) or folate[70].

Figure 12: E. coli Fic mutant G55R shows impaires cell division when incubated in high cAMP concentration resulting in filamentous cells. (Adapted from Komano et *al*.[70])

The enzymatic activity of FIC domains was only recently revealed with the work on the T3SS effector proteins VopS from *Vibrio parahaemolyticus*. This gram-negative halophilic bacterium, found in estuarine waters worldwide[74], is a causative agent of seafood-borne human gastroenteritis and can lead to death in immune-compromised or burdened patients[75]. These bacteria use a type III secretion system (T3SS) to deliver bacterial effector proteins into the host cell cytosol where they disrupt the signaling network[76] and induce autophagy, cell rounding and subsequent cell lysis[77]. In this context, VopS was implicated in the severe

rounding of cells by the inactivation of the Rho family guanosine triphosphatases (GTPases) members, Rac1, RhoA, and Cdc42[78]. Using pull-down experiments in combination with mass spectrometry analysis, Yarbrough et al[20], deciphered the mechanism of VopS-mediated GTPases inactivation. They observed that upon incubation with VopS, a specific threonine (T35 in Rac1 and Cdc42 and T37 in RhoA) in the switch I loop of the GTPases was covalently modified with an AMP moiety. Furthermore, they attributed AMP transferase activity to the FIC domain since mutation of the conserved histidine residue in the FIC active site motif (HxFx[D/E]GNGRxxR) rescued the rounded phenotype (Fig. 13). Owing to the critical role of the switch I region in downstream protein binding, the bulky AMP group disrupts the GTPase signaling cascade resulting in cell death[20].
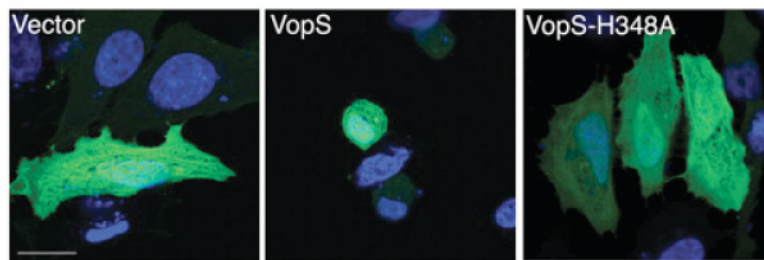
Figure 13: Human HeLa cells transfected with VopS round up. This phenotype, typical of actin cytoskeleton collapse, results from inactivation of Rho GTPase signaling. Mutant of the Fic active-site histidine displays normal cell shape, proving that the activity resides in the FIC domain. (Adapted from Yarbrough et al[20])

A parallel research on the *Histophilus somni* IbpA protein shed light on a similar Fic-catalyzed adenylylation and disruption of GTPase signaling. *H. somni* is a gram negative pathogen that causes a variety of syndromes in cattle and other ruminants, including pneumonia, septicemia, myocarditis, and abortion[79-82].

One of the virulence factors of *H. somni* is IbpA (immunoglobulin-binding protein A), a fibrillar surface antigen that contains two FIC domains at its C-terminus. The expression of the FIC domains in HeLa cells but also in the most physiologically relevant primary bovine alveolar type 2 (BAT2) cells led to a cell rounding and retraction phenotype[83]. As in the case of VopS, this phenotype originates from the blockade of GTPase signaling by Fic-mediated adenylylation of the Rho GTPases[19].

Interestingly, the modification was traced down to a tyrosine in the switch I region of Rho GTPases (Y32 in Rac1 and Cdc42 and Y34 in RhoA). It remains to be elucidated whether the variation in the type of amino acid that is modified has further implications in the regulation of Rho GTPase signaling events[84].

The same group also demonstrated that the huntingtin yeast-interacting protein E (HypE), the only human FIC-domain containing protein, can also adenylylate Rho GTPases *in vitro*[19]. It is unclear whether HypE inactivates GTPases *in vivo* and the physiological substrates remain to be identified. A yeast two-hybrid screening identified HypE as a binder of the huntingtin protein[85]. It is therefore tempting to speculate that HypE might play a role in the Huntington's disease by direct covalent modification of the huntingtin protein or through an indirect mechanism by targeting huntingtin-interacting partners.

The FIC domain is an evolutionary conserved domain that is ubiquitous in bacteria but also found in eukaryotes, archae and viruses but, noteworthy, not in plants and yeast. At the time of writing, more than 4500 sequences are present in the Pfam protein database[86] with the vast majority harboring the consensus HxFx[D/E]GNGRxxR motif (Fig. 14). For all residues of the motif except for the first glycine and first arginine, single point mutant have been generated in various Fic proteins that showed reduced adenylylation activity[19,20,87,88]. Therefore, integrity of the FIC motif is essential for catalysis. Fic proteins with degenerated active centers are likely to carry other or no enzymatic functions. The effector protein AnkX of *Legionella Pneumophila* provides a striking example of a Fic protein harboring a motif divergent from the consensus and exhibiting a related function. Its Fic motif (HxFxDANGRxxV) allows CDP-choline substrate binding and enables the Fic protein to dysregulate Rab1 GTPase by phosphocholination[18].


```
H.sapiens|Q9BVA6             AMNLHPVEFAALAHYKLVYIHPFIDGNGRTSR.LLMNLILMQAGYPPITIRKE
M.musculus|Q8BIX9            AMNLHPVEFAALAHYKLVYIHPFIDGNGRTSR.LLMNLILMQAGYPPITIRKE
C.elegans|Q23544            TLTIDPIERAAIAHYKLVLVHPFTDGNGRTAR.LLLNLIMMRSGFPPVILPVE
M.tuberculosis|A5WTK8        VGEGDLAGQVAYRYDYVNYAHPFREGNGRSTR.EFFDLLLSERGSGLDWGKTD
V.parahaemolyticus|Q87P32    ENPVADKDLGKHLFAGVIGYHGFTDGNGRMGRMLYAIAELRNDSFNPLAMNA.
N.meningitidis|E0N845        SGDLDPLIIMAAAHYQFEAIHPFTDGNGRTGR.ILNSLLLIEKGLLDLPILYL
H.somni_Fic1|Q06277          KQNVEPSVLAGLVYQRLIAYHPFAEGNGRMAR.VVVNKILLDAGYPPFTKFSS
M.maripaludis|A6VJ92         DKFIHPIIKGIIHFLIGYIHPFNDGNGRTARSLFYWYLLKNDYWLFEYMAI.
H.volcanii|D4GR67            GPRYAPLVDLALIHYQFETIHPFQDGNGRLGR.LLVMLALYKWELLPGPYLYP
WSSV|A2TEX9                  NDQIKIIKACAYFMYNFLTLHPFNDGNGRTAR.LLYSFLLKGNGIVPHFSPIT
Lactococcus_phage|Q9AZ49     RDDIPTYEKAFISHFYFENTHPFYDGNGRTGRFILCSYLARKLDYLSAVGVS.
```

Figure 14: Sequence alignment of selected proteins that define the diversity of these Fic proteins. Sequences are accompanied with their respective uniprot number and are colored in the following: magenta, eukaryotes; black, bacteria; cyan, archae; blue, viruses. WSSV stands for White spot syndrome virus. The consensus HxFx[D/E]GNGRxxR is highlighted with red triangles.

1.3.2   FIC domain: an α-helical fold topology

The FIC domain-containing Fic proteins are classified together with a second family of proteins, Doc (death on curing), in the Pfam protein database[86] and form the Fic/Doc protein family (Pfam: PF02661) or later named the fido superfamily[89]. Among the proteins present in the database, few have known three-dimension architectures (Fig. 15). Four structures solved by structural genomics efforts (PDB codes 2G03, 2F6S, 3EQX, and 3CUC) present the highly conserved FIC active site motif (HxFx[D/E]GNGRxxR) and a common α-helical fold topology composed of eight helices[87,88,90] (Fig 15 and 16). The α-helical fold consists of a four-helix bundle (α2-α5), representing the FIC core domain as defined by Pfam[86] that is decorated by one preceding helix (α1) and two additional helices (α6 and α7) stacking almost perpendicular to the bundle. The active site motif is located in the loop between helix α4 and α5. An additional helix (α') is visible in the four structures and adopts the same position with respect to the FIC active center. Interestingly, α' helix is found permuted and can be either found N-terminally (3EQX and 3CUC) or C-terminally (2G03 and 2F6S) to the FIC active site and completes the Fic fold (Fig. 15). Finally a β-hairpin loop structure located between helices α2 and α3 seems to be a common feature shared by Fic proteins.
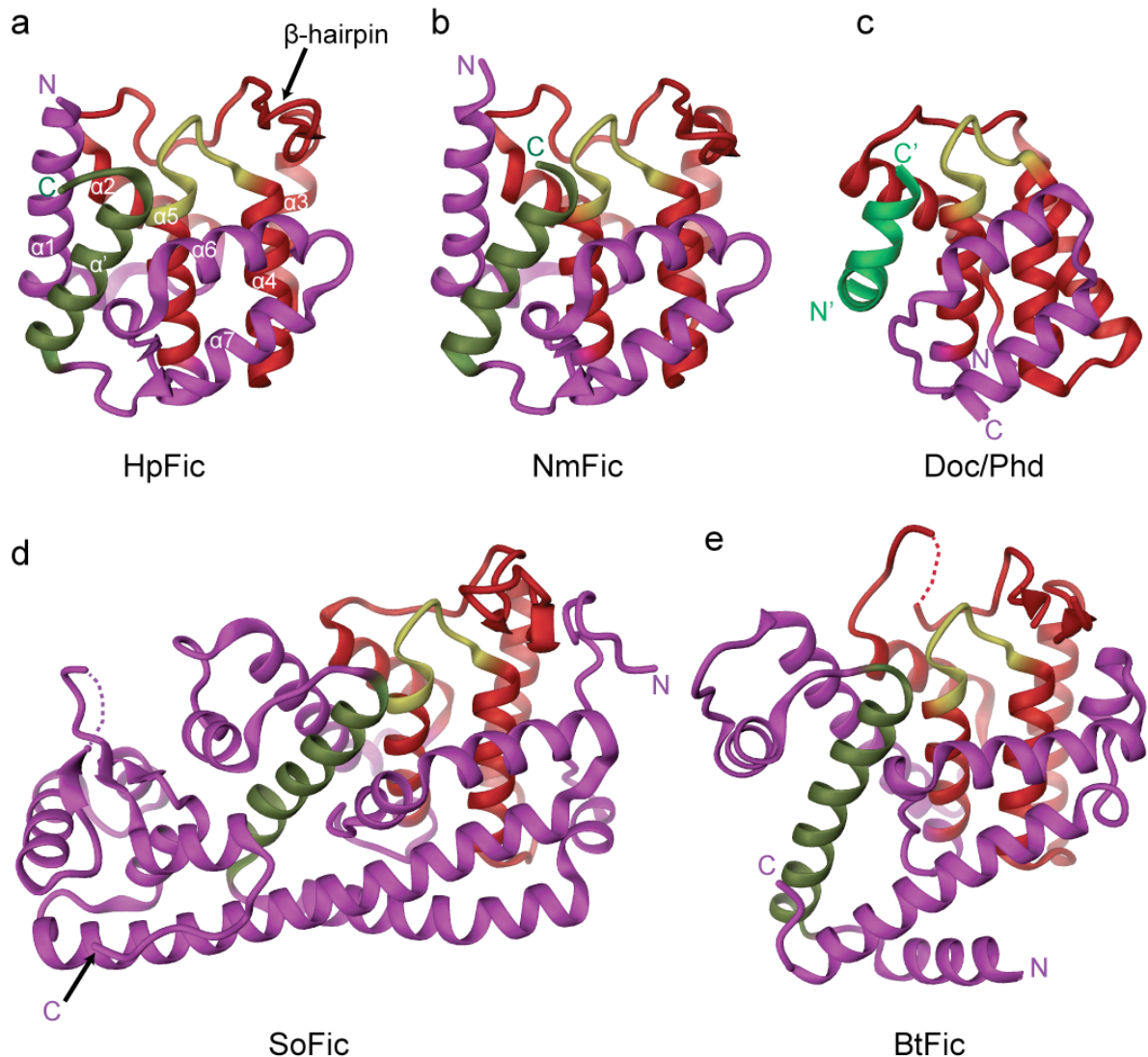
Figure 15: Overall topology of Fic and Doc proteins with known crystal structures. **a**) Fic *Helicobacter pylori* (PDB code 2F6S), **b**) Fic *Neisseria meningitidis* (PDB code 2G03), **c**) Doc of prophage P1 in complex with Phd (in green) (PDB code 3DD7), **d**) Fic *Shewanella oneidensis* (PDB code 3EQX), **e**) Fic *Bacteriodes thetaiotaomicron*. Color code is the following: red, core of Fic domain as defined by Pfam[86]; yellow, HxFxxGNGRxxR signature motif; green, permuted helix α'; magenta, remainder of structure. The same color code illustrates all Fic proteins throughout the manuscript. All Fic proteins harbor a β-hairpin structure implicated in target recognition (see chapter 1.3.4) that is absent in Doc. Helices α1 and α' are replaced by Phd in Doc/Phd structure.
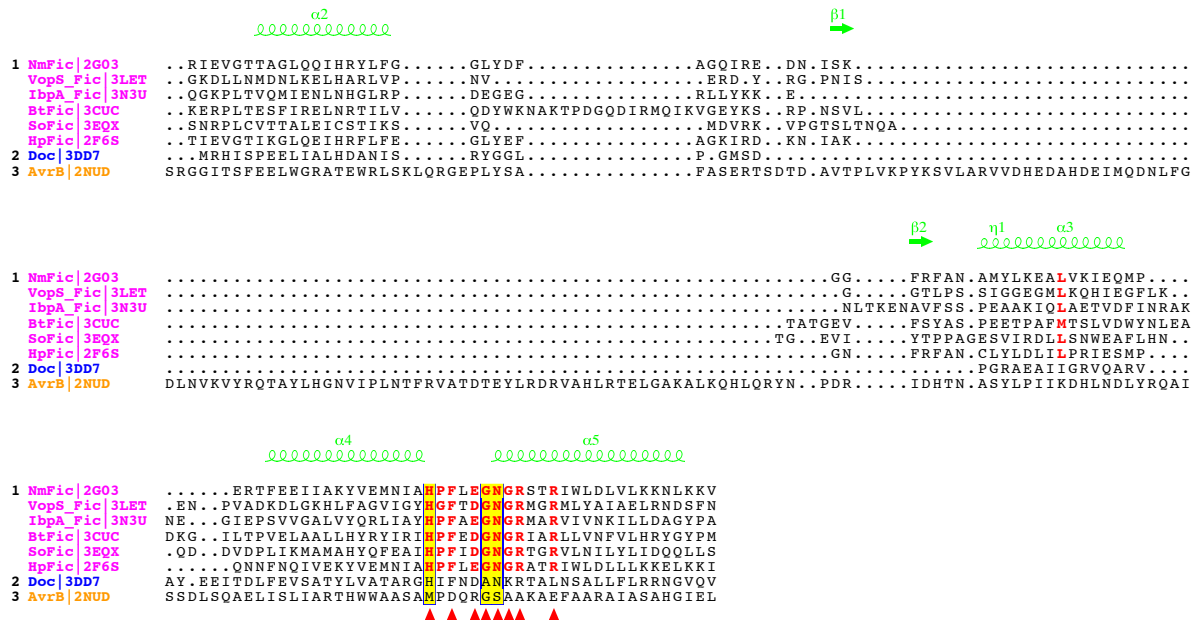
Figure 16: Structure-based multiple sequence alignment of the fido structures represented in Figures 15 and 17. The sequences were aligned using Expresso[91] and rendered using ESPript[92]. Fic-domain containing proteins, Doc and AvrB sequences are colored (magenta, blue, and orange, respectively) and labeled with PDB ID. The proteins were classified in 3 groups for similarity calculations using the BLOSUM62 Matrix[93] (group 1: Fic proteins, group2: Doc, group3: AvrB). In red characters are displayed highly similar residues within a group and yellow boxes reflect high global similarity between the different groups. Only the sequence part corresponding to the conserved FIC core domain (α2-α5 and β-hairpin) is shown. Red triangles highlight the residues defining the Fic active site (HxFx[D/E]GNGRxxR). The highly conserved Fic motif is degenerated in Doc and quasi absent in AvrB. AvrB displays a large extension between β1 and β2 strands that contacts the RIN4 target.

1.3.3   Fic is evolutionary related to Doc and AvrB

As mentioned above, Fic proteins are grouped with Doc related proteins in the same superfamily. In the *E. coli* phage P1 the Doc protein and the small PhD protein form a toxin-antitoxin module that serves as a postsegregational killing system (PSK) to insure stable inheritance of plasmid in the bacterial population. In the absence of the PhD antitoxin, Doc induces cell growth arrest by inhibition of translation via interaction with the ribosome[94,95]. The structure of Doc in complex with a C-terminal fragment of Phd (PDB code 3DD7) reveals the same overall α-helical topology (α2-α7) but with helices α1 and α' missing (Fig. 15). Interestingly, in the complex structure, Phd antitoxin is found at the place of the two missing α-helices, and was suggested to complete the Fic core through fold

30

complementation[95]. Comparison of Doc with the Fic proteins reveals a divergent active site motif (HxFxDANKRxxL) (Fig. 16). The histidine that has a catalytic role in the Fic proteins is also essential for Doc toxicity[95], pointing towards a conserved catalytic site in both families. However, besides the fact that, no adenylylation activity has been associated with Doc, the presence of a degenerated active site motif might indicate that Doc family proteins have adopted new enzymatic function. Given the similar topology and functional active site location, Fic and Doc are likely to have a common origin and have developed distinct functional properties later during evolution.

In addition to Doc, structural comparison of Fic proteins with existing structures in the PDB data bank reveals similarities with the T3SS effector protein avirulence protein B (AvrB) from *Pseudomonas syringae*. *P. syringae* is a gram-negative bacterial pathogen of plants that uses a T3SS to deliver proteins into host cell to hijack cellular functions.

Functional studies in the model host plant *Arabidopsis thaliana* have revealed a defense response triggered by AvrB. Upon translocation of AvrB the plant cells organizes a defense mechanism that relies on the recognition of the pathogen factor by the plant resistance gene product RIN4 (RPM1 interacting protein). This detection further leads to RPM1 activation resulting in effector-triggered immunity (ETI)[96,97]. The delivery of AvrB induces phosphorylation of RIN4 at threonine 166 that is a prerequisite for the RPM1-mediated resistance[97-100]. The mechanism of AvrB-dependent phosphorylation of RIN4 is unclear but two structures of AvrB in complex with either a RIN4 peptide or ADP suggest a kinase activity for AvrB[98]. Indeed, ADP, the end-product of the phosphorylation reaction, as well as the phosphorylated threonine of RIN4 are oriented towards the putative AvrB active site and face each other (Fig. 17A).

However, no kinase activity for AvrB has been detected *in vitro* and the ADP moiety rather seems to be a substrate for the activation of AvrB via phosphorylation by the *Arabidopsis* kinases RIPK (RPM1-induced protein kinase) and MPK4 (mitogen activated protein kinase 4)[98-100]. Moreover, RIPK and MPK4 can directly phosphorylate RIN4 in an AvrB-dependent manner in-vivo arguing for an indirect phosphorylation process triggered by AvrB[99,100]. Thus, more structural and functional data are needed to elucidate the exact function of AvrB in the RIN4-dependent RPM1 activation.
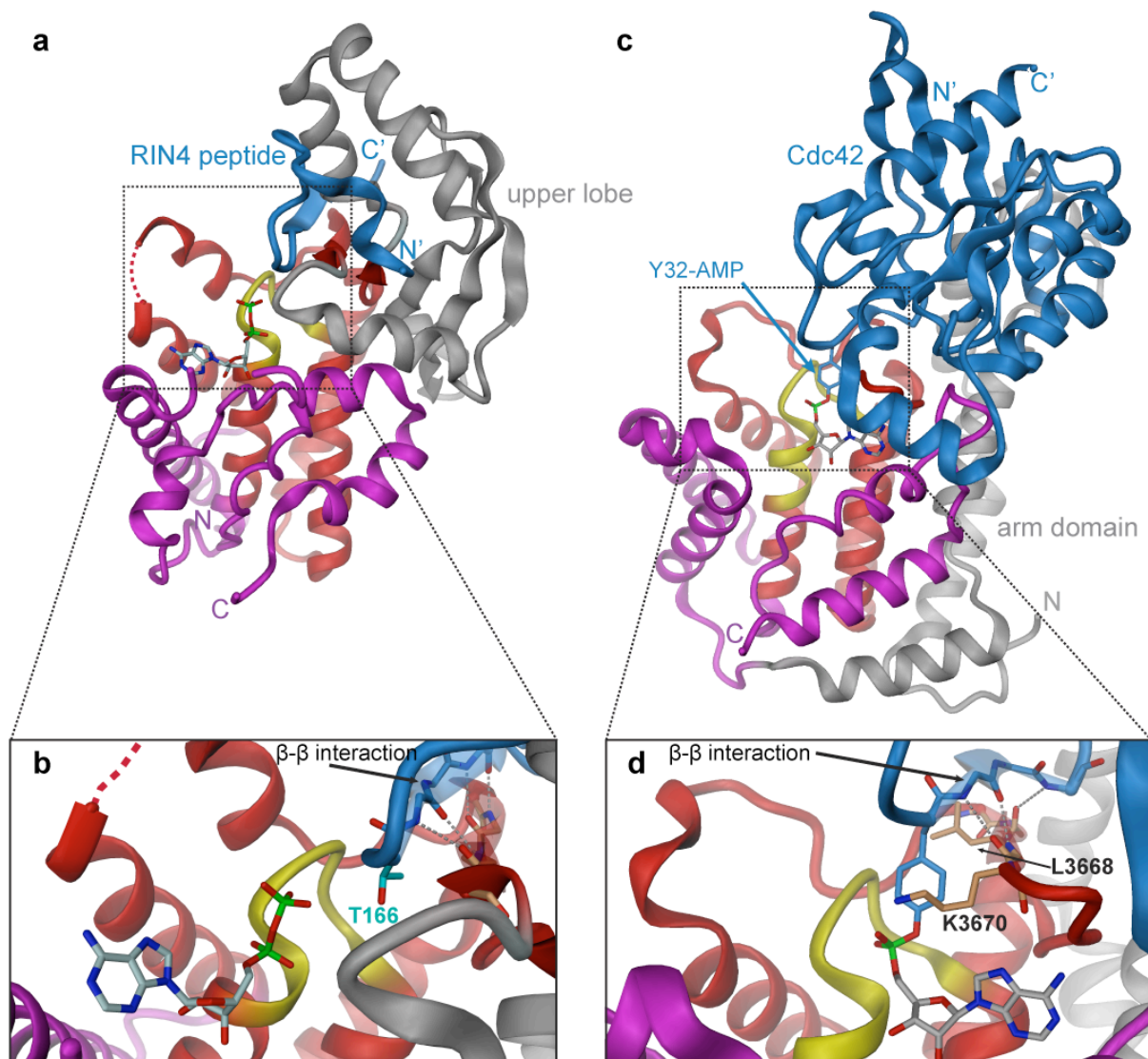
Figure 17: Structural elements important for target recognition illustrated by AvrB and IbpA. **a**) Structure of AvrB in complex with ADP and a RIN4 peptide obtained after superposition of structures of AvrB/ADP (PDB code 2NUN) and AvrB/RIN4 peptide (PDB code 2NUD). The upper lobe involved in RIN4 recognition and RIN4 are displayed in grey and blue, respectively. **b**) Close-up view of the AvrB active-site. The sequence-independent main chain-main chain hydrogen bonds are marked by stippled lines in grey. Side-chains have been omitted for clarity reason. The RIN4 Threonine 166, which is phosphorylated upon *P. syringae* infection, intrudes in the AvrB active site and is located close to ADP, supporting a kinase activity for AvrB. **c**) Structure of IbpA in complex with Cdc42 (PDB code 3N3V). Tyrosine 32 of Cdc42 is found adenylylated in the structure. The arm domain important for Cdc42 binding is shown in grey with Cdc42 in blue. **d**) Close-up view of IbpA active-site. Like AvrB, sequence-independent anti-parallel β-strand interactions are formed between Fic and target to position the target tyrosine in the Fic active site for catalysis. Additionally, side-chains of the β-hairpin residues Leu3668 and Lys3670 form a clamp that locks the tyrosine side-chain in the correct orientation for modification.

1.3.4   Target recognition: similarities and specificities

A closer inspection of the various Fic protein structures reveals a common β-hairpin element located between helices α2 and α3 (Fig. 15 and 17). In AvrB, this segment forms anti-parallel beta-strand hydrogen-bonding interactions with the RIN4 peptide and mutation of distinct β-hairpin residues impaired the ability of AvrB to promote immune response in plant[98] (Fig. 17a). The importance of this element for target recruitment has been confirmed by the structure of IbpA-Fic2 in complex with its target protein Cdc42 where a similar β-β interaction unit is visible[88] (Fig. 17b). Furthermore, the same type of contact is found in the crystal of the *Shewanella oneidensis* Fic protein (PDB code 3EQX) between two neighboring molecules[90], though, this interaction is of artificial origin. Notably, this β-hairpin is present in all Fic structures solved so far, but absent in Doc (Fig. 15), which might further confirm that Doc harbors another enzymatic function.

Target recognition via sequence-independent main chain-main chain hydrogen bonds implies the presence of additional structural elements to guide substrate specificity in Fic proteins. Consistently, the complex structure of IbpA-Fic2 with Cdc42 reveals a large interaction interface between the switch 2 region of the small GTPase and the so-called "arm domain" of IbpA[88] (Fig. 17b). An equivalent peripheral region is found in VopS Fic protein that is likely to confer the same specificity towards the small Rho GTPases[87,88]. Similarly, in AvrB, the Fic core β-hairpin structure is extended into a five-stranded antiparallel beta sheet connected to three α-helices. This domain, referred to as upper lobe, provides a large surface for RIN4 binding[98,101] (Fig. 17a). Apart from the Fic proteins of *Shewanella oneidensis and Bacteriodes thetaiotaomicron* that have large helical extensions at the N- and C-termini to possibly recognize targets, these additional satellite domains are not present in the remaining Fic proteins with known crystal structures. However, beside the four-helix bundle Fic core domain that, for catalytic reason, forms a rigid entity, the peripheral segments that regroup the helices α1, α6-α7 and α' are found in diverse dispositions in the different Fic structures (Fig. 15). These components, together with further loop extensions are probably relevant for target specificity.

1.3.5 Catalytic mechanism

Considering its analogy with phosphorylation, two scenarios were proposed for the Fic-mediated adenylylation in terms of mechanism, both relying on the central active site histidine[87-89]. The first is a ping-pong mechanism involving a high-energy adenylylated enzyme intermediate. In this scenario, the histidine is first modified before subsequent transfer of the AMP moiety onto the target protein. The second implies the formation of a ternary complex between the Fic protein and its two substrates. In this case, the histidine will serve as a general base to abstract a proton from the hydroxyl side chain of the target residue. The activated hydroxyl group can then performed a nucleophilic attack on the α-phosphate of ATP (Fig. 18).

Support for a histidine modification, which would implicate a ping-pong process, arised from the observation that Fic proteins can be auto-adenylylated[88,89]. The biological meaning of this phenomenon is unclear and it remains to be elucidated whether it has any functional relevance during adenylylation.

However, initial velocity experiments achieved on VopS strongly reject the ping-pong mechanism and argue for a ternary complex formation[87]. Additionally, steady-state kinetic experiments showed that mutation of the histidine does not entirely eliminate the adenylylation activity but significantly reduce it by nine orders of magnitude[87]. This residual activity of the point mutant was also confirmed by *in vitro* adenylylation assays[19]. Finally, Cdc42 was found adenylylated in the co-crystal obtained with the active site histidine to alanine mutant of IbpA-Fic2[88].
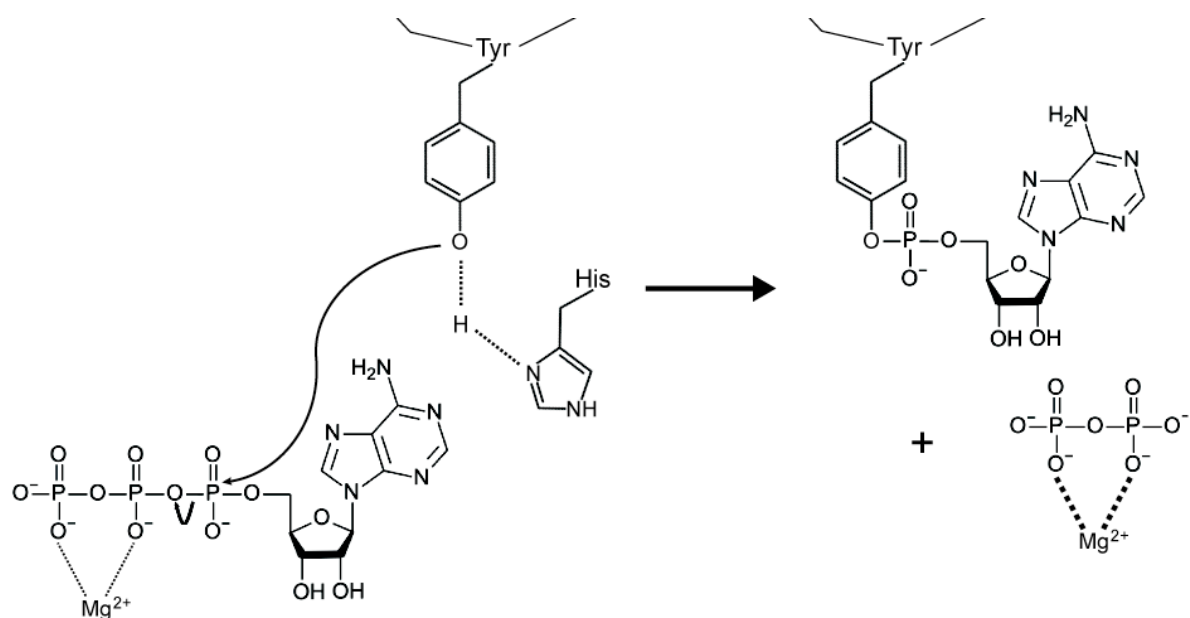
Figure 18: Mechanism of Fic-mediated adenylylation. A ternary complex is formed between Fic, ATP and the target protein (Tyr). The Fic histidine is ideally positioned to abstract a proton from the target hydroxyl group. The activated oxygen acts as a nucleophile, attacking the α-phosphorous of ATP and expelling pyrophosphate.

1.4 *Bartonella*: a suitable model for structure/function analysis of the FIC domain

1.4.1 General introduction

Bartonella species are gram-negative facultative intracellular pathogens that belong to the α2 subgroup of the proteobacteria and are closely related to the genera *Brucella* and *Agrobacterium*. Up to now, more than twenty species have been identified and ten have their genomes sequenced.

Bartonella species appear to have a specific mammalian species as a host, with *B. bacilliformis* and *B. quintana* being adapted to human. *B. bacilliformis* is highly virulent in its human reservoir host where it causes Carrion's disease which is characterized by severe hemolytic anemia during the acute phase of the infection[102]. Except for the acute infection phase of *B. bacilliformis*, bartonellae are typically considered as stealth pathogens in the sense that they circumvent host immune responses to establish long-term persistence. In most cases, infection of the reservoir hosts only results in mild symptoms[103].

The hallmark of such a reservoir host infection is a chronic intraerythrocytic bacteremia, which appears to be a specific adaptation to the mode of transmission by blood-sucking arthropods. Blood-sucking arthropods vectors, such as fleas, body lice and ticks transmit the bacteria from its reservoir to new incidental hosts (e.g. from zoonotic pathogens to humans) where it usually causes only benign symptoms. However, upon infection of immunocompromised patients, *Bartonella* can lead to more severe symptoms including tumor angiogenesis or meningitis and is therefore considered as an opportunistic pathogen.

1.4.2 Type IV secretion system and *Bartonella* evolution

A phylogenetic analysis based on four housekeeping genes and core genes of the available *Bartonellae* genomes allowed the classification of the species into four different lineages[104] (Fig. 19). This analysis placed *B. bacilliformis* as the sole representative of an ancestral lineage (lineage 1) from which most of the other species have evolved by radial speciation. The diversification of these species from a B. bacilliformis-like ancestor has likely arisen after horizontal acquisition of three different type IV secretion systems (T4SSs) that

facilitated adaptation to distinct mammalian reservoirs[104,105] (Fig. 19). Species of lineage 2, which exclusively are adapted to ruminant hosts, only harbor the so-called VirB-homologous (Vbh) T4SS. In contrast, lineages 3 and 4 harbor the translocator VirB/D4 T4SS which presumably has been acquired via HGT in an common ancestor. the most species rich lineage, lineage 4, harbors a third T4SS, the Trw (Fig. 19) involved in the attachement to erythrocytes.

The VirB-like (VirB or Vbh) T4SS is conserved in all *Bartonella* species of lineage 4 and presumably enable the bacteria to fine-tune the interactions with their respective mammalian hosts via translocation of a cocktail of different effector proteins. The effector proteins can be translocated into endothelial cells where they subvert cellular functions - a process likely essential for establishing a chronic infection[106].

The Trw system, that is only present in a sub-branch of the modern lineage, does not translocate any substrates into host cells and functions as an adherence factor mediating the invasion of erythrocytes[106].

In the following paragraphs, I will focus on the VirB-like T4SS and the current knowledge about the roles of the effector proteins in host subversion.
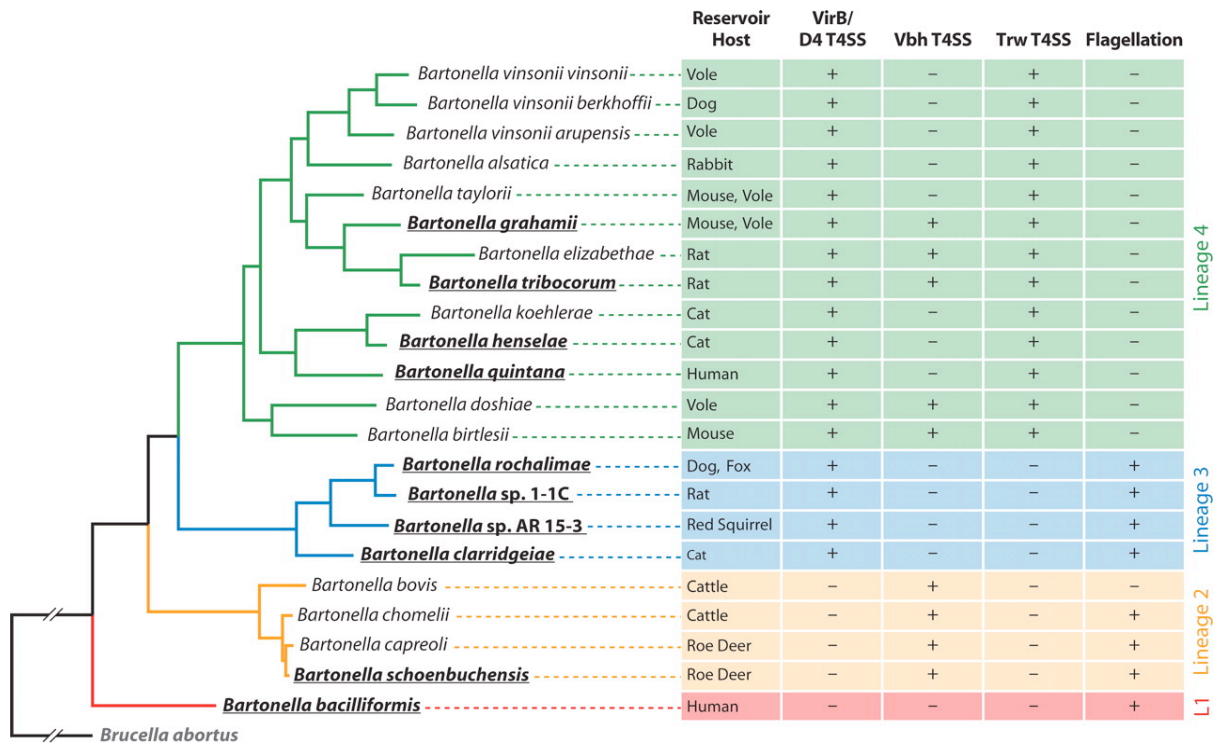
Figure 19: Phylogeny of *Bartonella*. The phylogenetic tree is based on a maximum-likelihood analysis using an alignment of 478 genes from the core genome of the ten sequenced *Bartonella* species (bold and underlined) and *Brucella abortus*. Additional *Bartonella* species have been added using the sequences of four housekeeping genes. The primary mammalian host as well as key virulence factors are indicated for each species. The deadly human pathogen *B. bacilliformis* forms a solely deep-branching ancestral lineage. All modern species harbor a type IV secretion system as a key innovation allowing the exploration of new niches. The Vbh and Trw T4SSs are characteristics of lineage 2 and 4 respectively. (Taken from Harms & Dehio[107]).

1.4.3 Versatility of the type IV secretion system

The bacterial T4SS is a multicomponent system, which protrudes the cell envelope and is used by many bacteria to secrete or translocate effectors into adjacent cells. It is the most versatile family of secretion systems, which serves for the delivery of macromolecules into various target cells, e.g., the transfer of DNA into bacteria or the transfer of effector proteins into eukaryotic target cells[108].

The exchange of genetic material through horizontal gene transfer plays an important role in genome plasticity and enables the bacteria to cope with environmental changes, as encountered during infection of the human host. It is also the basis for the spread of antibiotic resistance as well as virulence traits[109,110].

Many bacterial pathogens such as *Bordetella pertussis*, *Legionella pneumophila* or *Bartonella spp*. rely on T4SSs for their pathogenicity. They utilize this secretion machinery to translocate protein substrates into targeted host cells, thereby modulating a wide range of host cell functions[111-114].

1.4.4 Typical type IV secretion architecure

Although variations exist, many of the T4SS found in Gram-negative bacteria are similar to the well-characterized archetypical VirB/D4 system of *A. tumefaciens*, which comprises twelve proteins, named VirB1 to VirB11 and VirD4. A combination of functional and structural analysis on various T4S systems has helped to elucidate the overall architecture and details of the secretion mechanism. 14 copies of each VirB7, VirB9 and VirB10 proteins form a tight core complex anchored in the two membranes of the Gram-negative bacteria (Fig. 20). This core structure creates a channel that spans the entire bacterial cell envelope with a 55Å opening at the base of the core complex and a small 10Å opening to the extracellular milieu[115]. This narrow outer membrane pore suggests that large structural movements have to take place to allow substrate transfer. Unexpectedly, VirB10 inserts in both bacterial membranes and serves as a structural scaffold that bridges inner and outer membrane leaflets[115-117]. Thus, VirB10 may sense conformational changes upon substrate binding in the cytosol and propagate the signal to regulate outer membrane channel opening and closing. The VirB6 and VirB8 are integral membrane subunits that directly contact the substrate during secretion in *A. tumefaciens* and are likely to form, together with VirB10, the inner membrane channel[118]. VirB4, VirB11 and VirD4 harbor ATPase activity that energize substrate secretion and possibly assist the assembly of the machinery[119-121]. VirD4, also named coupling protein (CP), recruits substrates and guide them to specific subunit of the channel for secretion[108]. The extracellular pilus is built up by VirB2 and VirB5 subunits where VirB2 is the major component that is thought to form a channel between the donor and host cells for substrate transfer and VirB5 might function as an adhesin[122,123].

The roles of VirB1 and VirB3 in this complex are still unclear. But VirB1 seems to be essential for T-pilus biogenesis in *A. tumefaciens* and VirB3 is believed to function synergistically with VirB4[124,125].
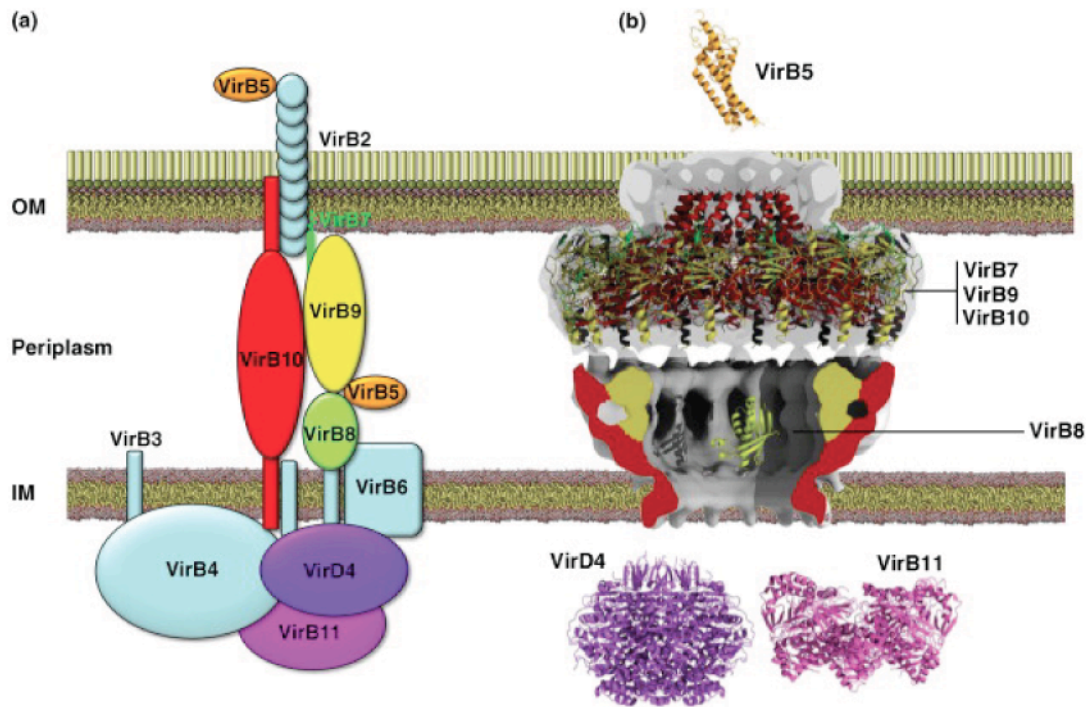
Figure 20: Schematic model and structure of the VirB/D4 type IV secretion system. (**a**) The experimentally predicted (i.e. copurification, immunoprecipitation) locations of the VirB/D4 subunits. (**b**) Structures of the T4SS components solved to date. The atomic structures are shown in ribbon. The cryo-electron microscopy picture of the T4SS core complex comprising VirB7-9-10 subunits is displayed as a cut-out volume (in grey). The crystal structure of the outer membrane complex is composed of the full-length VirB7 (green), the C-terminal domain of VirB9 (yellow), and the C-terminal domain of VirB10 (red). (Taken from Waksman & Fronzes[126]).

1.4.5 *Bartonella*-specific VirB/D4 type IV secretion systems and associated effectors

The horizontal acquisition of the VirB type IV secretion system (T4SS) was shown to play a central role for the host adaptation of *Bartonella* species[104]. This machinery is present in lineages 3 and 4 and is dedicated to the translocation of an arsenal of effector proteins to manipulate host cell signaling pathways during infection[111,112,127]. *Bartonella* effector proteins (Beps) display a modular architecture consisting of a variable N-terminal moiety and a conserved translocation signal at the C-terminus that is formed by the BID (Bep intracelluar delivery) domain and a positively charged C-tail. Both components constitute a bipartite signal recognized by the VirD4 coupling protein for efficient VirB-dependent protein transfer[112]. The N-terminal domain is highly diverse among Beps. It can be composed of,

additional BID domains that have acquired functions within the host, tandem-repeated tyrosine-phosphorylation motifs, or FIC domains.

A phylogenetic analysis inferred from the different sequences of the *bep* genes allowed classification of orthologous Beps from lineage 3 and lineage 4 in different clusters (clades 1 to 10 for lineage 3 and clades A to I for lineage 4) (Fig. 21). These clades evolved independently by duplications, diversification and reshuffling from a single ancestral Bep with a FIC-BID domain topology[105].

Figure 21: Domain architecture of *Bartonella* effector proteins. Domain architecture for (**a**) lineage 3 and (**b**) lineage 4. FIC and BID domains are displayed in orange and violet, respectively. The FIC motif (HxFx[D/E]GNGRxxR) is depicted. Motifs differing from the consensus sequence are indicated with the conserved and divergent residues in red and black, respectively. The far majority of FIC domains harbor degenerated FIC active site motifs, suggesting diverse functionalities. (Adapted from Engel et *al.*[105]).

Among the various *Bartonella spp.*, *B. henselae* is the best characterized member and the role of the VirB/D4 T4SS and the translocated effectors in *B. henselae* pathogenesis has been highly investigated in our laboratory. *B. henselae* causes an asymptomatic bacteremia in its feline reservoir host and can be incidentally transmitted to human by a bite / scratch of a cat or by indirect bite of cat fleas leading to Cat scratch Disease (CSD). In immunocompetent patients, the disease manifests by a benign lymphadenopathy in the proximal region of the bite / scratch and resolves spontaneously within 2-4 months[128]. In contrast, infection of immunocompromised patients leads to angioproliferative disorders such as bacillary angiomatosis or bacillary peliosis hepatitis that is characterized by vasoproliferative tumors through direct or indirect effects on endothelial cells[129,130].

Detailed *in vitro* studies using Primary Human Umbilical Vein Endothelial Cells (HUVEC) as an *in vitro* infection model[131,132] demonstrate that the VirB/D4 T4SS is an essential virulence determinant for vascular endothelial cell subversion[126,127]. *B. henselae* harbors a set of seven effectors, named BepA to BepG, which are encoded together with the VirB/D4 genes in a pathogenicity island (PAI)[112]. While BepA, B and C adopt the original FIC-BID domain architecture, the others display additional N-terminal BID domain (BepG), tyrosine-phosphorylation motifs (BepD) or a composite of both (BepE and BepF) (Fig. 21b).

A wide array of VirB/D4 effector-dependent phenotypes have been found such as, massive cytoskeletal rearrangements[133-135], inhibition of cell apoptosis[111,136] and induction of proinflammatory phenotype by activation of NF-κB[111]. The peculiar role of the effectors in these processes has been scrutinized. BepA is targeted to the endothelial cell plasma membrane where it inhibits apoptosis via a rise in the cytosolic concentration of the second messenger cAMP[136]. Furthermore, it promotes sprout formation in an *in vitro* angiogenesis assay[137]. Interestingly, the BID domain is sufficient to trigger these cellular phenotypes additional to its secretion signal function. No phenotypes have so far been associated with BepB and BepE. The combined action of BepC and BepF can mediate F-actin-dependent bacterial uptake through invasome formation, similarly to BepG[133-135]. BepD possesses a weak pro-angiogenic activity and becomes phosphorylated after translocation into host cells[137,138], but the cellular effects of tyrosine-phosphorylation sites remain to be defined. Remarkably, almost all phenotypes elicited by *B. henselae* effector proteins are solely dependent on the BID domains and no contribution of any FIC modules has been conclusively demonstrated *in vivo* and *in vitro*.

1.4.6 The Vbh T4SS: evolutionary link between conjugation machinery and host-interacting T4SS

The VirB homologous (Vbh) T4SS is characteristic of lineage 2 that strictly comprises ruminant-infecting species (Fig. 19). Comparative analyses with the VirB T4SS revealed that homologous genes of all VirB T4SS subunits are conserved in this ruminant-adapted species including the coupling protein-encoding gene. However, a remarkable divergence in gene content of the downstream located effector region is evident. Indeed, the vbh locus displays all genetic features (origin of transfer (*oriT*), and *traC* and *traD* involved in the processing of the *oriT* during conjugation) that make Vbh T4SS a functional conjugation system. It only contains one single *bep* gene, named *vbhT*, with the FIC-BID domain architecture as found for many effector proteins of the VirB T4SS (Fig. 21). Thus, Vbh T4SS appears to display an evolutionary link between conjugation machineries and T4SSs adopted for host interaction[139].

1.4.7 *Bartonella*: a tremendous playground to explore FIC domain structure/function

Twelve out of the seventeen Bep clades of the *Bartonella* lineages 3 and 4 contain FIC domains. In total, they represent 55 proteins that harbor FIC domains with disparate active site motifs (Fig. 21). Owing to the omnipresence of FIC, *Bartonella* is an ideal model not only to examine the biological importance of FIC-mediated adenylylation during bacterial pathogenesis but also to explore new functions adopted by FIC domain with degenerated active site motif during evolution.

2 Aim of the Thesis

The FIC domain is widely spread in all kingdoms of life and regulates fundamental pathways in bacteria to higher organisms by adenylylation of target proteins. Only recently, the domain was functionally characterized, yet little is known about the catalytic mechanism. A major breakthrough to understanding FIC-mediated AMP transfer arose from the crystal structure of IbpA-Fic2 in complex with the adenylylated form of the small Rho-GTPase Cdc42[88]. This snapshot of a Fic protein with an end product of the enzymatic reaction provided structural insight into the determinant of target recognition. However, to decipher the FIC-catalyzed adenylylation in its entirety, details of ATP binding mode in the FIC active site is indispensible.

Apart from catalysis, the regulation of Fic-mediated adenylylation is the second prominent question of biological relevance that needs to be addressed. The function of the Fic domains have been discovered in the context of bacterial effector proteins that hijack cellular processes of the host upon infection. However, the vast majority of Fic proteins found in the genomes of various prokaryotes do not have any apparent link to host interaction and the potentially lethal activity carried by these proteins should be tightly regulated.

During my PhD, I tried to unravel the structural elements and mechanisms governing catalysis and regulation of the Fic AMP transferase family. To this end I applied X-ray crystallography and biochemical techniques and, focussed mainly on FIC domains from *Bartonella* species with occasionally taking resort to Fic proteins from *Shewanella oneidensis* and *Neisseria meningitidis*.

3 Results

3.1   Research article I

Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA

Dinesh V. Palanivelu, Arnaud Goepfert, Marcel Meury, Patrick Guye, Christoph Dehio and Tilman Schirmer

Statement of the own participation

I contributed to this publication by performing *in vitro* adenylylation assays with the different constructs of BepA from *B.henselae*. Constructs were cloned by Patrick Guye. Structural investigations were carried out by Dinesh Palanivelu and Marcel Meury. The manuscript was written by Christoph Dehio and Tilman Schirmer.

Statement of the own participation

Fic domain-catalyzed adenylylation: Insight provided by the structural analysis of the type IV secretion system effector BepA

Dinesh V. Palanivelu,[1] Arnaud Goepfert,[1,2] Marcel Meury,[1] Patrick Guye,[2] Christoph Dehio,[2] and Tilman Schirmer[1]*

[1]Core program of Structural Biology and Biophysics, Biozentrum, University of Basel, CH-4056 Basel, Switzerland
[2]Focal Area of Infection Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

Abstract: Numerous bacterial pathogens subvert cellular functions of eukaryotic host cells by the injection of effector proteins via dedicated secretion systems. The type IV secretion system (T4SS) effector protein BepA from _Bartonella henselae_ is composed of an N-terminal Fic domain and a C-terminal _Bartonella_ intracellular delivery domain, the latter being responsible for T4SS-mediated translocation into host cells. A proteolysis resistant fragment (residues 10–302) that includes the Fic domain shows autoadenylylation activity and adenylyl transfer onto Hela cell extract proteins as demonstrated by autoradiography on incubation with α-[^{32}P]-ATP. Its crystal structure, determined to 2.9-Å resolution by the SeMet-SAD method, exhibits the canonical Fic fold including the HPFxxGNGRxxR signature motif with several elaborations in loop regions and an additional β-rich domain at the C-terminus. On crystal soaking with ATP/Mg^{2+}, additional electron density indicated the presence of a PP$_i$/Mg^{2+} moiety, the side product of the adenylylation reaction, in the anion binding nest of the signature motif. On the basis of this information and that of the recent structure of IbpA(Fic2) in complex with the eukaryotic target protein Cdc42, we present a detailed model for the ternary complex of Fic with the two substrates, ATP/Mg^{2+} and target tyrosine. The model is consistent with an in-line nucleophilic attack of the deprotonated side-chain hydroxyl group onto the α-phosphorus of the nucleotide to accomplish AMP transfer. Furthermore, a general, sequence-independent mechanism of target positioning through antiparallel β-strand interactions between enzyme and target is suggested.

Keywords: FIC domain; AMPylation; adenylylation; AMP transfer; type IV secretion system

51

Introduction

Adenylylation, that is, AMP-transfer from an ATP substrate onto a target molecule, is a common process in biology. This enzymatic reaction occurs during nucleotide polymerization and on priming carboxylate substrates for further chemical modification as, for example, catalyzed by acyl-CoA synthetase.[1] However, proteins also can become post-translationally modified by attaching the AMP moiety onto various types of side-chains. In this way, a lysyl side-chain is transiently adenylylated in polynucleotide ligase, followed by adenylyl transfer onto the substrate.[2] Glutamine synthetase adenylyl transferase controls the activity of an important metabolic enzyme by tyrosyl modification.[3] A related protein from the pathogen *Legionella*, DrrA, adenylylates the membrane traffic regulator Rab1b on translocation into eukaryotic host cells.[4] Only recently, the Fic (filamentation induced by cAMP) domain (Pfam 24.0[5]) found in almost all kingdoms of life and annotated in currently more than 2000 sequences has been recognized to catalyze adenylyl transfer onto specific protein hydroxyl side-chains of interacting target proteins. This has been demonstrated for two bacterial effector proteins carrying a Fic domain, VopS of *Vibrio parahaemolyticus*, and IbpA of *Histophilus somni*, which on translocation into mammalian host cells inactivate small GTPases of the Rho-family by adenylylation of conserved threonine or tyrosine residues, respectively, resulting in a collapse of the cellular actin cytoskeleton.[6,7]

Fic-domain containing effector proteins are also encoded by *Bartonella henselae*.[8] This pathogen translocates via a dedicated protein secretion system, the type IV secretion system VirB, seven effector proteins known as BepA-BepG into human endothelial cells.[9] The resulting cellular changes include rearrangements of the actin cytoskeleton, the inhibition of programmed cell death (apoptosis), the activation of a proangiogenic process, and the initiation of an inflammatory response.[10] All seven Beps show one or several *Bartonella* intracellular delivery (BID) domains in the C-terminal part that are responsible for protein delivery into host cells,[9] whereas the N-terminal part of the homologous BepA, BepB, and BepC proteins comprise a Fic domain (see also Fig. 1, top). The BID domain of BepA mediates inhibition of apoptosis,[11] whereas the precise function of the N-terminal Fic domain is yet unknown. Here, we present the crystal structure of the BepA Fic domain in complex with pyrophosphate, the side product of the adenylylation reaction. Combined with other published structural information,[12] the structure of a generic ternary substrate complex is inferred and a detailed catalytic mechanism of the adenylyl transfer onto the target hydroxyl group is proposed.
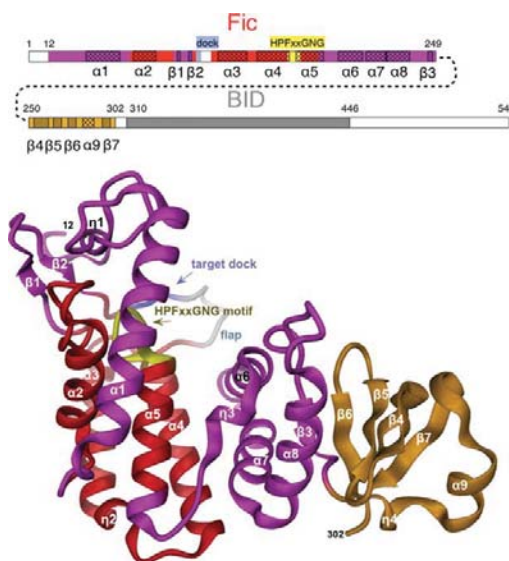
Figure 1. Crystal structure of BepA$_{tr}$ (residues 12–302) and scheme of the BepA sequence (top). The following color code is used throughout: red, core of Fic domain as defined by Pfam[5]; blue, peptide segment (103–105) in extended conformation that has been shown to interact with target protein/peptide in structural homologs IbpA,[12] AvrB,[19] and SO2466[13]; light-grey, part of the flap that is not resolved by electron density (106–111, light-grey); yellow, HPFxxGNG signature motif; gold, OB-fold (see also Supporting Information Fig. S2); magenta, remainder of structure. Secondary structure elements are numbered sequentially, α, α-helices, η, 3$_{10}$-helices, and β, β-strands.

Results and Discussion

Crystal structure of a BepA fragment comprising the Fic domain

An N-terminal His$_6$-tagged construct of full-length BepA (64 kDa) was expressed in *E. coli* and was found to be proteolytically truncated at the C-terminus to yield a fragment (BepA$_{tr}$) with a mass of about 40 kDa as evidenced by size-exclusion chromatography and anti-His Western blot. Removal of the His-tag, which proved essential for crystallization, was achieved by proteolytic treatment with thrombin. The two molecules of the asymmetric unit form a weak, probably nonphysiological, symmetric dimer (Supporting Information Fig. S1). Figure 1 shows the overall structure of BepA$_{tr}$ with residues 12–302 well defined by electron density except for residues 106–111 of a loop close to the signature motif (flap) and a few side-chains. The structure can be grouped in an N-terminal, predominantly α-helical part (12–245), and the remaining C-terminal end exhibiting the OB (oligonucleotide/oligosaccharide binding)-fold (Supporting Information Fig. S2). Residues 62–177 [helices α2–α5, see also Fig. 2(A)] are classified as Fic domain in the Pfam database.[5] There is a small
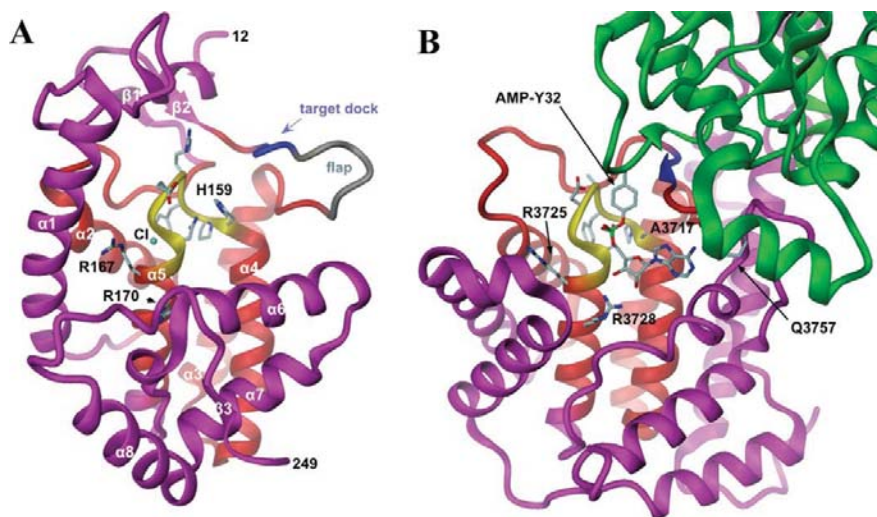
Figure 2. Structural comparison of (A) $BepA_{tr}$ (PDB code: 2vy3) with a chloride ion bound to the N-cap of helix $\alpha5$ and (B) IbpA(Fic2, H3717A) in complex with Cdc42 (green) that has its modified AMP-Y_{32} residue bound to the Fic active site (PDB code: 3n3v[12]). Same color code as in Figure 1, view rotated by 90° about the vertical direction. Important residues are labeled. The OB-domain of $BepA_{tr}$ has been removed for clarity.

β-sheet (β1/β2) inserted into the α2–α3 loop that is not part of the Fic profile. The Fic core structure is flanked N-terminally by α1 and irregular loop structure. C-terminally it is followed by a three helix bundle (α6–α8) and β3 that is an extension of the β-sheet of the OB-fold domain.

Figure 2(A) shows that the H_{159}PFxxGNG signature motif of the Fic profile locates to the α4–α5 loop and the following N-cap of helix α5. A chloride ion is bound to the N-cap in both chains of the asymmetric unit. The N_{165} side-chain forms two H-bonds with the backbone amide and carbonyl group of the phenylalanine residue [F_{161}, see also Fig. 3(A)] and may thus be critical for defining the exact geometry of the loop. The phenylalanine side-chain is tucked away in the hydrophobic core of the domain.

Structural comparisons

Several structures of Fic proteins have been determined by structural genomics initiatives (PDB codes: 2f6s, 2g03, 3cuc, and 3eqx[13]). $BepA_{tr}$ exhibit close structural resemblance with these proteins (Z-scores 11–13 as determined by Dali,[14] see also Supporting Information Figs. S3 and S4) showing the same arrangement for helices α1 to α7. Helix α8 is packed peripherally, together with β3, onto α6 and α7 and connects to the OB-domain (Fig. 1). In the Fic proteins and in BepA, an extended HPFxxGNGRxxR motif is strictly obeyed (Supporting Information Fig. S4a). More distant structural homologs are the type III effectors AvrB and VopS, and the Doc toxin (Supporting Information Figs. S3 and S4) that give moderate Z-scores of 8, 5, and 6, respectively. Interestingly, in

$BepA_{tr}$ as in Doc toxin, a helix, that runs antiparallely to α5 and α1 is missing. This leaves a groove in the structure that is taken up by the Phd antitoxin in the Doc/Phd complex.[15] Notably, there is no overall significant sequence similarity of AvrB and Doc with BepA and the sequences of the signature loop are degenerated (Supporting Information Fig. S4). Still these segments show the canonical conformation (Supporting Information Fig. S3). A thorough structural comparison of all these structural homologs of BepA has appeared recently.[16] The recently reported structure of the Fic2 domain of IbpA[12] [Fig. 2(B)] also exhibits significant ($Z = 8$) similarity with BepA and the other Fic proteins (Supporting Information Fig. S4B).

$BepA_{tr}/PP_i/Mg^{2+}$ product complex

At the time the BepAtr structure was solved, there was no clue about the molecular function of Fic domains. However, the N-cap of α5 (residues 165–167) together with the preceding residue G_{164} appeared well suited for polyanion binding. In fact, the three backbone amide nitrogens of residues G_{164}, N_{165}, and G_{166} form a well-defined so-called anion binding "nest" of type LR[17] as noted before for the Doc protein.[15] Therefore, BepAtr crystals were used to test for nucleotide binding. On soaking of crystals with ATP/Mg^{2+}, additional electron density was found near the anion binding nest of the signature motif [Fig. 3(A)]. The density was modeled satisfactorily with a PP_i/Mg^{2+}. Crystallographic statistics are given in Tables I and II. Although one of the phosphates occupies the 164–166 nest, the other interacts with backbone amide 167. Both phosphate moieties and

Crystal Structure of BepA

Figure 3. Structural details of BepA. (A) Signature loop (yellow) with bound PP_i/Mg^{2+} and corresponding $F_o–F_c$ map (with H_{159} side-chain and PP_i/Mg^{2+} omitted for phasing) contoured at 3.5σ. The structure was obtained by soaking a $BepA_{tr}$ crystal overnight 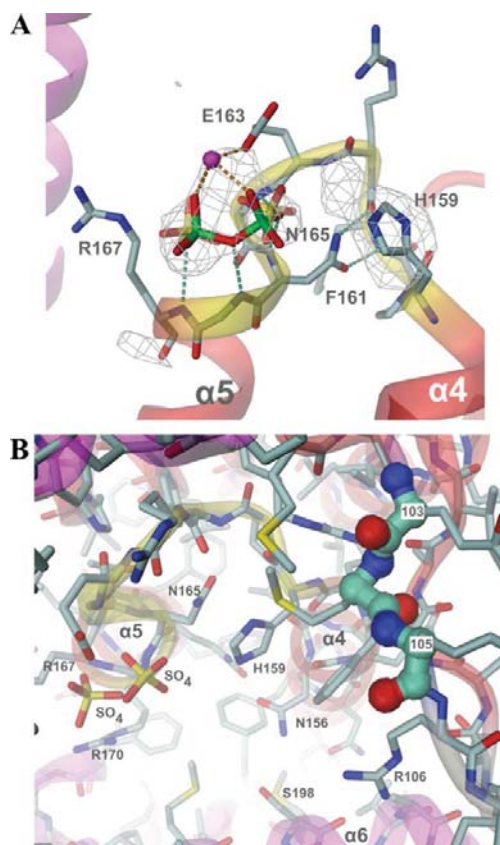in 10 mM ATP and 5 mM $MgCl_2$. Also shown, in the background, are the two sulfate anions of the $BepA_{10-303}$ structure (D-chain) bound to two mutually exclusive positions at the N-cap of helix $\alpha5$, after superposition of the two protein structures. Note the nonmodeled elongation of the electron density of the H_{159} side-chain (see main text). (B) Putative target docking site of $BepA_{10-303}$, view from the top of Figure 1 with helix $\alpha6$ running horizontally at the bottom of this figure. Full atom representation with fold cartoon superimposed. The main-chain atoms of the "target dock" in extended conformation [blue in Figs. 1 and 2(A)] of the flap ($\beta2–\alpha4$ loop) are shown in ball-and-stick representation. Note, that main-chain amide and carbonyl groups 103 and 105 are fully exposed. The large pocket next to the "target dock" in the middle of the figure accommodates the modified AMP-Tyr side-chain in IbpA(Fic 2), see Figure 2(B). Also shown are the sulfate anions, as in panel (B).

the carboxylic side-chain of E_{163} coordinate the divalent cation. The site is probably not fully occupied as indicated by elevated B-factors of the ligand.

Only when it was reported that the Fic domain of VopS exhibits adenylylation activity,[6] it became clear

that the observed moiety may represent the side-product of this reaction. In absence of a protein target, ATP may either have been hydrolyzed between the α- and β-phosphate or the AMP moiety may have been transferred to H_{159} with subsequent partial hydrolysis. Indeed, additional density at the imidazole side-chain [Fig. 3(A)] seems to corroborate later scenario.

For further studies and to avoid the problems associated with the uncontrolled processing of BepA that hampered crystal reproduction, a tailored C-terminally His-tagged construct comprising residues 10–303 of BepA ($BepA_{10-303}$) was expressed and crystallized (Supporting Information Tables S1 and S2). Unfortunately, on soaking of various ATP analogs, no additional density was found. The high salt concentration of the crystallization condition (0.5M NaCl and 1.5M $(NH_4)_2SO_4$) probably hindered nucleotide binding. Instead, as in the native $BepA_{tr}$ crystals, putative sulphate ions were found within the anion nest 164–166 and the following residue 167 [Fig. 3(A)].

BepA-catalyzed adenylylation

On incubation with α-[^{32}P]-ATP radioactivity was transferred onto the $BepA_{10-303}$ fragment [Fig. 4(A)]. This autoadenylylation was dependent on the presence of the histidine of the signature motif as evidenced by the negative result obtained with the H159A point mutant, which was properly folded as demonstrated by its crystal structure (not shown). Autoadenylation also occurred with the shortened fragment $BepA_{10-218}$ [Fig. 4(B)] that was designed to probe the activity of the Fic domain without helix $\alpha8$ and the following OB-fold. Thus, the activity resides within the canonical Fic fold. Furthermore, the activity was not base-specific. The guanyl- and cytosyl-nucleotides were incorporated with virtually equal efficiency as ATP [Fig. 4(B)]. In the presence of

Table I. *Crystallographic Data*

	$BepA_{tr}$[a]		$BepA_{10-303}$[b]
	SeMet	ATP/Mg^{++}	
Wavelength (Å)	0.97935	0.90000	1.0001
Resolution (Å)	2.9	3.0	3.20
Outer resolution shell (Å)	3.06–2.90	3.16–3.00	3.37–3.20
Space group	$P2_12_12_1$	$P2_12_12_1$	H32
a (Å)	73.3	72.9	230
b (Å)	84.4	82.9	230
c (Å)	126.7	126.7	309
R_{merge}(%)	10.3 (26.1)	9.8 (33.2)	9.3 (42.1)
Redundancy	5.6 (2.7)	3.4 (3.1)	2.8 (2.6)
I/Γ	12.8 (1.9)	6.3 (2.3)	7.3 (2.3)
Completeness (%)	92 (68)	94 (92)	96 (94)
Anomalous completeness (%)	92 (52)	–	–

Data in brackets refer to the outer resolution shell.
[a] Full-length BepA GSHMLEM$_1$–S$_{544}$ proteolytically truncated at C-terminal end during expression/purification.
[b] Designed BepA fragment MN$_{10}$–A$_{303}$HHHHHH.

Table 2. *Model refinement*

	BepA$_{tr}$[a]		BepA$_{10-303}$[b]
	SeMet	ATP soak	
Resolution (Å)	30.0–2.90	30.0–3.00	30–3.2
No. of unique reflections	16,307	14,807	46,664
R / R$_{free}$[c] (%)	21.3/27.4	21.1/27.2	23.9/27.7
Average B (Å^2)	63	66	81
r.m.s. deviation from ideal values			
Bond lengths[d] (Å)	0.011	0.012	0.009
Bond angles[d] (°)	1.1	1.3	1.2
r.m.s.d. NCS related C$_\alpha$-positions (Å)	0.4[e]	0.4[e]	0.05–0.2[f]
Molecules in asymmetric unit	2	2	8
Protein atoms	4449	4458	17898
Pyrophosphate atoms	-	9	-
Water molecules	7	7	-
Ion atoms	6	6	90
Residues in Ramachandran core (%)	98 %	97 %	96 %
Residues in disallowed region (%)	0.4 %	0.5 %	0.9 %
PDB ID	2vy3	2jk8	2vza

[b] See Table 1.

[c] R_{free} calculated with 5 % of the data that were not used for refinement.

[d] r.m.s.d. from ideal stereochemistry.

[e] For residues 12–101 and 117–302 after separate superposition.

[f] r.m.s.d. values for residues 11–101, 117–223, and 239–306 of chain A after individual superpositions on chains B to H, respectively.

HeLa cell extract, adenylyl transfer onto target proteins was evident through the appearance of two additional bands corresponding to sizes of about 40 and 50 kDa on the autoradiograph [Fig. 4(A)]. Because of the different apparent molecular weight from Rho-family GTPases representing the only known protein substrates of Fic domain-containing proteins[6,7] these proteins likely represent novel Fic domain substrates. Their identification is underway, but beyond the scope of this study.

Model of ternary substrate complex and enzymatic mechanism

Based on the two known binary Fic domain/product complex structures, that is, IbpA(Fic2)/AMP-Y$_{32}$ and BepA$_{tr}$/PP$_i$/Mg^{2+} [Figs. 2(B) and 3(A), see also Supporting Information Fig. S5], it was possible to derive a generic model for the Fic/ATP/Mg^{2+} substrate complex (Fig. 5). For this, we reasoned that the observed PP$_i$ moiety [Fig. 3(A)] most likely maps the β- and α-phosphate subsites of the ATP substrate and not the (original) γ- and β-sites. The assumption

appears reasonable considering that the nest (α-phosphate binding subsite) represents a high-affinity site, which becomes available for PPi binding on dissociation of the adenylyl moiety. Indeed, after superposition of the protein scaffolds, the PP$_i$ phosphate in the nest and the AMP α-phosphate have a distance of only 1.3 Å (cf. Supporting Information Figs. S5b and S5c). Thus, for model building, the PP$_i$ had to be simply linked to the adenosine moiety of AMP-Tyr and, on the other end, extended by an additional phosphate to yield the triphosphate moiety. Furthermore, the Mg^{2+} cation was repositioned such that it became coordinated by the γ- and α-phosphates and the E163 carboxylate. Finally, the conformation of the R$_{167}$ side-chain was altered to stabilize the γ-phosphate. Figure 5 shows the resulting model of the ternary complex.

Both conserved arginines, R$_{167}$ and R$_{170}$, are predicted to be important for ATP substrate binding. However, in five of the seven known Fic structures (not counting the distantly related AvrB and Doc)
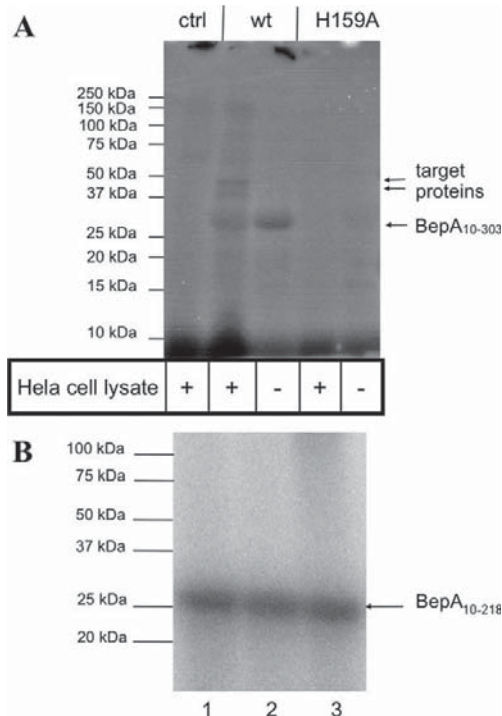
Figure 4. Autoadenylylation of BepA fragments and AMP transfer onto HeLa cell extract. BepA variants were incubated with a mixture of hot (α-[^{32}P]) and cold nucleosyl-tri-phosphate. (A) BepA$_{10-303}$ (wild-type and mutant H159A) incubated with 20-μM ATP , α-[^{32}P]-ATP (10 μCi) in the presence or absence of 150 μg of S100 HeLa cell lysate. The control (ctrl) was performed in the absence of BepA. (B) BepA$_{10-218}$ incubated with 20-μM nucleosyl-tri-phosphate and α-[^{32}P]-nucleosyl-tri-phosphate (5 μCi); lane 1: ATP, lane 2: GTP, lane 3: CTP.
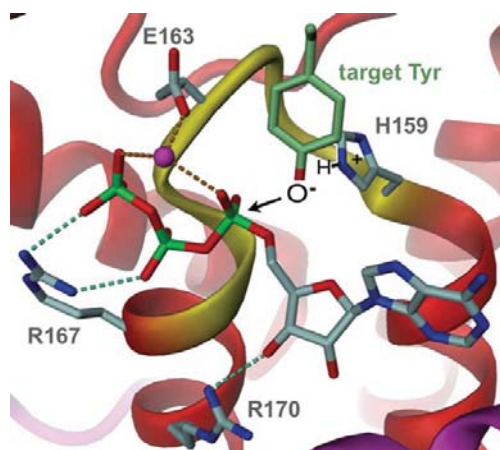
Figure 5. Model of the ternary complex of BepA(Fic) with ATP/Mg^{2+} and the modifiable tyrosine side-chain (light-green) of the target protein. The model was obtained by combining structural information of the BepAtr/PP$_i$/Mg^{2+} (2jk8) and the IbpA(Fig2)/cdc42 (3n3v) complexes, see also Supporting Information Figure S5. The structures of the two Fic domains were superimposed, the pyrophosphate moiety of the BepAtr complex extended to a triphosphate and the cation positioned in α-, γ-phosphate bridging location (guided by the structure of ATP in a myosin kinase complex, PDB code: 3lmi). Subsequently, the adenosine moiety of the adenylylated Y$_{32}$ of the IbpA complex was linked to the α-phosphate and the energy of the structure minimized with program REFMAC[27].

the arginines are engaged in salt-bridges to carboxylic side-chains of peripheral α-helices. It is tempting to speculate that on target binding, conformational changes are induced in the Fic domain that cause disruption of these salt-bridges. Such a mechanism would disfavor ATP binding and, thus, futile hydrolysis in the absence of a target. Further kinetic studies are needed to scrutinize this notion. For BepA, the nucleotide base is predicted to bind to a pocket that is structurally equivalent to that of IbpA [Fig. 2(B)]. We were not able to identify hydrophilic residues that would confer base specificity. Indeed the promiscuity observed in the functional assay [Fig. 4(B)] may indicate that there are no H-bonds formed between nucleotide base and enzyme.

The accumulated structural information allows to propose a detailed mechanism for the adenylylation reaction catalyzed by Fic domains. Thereby, as proposed before,[12,16] the histidine of the motif would act as a general base for deprotonation of the tyrosine hydroxyl group. In the model, the hydroxyl group of the tyrosine is in-line with the scissile P$_\alpha$–O$_{3\alpha}$ phosphoester bond and has distance of 2.8 Å from the phosphorus atom. Nucleophilic attack will generate the penta-coordinated transition state that is probably stabilized by the divalent cation. This would be followed by phosphoester bond cleavage to

yield AMP-tyrosine and pyrophosphate, whereby the PP$_i$ leaving group would again be stabilized by the divalent cation. Nucleotide polymerases, DNA ligases,[2] and carboxylate:adenylyl transferases[1,18] operate by the same general mechanism.

Although this mechanism should be representative for most Fic proteins with adenylyl transfer activity, the situation is different in AvrB that seems to be a kinase.[19] Although the modifiable hydroxyl side-chain of the RIN4 is in a similar position as in IbpA(Fic2)/Cdc42, the ADP nucleotide shows a binding mode different to that of ATP in our model. Although the polyphosphate is still bound to the anion binding nest, the direction of the polyphosphate is essentially reversed. As the Thr$_{166}$ side-chain of RIN4 is in a similar position than Y$_{32}$ of Cdc42, this implies that the γ-phosphate of an ATP substrate could be transferred onto Thr$_{166}$ (after deprotonation by H$_{217}$,[19] a residue not homologous to the usual signature histidine, which is absent in AvrB). Thus, two different but related activities have evolved based on the Fic fold.

In absence of a target, autoadenylylation of the Fic domain has been reported for VopS and a eukaryotic Fic domain,[16] which is also observed for BepA$_{10-303}$ and BepA$_{10-218}$ (Fig. 4). It is well conceivable that hereby, in a side-reaction, the AMP moiety is transferred onto the active histidine. The orientation of the imidazole side-chain appears suitable for an in-line attack of the N$_{\varepsilon 2}$ onto the P$_\alpha$, but the distance is comparatively large (>4.3 Å) suggesting that this would represent an inefficient side-reaction. Autoadenylylation could be the first step in a ping-pong mechanism for AMP transfer onto a target. However, consistent with our model, this has been convincingly disproven for VopS by analysis of the enzyme kinetics.[20]

Target recognition

Fic domain proteins exhibit a rather small conserved core structure. The helical α4/α5 hairpin and the intervening loop seem to be tailored for ATP/Mg^{2+} binding and catalysis of AMP transfer as discussed above. Helices α2 and α3 of the core pack against the α4/α5 hairpin by predominantly hydrophobic interactions with the phenylalanine of the motif (F$_{161}$) being part of the hydrophobic core. Although helix α1 and helices α6 to α8 are present in most Fic protein structures, they show considerable variation in location and orientation. Moreover, relative extensions/deletions at the termini or of loop structure are common and it is self-evident that these may participate in target recognition and binding as exemplified by the IbpA(Fic2)/Cdc42 complex structure.[12] Furthermore, as speculated in Ref. [12], the variability of these peripheral structural elements suggests that Fic structures have evolved to recognize (or rather hijack) distinct targets.

Nevertheless, closer inspection of the accumulated structural information about Fic proteins (Supporting Information Fig. S3) reveals a common feature that may be pertinent for target positioning. All known structures (except for Doc) reveal a β-hairpin or loop structure ["flap" in Figs. 1 and 2(A)] preceding helix α3. It is found in various orientations (Supporting Information Fig. S3; in BepA$_{tr}$ it is partly disordered), but may, on substrate (ATP and/or protein target) binding, close down onto the active site as commonly observed for kinases to prevent futile ATP hydrolysis. The N-terminal part of the flap may be involved, however, in another important function. This segment, which shows extended backbone conformation in all Fic structures [Supporting Information Fig. S3, see also Fig. 3(C)], has been observed in two cases to form intermolecular antiparallel β-strand interactions with a cognate eukaryotic target protein (IbpA(Fic2)/Cdc42[12] and AvrB/RIN4-peptide complexes[19]). Furthermore, in SO2466, this segment form an artifactual interaction of the same kind with a neighboring molecule in the crystal.[13] While this "target dock" appears well suited to recognize, in a sequence independent manner, target segments in extended conformation, its main role may be to contribute to the precise positioning of the modifiable target residue in the active site of the Fic domain. In fact, residues Y$_{32}$ of Cdc42 and T$_{166}$ of RIN4 are directly preceding the target dock (residues 33–35 and 167–169, respectively). Such coupling of sequence independent β-strand recognition and productive binding of the susceptible part of the target has also been observed in serine proteinases.[21]

Materials and Methods

Protein expression and crystallization

Full-length BepA with an N-terminal His$_6$-tag and a thrombin cleavage site was overexpressed in *E. coli* strain BL21 (DE3). Purification was performed on a Ni-NTA column (Qiagen AG, Switzerland). The His$_6$-tag was cleaved using bovine thrombin (Amersham) at 4°C for 6 h. Uncleaved material was removed by using a second Ni-NTA column. Subsequently, benzamidine sepharose beads (GE Healthcare) were used to remove thrombin from the sample. Pooled fractions were dialyzed against 20 mM TrisHCl (pH 8.0), 500 mM NaCl, and 2 mM β-mercaptoethanol (buffer A) at 4°C, concentrated and loaded onto a Superdex 75 HR 26/60 (GE Healthcare) column. The protein that turned out to be proteolytically cleaved at the C-terminal end was designated BepA$_{tr}$. The elution volume indicated monomeric protein. Selenomethionine-substituted BepA was expressed using the metabolic inhibition pathway[22] and purified the same way as native BepA. Crystals of wild-type and Se-Met protein were obtained by mixing 1-μL protein (17 mg/mL) in buffer A with 1.5-μL 22% polyethylene glycol 4000, 0.1M tris pH 8.7, 5 mM NiCl$_2$ reservoir solution.

A BepA construct comprising residues 10–303 and a C-terminal His$_6$ tag (BepA$_{10\text{-}303}$) was expressed and crystallized by mixing 1-μL protein (27 mg/mL, in 20 mM Tris pH 8.0, 0.5M NaCl, and 2 mM β-mercaptoethanol) with 1 μL of 1.5M (NH$_4$)$_2$SO$_4$, 0.1M Tris pH 8.5, and 0.1M Li$_2$SO$_4$.

Crystal structure determination

Diffraction data were collected at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland and were processed with MOSFLM/SCALA[23]; details are given in Table I. SHELXD[24] was used to identify 18 selenium sites of the Se-Met-substituted crystal from the Bijvoet differences determined at the wavelength of the X-ray absorption peak. Phase refinement was performed by SHARP/SOLOMON[25] and was followed by twofold averaging and phase extension using DM.[26] The model was build at the interactive graphics and refined by using REFMAC[27] imposing noncrystallographic twofold symmetry restraints on most of the chain. The resulting model encompasses residues 11–302 with the segment 106–111 and several side-chains undefined by electron density. The refinement statistics is given in Table II.

Activity assay

BepA variants (500 pmol in 20 mM Tris pH 7.5, 100 mM NaCl) were incubated with a mixture of hot (α-[^{32}P]) and cold nucleosyl-tri-phosphate for 30 min at 30°C. The reaction was stopped by the addition of the SDS sample buffer. Samples were boiled for 5 min at 95°C, subjected to SDS-PAGE and analyzed by autoradiography. S100 HeLa cell lysate was prepared as in Ref. 28.

Conclusions and Outlook

The structure of BepA$_{tr}$ and accompanying data allowed us to derive a generic model for the Fic/ATP/Mg^{2+} substrate complex and to propose a detailed mechanism for the adenylylation reaction catalyzed by Fic domains. The identification of the proteins adenylylated by BepA$_{tr}$ in Hela cell extracts will establish a novel class of Fic target proteins different from the known Rho-family GTPases. A subsequent structure/function analysis of the BepA/target complex may allow to prove the proposed adenylylation mechanism and, moreover, will address a putative role of the OB fold in target recognition. Together, these data will enhance our understanding of the molecular function of the Fic domain and, moreover, shed light on the role of the BepA effector protein in subverting endothelial cell functions during chronic *B. henselae* infection.

Crystal Structure of BepA

57

References

1. Schmelz S, Naismith JH (2009) Adenylate-forming enzymes. Curr Opin Struct Biol 19:666–671.
2. Shuman S (2009) DNA ligases: progress and prospects. J Biol Chem 284:17365–17369.
3. Xu Y, Carr PD, Vasudevan SG, Ollis DL (2010) Structure of the adenylylation domain of *E. coli* glutamine synthetase adenylyl transferase: evidence for gene duplication and evolution of a new active site. J Mol Biol 396:773–784.
4. Müller MP, Peters H, Blümer J, Blankenfeldt W, Goody RS, Itzen A (2010) The *Legionella* effector protein DrrA AMPylates the membrane traffic regulator Rab1b. Science 329:946–949.
5. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A (2010) The Pfam protein families database. Nucleic Acids Res 38:D211–D222.
6. Yarbrough ML, Li Y, Kinch LN, Grishin NV, Ball HL, Orth K (2009) AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. Science 323:269–272.
7. Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, Mendez JC, Zekarias B, Lazar C, Dixon JE (2009) The fic domain: regulation of cell signaling by adenylylation. Mol Cell 34:93–103.
8. Dehio C (2008) Infection-associated type IV secretion systems of *Bartonella* and their diverse roles in host cell interaction. Cell Microbiol 10:1591–1598.
9. Schulein R, Guye P, Rhomberg TA, Schmid MC, Schröder G, Vergunst AC, Carena I, Dehio C (2005) A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. Proc Natl Acad Sci USA 102:856–861.
10. Schmid MC, Schulein R, Dehio M, Denecker G, Carena I, Dehio C (2004) The VirB type IV secretion system of *Bartonella henselae* mediates invasion, proinflammatory activation and antiapoptotic protection of endothelial cells. Mol Microbiol 52:81–92.
11. Schmid MC, Scheidegger F, Dehio M, Balmelle-Devaux N, Schulein R, Guye P, Chennakesava CS, Biedermann B, Dehio C (2006) A translocated bacterial protein protects vascular endothelial cells from apoptosis. PLoS Pathog 2:1083–1097.
12. Xiao J, Worby CA, Mattoo S, Sankaran B, Dixon JE (2010) Structural basis of Fic-mediated adenylylation. Nat Struct Mol Biol 17:1004–1010.
13. Das D, Krishna SS, McMullan D, Miller MD, Xu Q, Abdubek P, Acosta C, Astakhova T, Axelrod HL, Burra P, Carlton D, Chiu H-J, Clayton T, Deller MC, Duan L, Elias Y, Elsliger M-A, Ernst D, Feuerhelm J, Grzechnik A, Grzechnik SK, Hale J, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kozbial P, Kumar A, Marciano D, Morse AT, Murphy KD, Nigoghossian E, Okach L, Oommachen S, Paulsen J, Reyes R, Rife CL, Sefcovic N, Tien H, Trame CB, Trout CV, van den Bedem H, Weekes D, White A, Hodgson KO, Wooley J, Deacon AM, Godzik A, Lesley SA, Wilson IA (2009) Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gi|24375750) from *Shewanella oneidensis* MR-1 at 1.6 A resolution. Proteins 75:264–271.
14. Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res 38 (Suppl): W545–W549.
15. Garcia-Pino A, Christensen-Dalsgaard M, Wyns L, Yarmolinsky M, Magnuson R, Gerdes K, Loris R (2008) Doc of prophage P1 is inhibited by its antitoxin partner Phd though fold complementation. J Biol Chem 283: 30821–30827.
16. Kinch LN, Yarbrough ML, Orth K, Grishin NV (2009) Fido, a novel AMPylation domain common to fic, doc, and AvrB. PLoS ONE 4:1–9.
17. Watson JD, Milner-White EJ (2002) A novel main-chain anion-binding site in proteins: the nest. A particular combination of phi,psi values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions. J Mol Biol 315:171–182.
18. Desogus G, Todone F, Brick P, Onesti S (2000) Active site of lysyl-tRNA synthetase: structural studies of the adenylation reaction. Biochemistry 39:8418–8425.
19. Desveaux D, Singer AU, Wu A-J, McNulty BC, Musselwhite L, Nimchuk Z, Sondek J, Dangl JL (2007) Type III effector activation via nucleotide binding, phosphorylation, and host target interaction. PLoS Pathog 3:456–459.
20. Luong P, Kinch LN, Brautigam CA, Grishin NV, Tomchick DR, Orth K (2010) Kinetic and structural insights into the mechanism of AMPylation by VopS FIC domain. J Biol Chem 285:20155–20163.
21. Wilmouth RC, Clifton IJ, Robinson CV, Roach PL, Aplin RT, Westwood NJ, Hajdu J, Schofield CJ (1997) Structure of a specific acyl-enzyme complex formed between beta-casomorphin-7 and porcine pancreatic elastase. Nat Struct Biol 4:456–462.
22. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol 229:105–124.
23. Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. In Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography, no. 27. pp 30–30.
24. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. Acta Cryst D 58:1772–1779.
25. De La Fortelle E, Bricogne G (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol 276:472–493.
26. Cowtan KD, Zhang KY (1999) Density modification for macromolecular phase improvement. Prog Biophys Mol Biol 72:245–270.
27. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Cryst D 53:240–255.
28. Mukherjee S, Keitany G, Li Y, Wang Y, Ball HL, Goldsmith EJ, Orth K (2006) *Yersinia* YopJ acetylates and inhibits kinase activation by blocking phosphorylation. Science 312:1211–1214.

Palanivelu et al.   PROTEIN SCIENCE | VOL 20:492–499   **499**

58

Supplemental material to

Fic domain catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA

Dinesh V. Palanivelu[1,2], Arnaud Goepfert[1,4], Marcel Meury[1,3], Patrick Guye[4,5],
Christoph Dehio[4] and Tilman Schirmer[1]

[1] Core program of Structural Biology and Biophysics, and [4]Focal Area of Infection Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

[2] Current address: Departments of Biochemistry and Biophysics and Cellular and Molecular Pharmacology, University of California, QB3- Minor Lab Box 2532, 1700 4th Street, San Francisco, CA 94158-2532, U.S.A.

[3] Current address: Institute of Biochemistry and Molecular Medicine, University of Bern, Bühlstrasse 28, CH-3012 Bern, Switzerland

[5] Current address: MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge, MA 02139, USA

Corresponding author: tilman.schirmer@unibas.ch
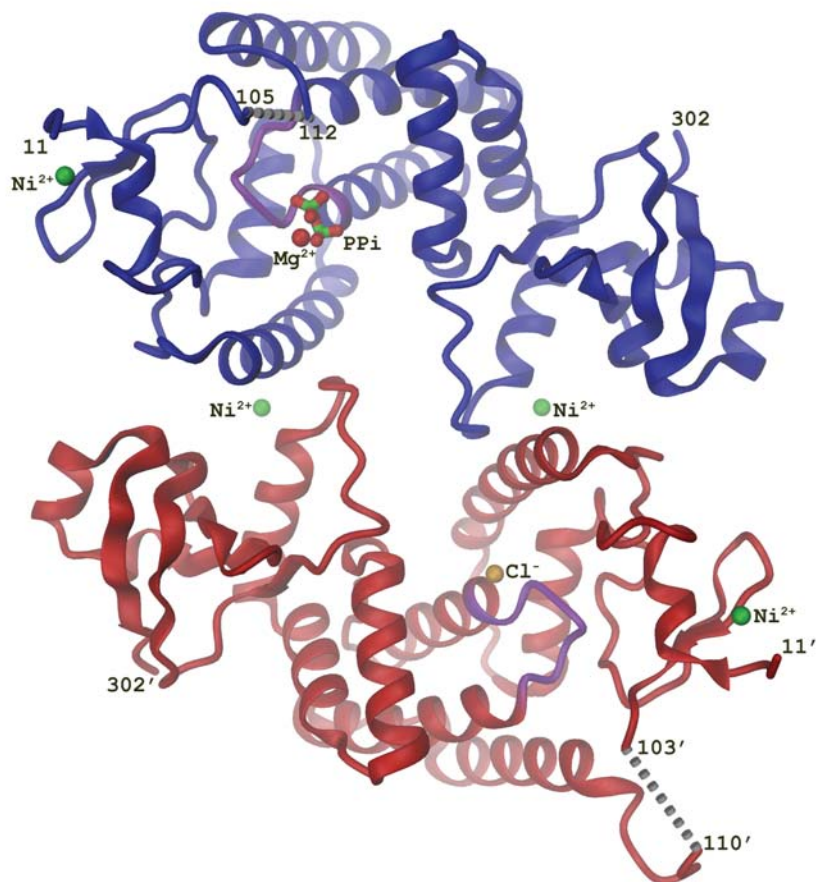
Fig. S1. Crystal structure of BepA$_{tr}$/PP$_i$/Mg^{2+} complex (2jk8). As for the isomorphous native BepA$_{tr}$ crystal (2vy3), the asymmetric unit contains two molecules arranged to a 2-fold symmetric dimer. The two chains show a rmsd of 0.69 Å for 275 C$_\alpha$ positions after superposition. The dimer is stabilized by a putative nickel ion that is tetrahedrally coordinated by residues H45, K49, H231' (the prime indicates a residue of the adjacent monomer) and a water molecule. Due to the local two-fold symmetry, this interaction occurs twice. In addition, there is a second nickel binding site in each monomer found near the N-terminus coordinated by residues H17, H18 and two water molecules. The X-ray fluorescent spectrum confirmed the enrichment of nickel within the crystal (data not shown). The blue subunit carries a PP$_i$/Mg^{2+} moiety bound to the signature loop (magenta), whereas the other has a chloride ion bound to the same location.

60
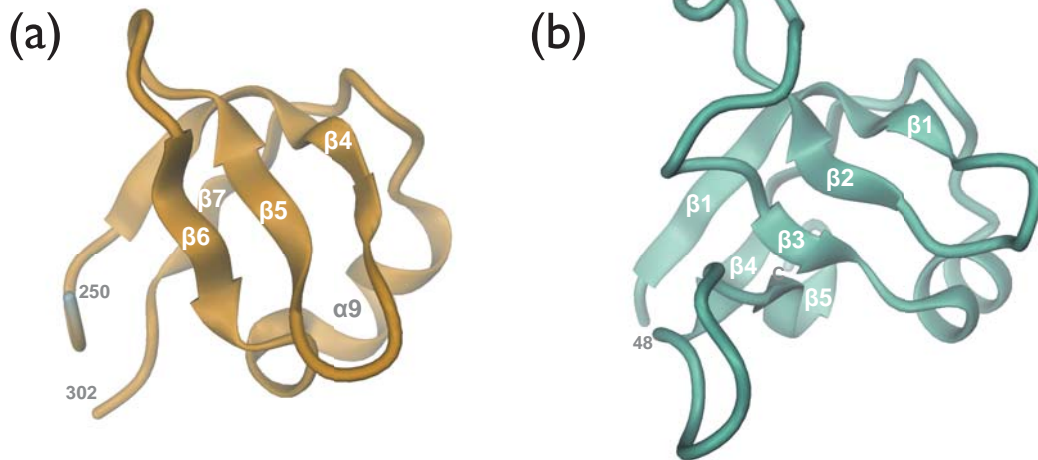
Fig. S2. OB-fold of (a) residues 250 - 302 of BepA$_{tr}$ (2vy3) and (b) of a Rho Transcription Termination Factor (2a8v), for comparison. Note, that the C-terminal strand β5 has no equivalent in BepA$_{tr}$.
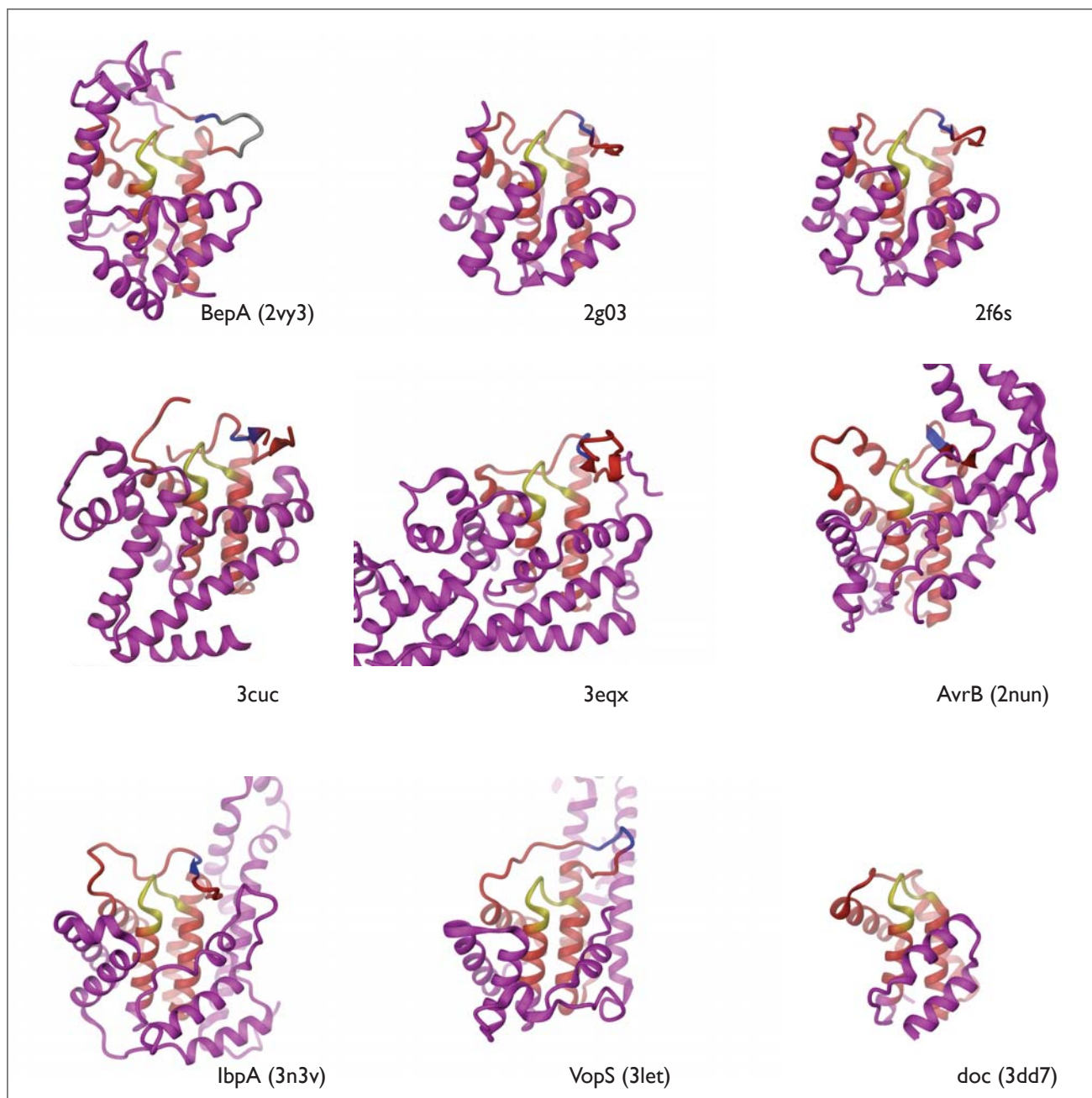
Fig. S3. Gallery of protein structures with Fic fold. Color code as in Fig. 1 of main text. PDB codes are indicated.

(a)

protein	PDB-code	Z-score	signature motif
BepA	2vy3	–	1.....H_{159} PFREGNGRTQR..... 544
	2f6s	13.5	1.....H_{96} PFLEGNGRATR..... 177
	2g03	13.4	1.....H_{107} PFLEGNGRSTR..... 191
	3cuc	11.7	1.....H_{192} PFEDGNGRIAR..... 479
SO2466	3eqx	11.5	1.....H_{198} PFIDGNGRTGR..... 372
AvrB	1nh1	8.0	1.....M_{262} PDQRGSAAKAE..... 321
IbpA	3n3u	8.0	1.....H_{3717} PFAEGNGRMAR..... 4095
Doc	3dd7	5.8	1.....Y_{66} IFNDANKRTAL..... 126
VopS	3let	4.7	1.....H_{348} GFTDGNGRMGR..... 387

(b)

Cross-structure statistics

RMSD

structure:	1	2	3	4	5	6	7	8	9
1 PDB 2vy3:A		3.408	3.104	2.403	2.208	2.064	2.058	2.607	2.646
2 PDB 3let:A	3.408		3.279	3.403	3.104	3.231	3.252	3.287	3.060
3 PDB 3dd7:A	3.104	3.279		3.148	2.899	2.883	2.909	3.034	3.388
4 PDB 3eqx:A	2.403	3.403	3.148		2.337	2.460	2.435	2.862	2.796
5 PDB 3cuc:A	2.208	3.104	2.899	2.337		2.147	2.075	2.487	2.096
6 PDB 2g03:A	2.064	3.231	2.883	2.460	2.147		0.413	2.300	2.905
7 PDB 2f6s:A	2.058	3.252	2.909	2.435	2.075	0.413		2.287	2.890
8 PDB 1nh1:A	2.607	3.287	3.034	2.862	2.487	2.300	2.287		2.712
9 PDB 3n3u:A	2.646	3.060	3.388	2.796	2.096	2.905	2.890	2.712	

Q-score

structure:	1	2	3	4	5	6	7	8	9
1 PDB 2vy3:A		0.036	0.097	0.041	0.061	0.095	0.094	0.049	0.048
2 PDB 3let:A	0.036		0.089	0.029	0.044	0.063	0.061	0.037	0.040
3 PDB 3dd7:A	0.097	0.089		0.076	0.116	0.173	0.168	0.100	0.089
4 PDB 3eqx:A	0.041	0.029	0.076		0.047	0.067	0.066	0.036	0.036
5 PDB 3cuc:A	0.061	0.044	0.116	0.047		0.103	0.104	0.056	0.064
6 PDB 2g03:A	0.095	0.063	0.173	0.067	0.103		0.223	0.089	0.072
7 PDB 2f6s:A	0.094	0.061	0.168	0.066	0.104	0.223		0.088	0.072
8 PDB 1nh1:A	0.049	0.037	0.100	0.036	0.056	0.089	0.088		0.047
9 PDB 3n3u:A	0.048	0.040	0.089	0.036	0.064	0.072	0.072	0.047	

Sequence Identity

structure:	1	2	3	4	5	6	7	8	9
1 PDB 2vy3:A		0.165	0.059	0.153	0.153	0.271	0.247	0.094	0.153
2 PDB 3let:A	0.165		0.129	0.141	0.176	0.141	0.129	0.047	0.212
3 PDB 3dd7:A	0.059	0.129		0.106	0.129	0.165	0.153	0.106	0.118
4 PDB 3eqx:A	0.153	0.141	0.106		0.247	0.153	0.165	0.094	0.212
5 PDB 3cuc:A	0.153	0.176	0.129	0.247		0.141	0.129	0.141	0.318
6 PDB 2g03:A	0.271	0.141	0.165	0.153	0.141		0.682	0.059	0.188
7 PDB 2f6s:A	0.247	0.129	0.153	0.165	0.129	0.682		0.082	0.176
8 PDB 1nh1:A	0.094	0.047	0.106	0.094	0.141	0.059	0.082		0.082
9 PDB 3n3u:A	0.153	0.212	0.118	0.212	0.318	0.188	0.176	0.082	

Fig. S4. (a) Structural similarity Z-scores as determined by Dali[1] for BepA$_{tr}$ compared to all other Fic domain proteins. Also shown are the sequences of the active loop. (b) Sequence and structure similarity of proteins with Fic fold as determined by the PDBeFold server[2].
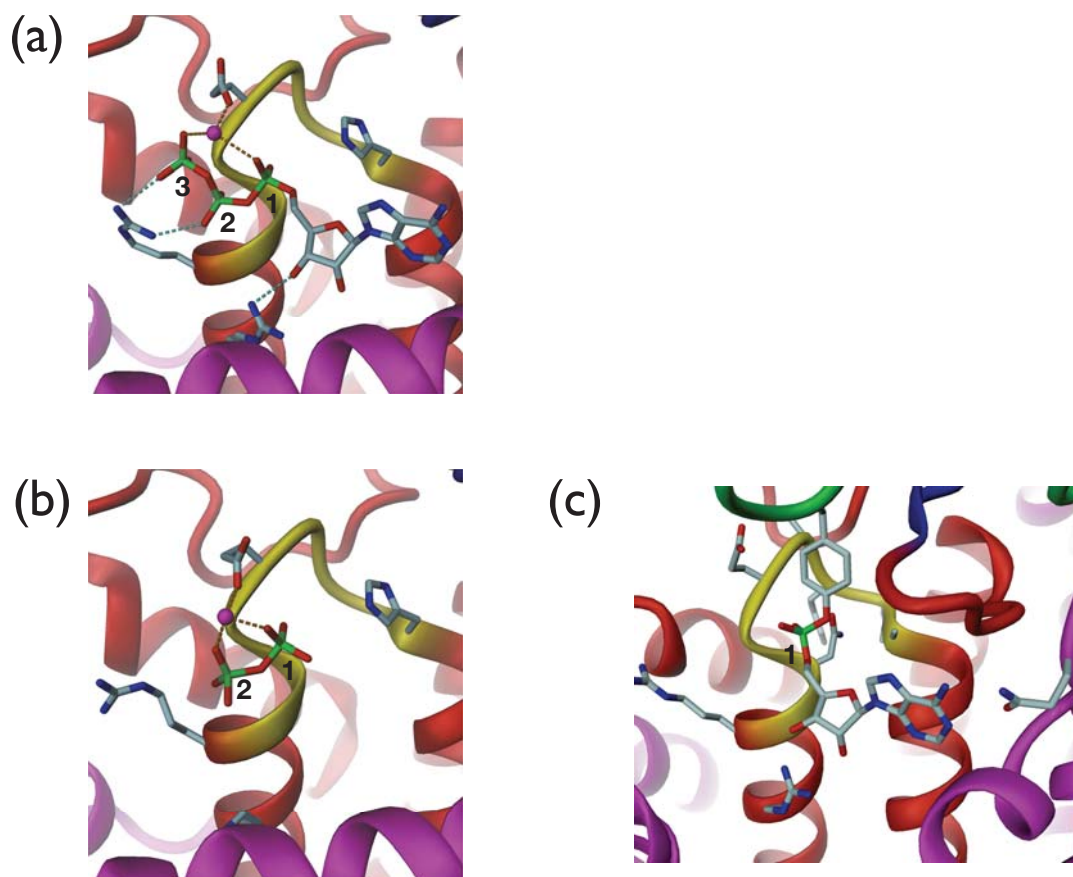
Fig. S5. Fic/ATP/Mg^{2+} substrate model and experimental Fic/product structures. (a) BepA(Fic)/ ATP/Mg^{2+} model with the substrate phosphates occupying binding sub-sites 1 (anion binding LR nest), 2, and 3. (b) BepA(Fic) in complex with PP$_i$/Mg^{2+} (2jk8) with phosphates bound to sub-sites 1 and 2. (c) Fic2 domain of IbpA(H3717A) in complex with AMP-Y$_{32}$ of cdc42 (green) with phosphate bound to the sub-site 1, PDB-code 3n3v.

References

1.   Holm, L. & Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38 Suppl**, W545-9.

2.   Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256-68.

3.2   Research article II

Adenylylation control by intra- or intermolecular active site obstruction in Fic proteins

Philipp Engel*, Arnaud Goepfert*, Frédéric V. Stanger, Alexander Harms, Alexander Schmidt, Tilman Schirmer, and Christoph Dehio

*These authors contributed equally to this work

Statement of the own participation

I contributed to this publication by expressing, purifying and crystallizing VbhA–VbhT(FIC), NmFic(SE/AA) and NmFic(Δ8), and by determining their structures. I also performed *in vitro* adenylylation assays together with Alexander Harms. Philipp Engel discovered and physiologically characterized VbhT–VbhA as a toxin–antitoxin module and carried out the bioinformatic analysis. Frédéric Stanger expressed, purified and crystallized NmFic with AMPPNP and determined the structure. Philipp Engel, Frédéric Stanger and Alexander Harms cloned recombinant plasmids. Alexander Harms carried out the growth curve experiment. Alexander Schmidt conducted the mass spectrometry analysis. The manuscript was written by me, Philipp Engel, Tilman Schirmer, and Christoph Dehio.

LETTER

Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins

Philipp Engel[1]†*, Arnaud Goepfert[1,2]*, Frédéric V. Stanger[1,2], Alexander Harms[1], Alexander Schmidt[3], Tilman Schirmer[2] & Christoph Dehio[1]

Fic proteins that are defined by the ubiquitous FIC (filamentation induced by cyclic AMP) domain are known to catalyse adenylylation (also called AMPylation); that is, the transfer of AMP onto a target protein. In mammalian cells, adenylylation of small GTPases through Fic proteins injected by pathogenic bacteria can cause collapse of the actin cytoskeleton and cell death[1,2]. It is unknown how this potentially deleterious adenylylation activity is regulated in the widespread Fic proteins that are found in all domains of life and that are thought to have critical roles in intrinsic signalling processes. Here we show that FIC-domain-mediated adenylylation is controlled by a conserved mechanism of ATP-binding-site obstruction that involves an inhibitory α-helix (α_{inh}) with a conserved (S/T)XXXE(G/N) motif, and that in this mechanism the invariable glutamate competes with ATP γ-phosphate binding. Consistent with this, FIC-domain-mediated growth arrest of bacteria by the VbhT toxin of *Bartonella schoenbuchensis* is intermolecularly repressed by the VbhA antitoxin through tight binding of its α_{inh} to the FIC domain of VbhT, as shown by structure and function analysis. Furthermore, structural comparisons with other bacterial Fic proteins, such as Fic of *Neisseria meningitidis* and of *Shewanella oneidensis*, show that α_{inh} frequently constitutes an amino-terminal or carboxy-terminal extension to the FIC domain, respectively, partially obstructing the ATP binding site in an intramolecular manner. After mutation of the inhibitory motif in various Fic proteins, including the human homologue FICD (also known as HYPE), adenylylation activity is considerably boosted, consistent with the anticipated relief of inhibition. Structural homology modelling of all annotated Fic proteins indicates that inhibition by α_{inh} is universal and conserved through evolution, as the inhibitory motif is present in ~90% of all putatively adenylylation-active FIC domains, including examples from all domains of life and from viruses. Future studies should reveal how intrinsic or extrinsic factors modulate adenylylation activity by weakening the interaction of α_{inh} with the FIC active site.

In two Fic proteins, IbpA and VopS, that are translocated by pathogenic bacteria into host cells, the ubiquitous FIC domain has been shown to catalyse adenylylation[1–4]. The crystal structure of the effector domain IbpA(FIC2) in complex with its adenylylated host target Cdc42 has been reported[5] and a catalytic mechanism has been proposed[5,6]. IbpA- or VopS-mediated adenylylation of Rho-family GTPases abolishes downstream signalling in human cells and, thus, causes actin cytoskeleton collapse and cell death[1,2]. By contrast, overexpression of a human Fic protein with similar target specificity, HYPE, had only a marginal effect[1]. This suggests that the potentially deleterious adenylylation activity is tightly regulated in HYPE and probably in most of the almost 3,000 Fic proteins that are proposed to have important roles in intrinsic signalling processes in bacteria, archaea and eukaryotes.

VbhT is a bacterial Fic protein of the mammalian pathogen *B. schoenbuchensis*[7,8]. It is composed of an N-terminal FIC domain

and a C-terminal BID domain (Fig. 1a). The BID domain facilitates protein translocation into mammalian or bacterial target cells through a type IV secretion system or conjugation machinery[9], respectively, but the target cell and functional role of VbhT are unknown. VbhT arrests growth when expressed in *Escherichia coli* (Fig. 1c and Supplementary Fig. 1a). Growth arrest is repressed by mutation of histidine to alanine (VbhT(H136A)) in the conserved FIC motif HXFX(D/E)GNGRXXR. In other Fic proteins, this signature motif has been shown to be essential for target protein adenylylation activity[1,2,10], therefore suggesting that toxicity is related to adenylylation of endogenous proteins. Indeed, wild-type VbhT, but not VbhT(H136A), catalysed *in vitro* adenylylation of a putative *E. coli* target protein of approximately 80 kilodaltons (kDa) (Fig. 1d). Furthermore, *E. coli* cells showed filamentation after expression of wild-type VbhT, but not of VbhT(H136A) (Supplementary Fig. 2). A similar phenotype has been described for what is thought to be a hyperactive mutant of the *E. coli* Fic protein[11]. Co-expression of VbhT with VbhA, encoded by the small open reading frame *vbhA*
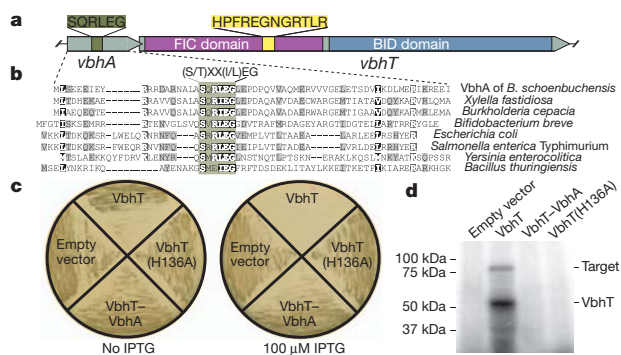
Figure 1 | The small protein VbhA represses the toxic effect (growth arrest) that is mediated by the adenylylation activity of VbhT in *E. coli*. a, Genetic organization of the type-IV-secretion-system-associated locus of *B. schoenbuchensis*, which is composed of the overlapping *vbhA* and *vbhT* genes. The FIC and BID domains that are encoded by *vbhT* are shown in different colours. Protein translations of the regions encoding the conserved motif of VbhA and the FIC motif of VbhT are depicted. **b**, Protein alignment of VbhA and a subset of the 158 homologues that are encoded upstream of *fic* loci in different bacteria (see also Supplementary Fig. 4). The most conserved region shows a (S/T)XX(I/L)EG consensus. Sequence accessions and strain designations are given in Supplementary Fig. 4. **c**, Growth of *E. coli* on lysogeny broth (LB) plates after IPTG-induced expression of different VbhT and VbhA constructs. Expression of VbhT shows a toxic effect, whereas bacterial growth is not affected when VbhT(H136A) is expressed or when VbhT and VbhA are co-expressed (VbhT–VbhA). All *E. coli* strains revealed normal growth without induction. Growth curve experiments in LB gave analogous results (Supplementary Fig. 1). **d**, Adenylylation assay with crude cell lysates of *E. coli* ectopically expressing the same constructs as in panel **c**.

[1]Focal Area Infection Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. [2]Core Program Structural Biology and Biophysics, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. [3]Proteomics Core Facility, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. †Present address: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520-8106, USA. *These authors contributed equally to this work.

69

immediately upstream of *vbhT* (Fig. 1a), completely repressed VbhT toxicity, as shown by wild-type-like bacterial growth, normal cell morphology, and inhibition of VbhT-dependent adenylylation (Fig. 1 and Supplementary Figs 1a and 2). We also observed VbhT-mediated toxicity and its repression by VbhA in *B. schoenbuchensis*, the natural carrier of this toxin and antitoxin, and in the related species *Bartonella henselae* (Supplementary Fig. 3).

The inhibitory action of VbhA on the VbhT toxin, and the genetic organization of the respective genes in an operon are reminiscent of toxin–antitoxin modules that are found in many bacterial genomes, often associated with mobile genetic elements[12]. A comprehensive analysis of the upstream region of FIC-domain-encoding genes (PFAM pf02661) identified 158 bacterial *vbhA* homologues that probably function as antitoxins. Although the sequences are rather diverse, a central (S/T)XX(I/L)EG motif is conspicuous (Fig. 1b and Supplementary Fig. 4). The high-resolution (1.5 Å) crystal structure of VbhA in complex with the FIC domain of VbhT (VbhT(FIC)) (Supplementary Table 1) shows that VbhA is folded into two anti-parallel helices that tightly embrace VbhT(FIC) (Fig. 2a) with the N-terminal helix (α_{inh}), adopting a location that is analogous but distinct to that of the antitoxin Phd in its complex with Doc[13] (Supplementary Fig. 5). Doc is a Fic protein with a degenerate, probably adenylylation-incompetent FIC motif[14] that may have adopted another toxic activity (Supplementary Information, section 1). The VbhA antitoxin motif locates to the C-terminal part of α_{inh} and is positioned close to the putative ATP-binding site[10] at the N-cap of the helix that follows the active loop of VbhT(FIC). This suggests that the antitoxin competes with ATP binding. VbhA residues Ser 20 and Glu 24 of the inhibitory motif form a hydrogen bond and a salt bridge, respectively, with the conserved Arg147 of VbhT following the active loop.

Intriguingly, structural comparison with other bacterial Fic proteins of known fold (Fic proteins from *S. oneidensis* (SoFic)[14,15] and from *N. meningitidis* (NmFic)[14] (Fig. 2b, c), and from *Bacteroides theataio-taomicron* (BtFic)[14] and from *Helicobacter pylori* (HpFic)[14] (Supplementary Fig. 6)) reveals that a structural equivalent of α_{inh} can be part of the FIC domain fold itself. Moreover, these proteins also show the SXXXE(G/N) inhibitory motifs that are, with respect to the FIC active site, arranged exactly as in the VbhA-VbhT(FIC) complex. Along the polypeptide chain, however, these α_{inh} occur at two distinct locations either in the N-terminal part (SoFic and BtFic) or at the C terminus (NmFic and HpFic). Thus, Fic proteins containing α_{inh} can be grouped into three classes (Fig. 2d) depending on whether α_{inh} is provided by an interacting antitoxin (class I) or whether it is part of the FIC fold as an N-terminal helix (class II) or a C-terminal helix (class III).

To investigate the distribution of class II and class III Fic proteins, we predicted the structures of all PFAM FIC domain entries by homology modelling (Supplementary Information, section 1). Including the class I proteins that are identified above, two-thirds of all Fic proteins were classified (Fig. 2e and Supplementary Table 2), with a strong dominance of class II, the class to which human HYPE belongs (Supplementary Tables 3 and 4). The proportion of classified Fic proteins increases to 90% when considering only adenylylation-competent Fic proteins that are defined by compliance with the HXFX(D/E)GNGRXXR motif (Supplementary Text). This suggests co-evolution of catalytic and inhibitory function. The inhibition motifs that are derived from Psi-Blast (class I) and structural predictions (classes II and III) are shown in Fig. 2f with the overall consensus being (S/T)XXXE(G/N). The strict conservation of the glutamate is striking and indicates that the observed ionic interaction with the second arginine of the conserved FIC motif (Figs 2a–c) is crucial for inhibition. Phylogenetic distribution of class I
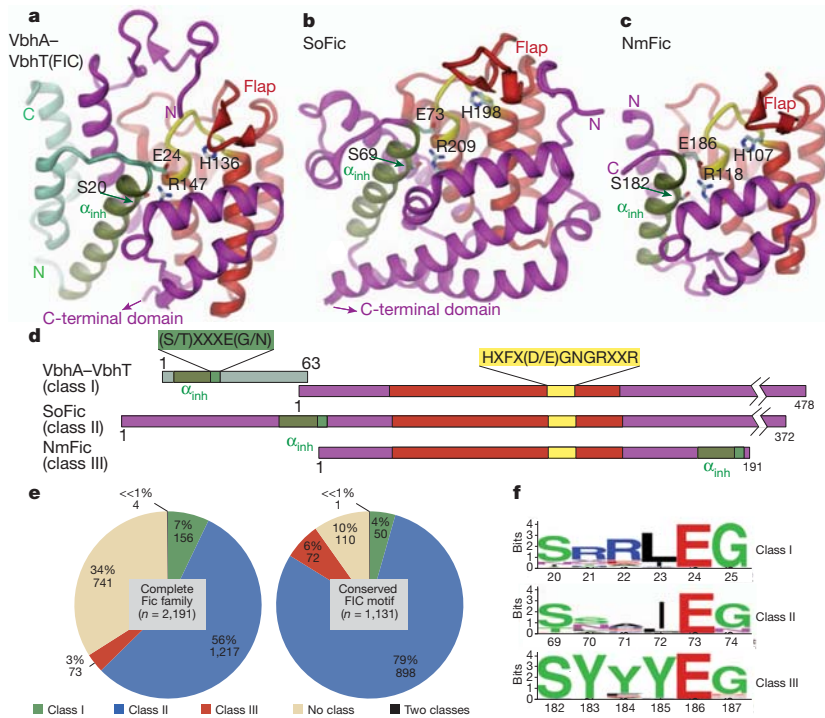
Figure 2 | Structures and classification of Fic proteins according to the position of the inhibitory motif along the polypeptide chain. Structures are shown in cartoon (FIC core as defined by PFAM, red; active site loop with histidine, yellow; inhibitory helix (α_{inh}) with SXXXE(G/N) motif, green) with important residues in full. **a**, Complex of VbhT(FIC) with antitoxin VbhA (green). **b**, SoFic (from *S. oneidensis*, PDB code 3EQX) with C-terminal domain omitted. **c**, NmFic (from *N. meningitidis*, PDB code 2G03). **d**, Linear organization of motifs in the proteins presented in panels **a–c**. **e**, Distribution of the three Fic protein classes for the entire family in the PFAM database (left) and the subset of Fic proteins with the conserved FIC motif HXFX[(D/E)GNGRXXR that is likely to convey adenylylation activity (right; see also Supplementary Information, section 1). **f**, Sequence profiles for the inhibition site of the three Fic protein classes.

70

and class III Fic proteins indicates that each is of monophyletic origin (Supplementary Fig. 7). Fic proteins with a degenerate FIC motif are dispersed over the tree, with the exception of the large cluster of Doc-like toxins. This suggests that there is recurrent degeneration of the conserved FIC motif with concomitant loss of adenylylation activity. Consistent with this, deterioration of the FIC motif seems to correlate with the absence of a recognizable inhibitory motif (Supplementary Fig. 7). Fic proteins with a degenerate FIC motif may display catalytic activities different from adenylylation, such as phosphocholination, as reported for the *Legionella pneumophila* effector AnkX[16].

Owing to our discovery of the prevalence of the inhibitory motif in Fic proteins, we carried out a detailed analysis of its functional and structural role. For this, NmFic (class III, Fig. 2c) was chosen, as it is the smallest active Fic family protein with known crystallization condition (Supplementary Table 1). As reported before[6], NmFic exhibits *in vitro* auto-adenylylation activity (Fig. 3d). The acceptor site was traced to Y183 of α_{inh} by mass spectrometry (Supplementary Information, section 1, and Supplementary Fig. 8). On the basis of the location of Y183 relative to the active site, auto-adenylylation is probably catalysed intermolecularly after partial unfolding or detachment of α_{inh}. Addition of *E. coli* lysate to NmFic did not reveal additional bands on the autoradiograph, and this indicated that there are no NmFic targets in *E. coli* or that the activity of NmFic is inhibited. The latter was shown to be correct as mutation of the inhibitory motif (S182A/E186A, NmFic(SE/AA)) resulted in transfer of radioactivity onto an ~80-kDa *E. coli* protein and enhanced auto-adenylylation with an additional acceptor site (Y188; Fig. 3d and Supplementary Fig. 9). Deletion of the entire α_{inh} helix (NmFic(Δ8)) led to similar target protein adenylylation, proving that the activity resides in the FIC domain core. However, only weak auto-adenylylation was apparent owing to the lack of the acceptor tyrosines in this deletion mutant (Fig. 3d and Supplementary Fig. 8c).

To investigate the inhibitory mechanism, crystals of NmFic proteins were soaked with the non-hydrolysable ATP analogue adenylyl

imidodiphosphate (AMPPNP) (Supplementary Table 1). The NmFic–AMPPNP structure revealed nucleotide binding but with the γ-phosphate disordered (Fig. 3a). Notably, the orientation of the α-phosphate seems to be non-productive, as the position that is in line with the scissile Pα-O3α bond is occluded by H107 and N113 (Supplementary Fig. 10a and Supplementary Movie). To reveal the situation in an inhibition-relieved mutant, the structure of NmFic(SE/AA) was determined to 3.0 Å (Supplementary Fig. 11). Electron density was lacking for α_{inh}, indicating disorder, whereas the nucleotide conformation was well defined. In the NmFic(Δ8)–AMPPNP structure, the same nucleotide conformation was observed (Fig. 3b) and, owing to its high resolution (1.7 Å), the structural basis for the observed relief of inhibition in these mutants became evident. Whereas the adenosine moiety adopts the same position as in the wild-type, the γ-phosphate of the nucleotide is bound to R118, occupying the same position as the carboxylate of the inhibitory E186 in the wild type. As a consequence, the α-phosphate is found to be re-oriented, and the new orientation permits in line attack of a target side chain onto the α-phosphorus to accomplish AMP transfer (Fig. 3c, Supplementary Fig. 10 and Supplementary Movie).

The exact role of the inhibitory glutamate was investigated further by mutagenesis of Fic proteins from the three regulatory classes. In wild-type NmFic, Cα and Cβ of the glutamate are close to the position that is attained by the γ-phosphate position in NmFic(Δ8) (Fig. 3c). Still, an E to G single point mutant may provide sufficient space and main-chain flexibility to allow γ-phosphate binding. Indeed, similar to NmFic(Δ8) and NmFic(SE/AA), the mutant NmFic(E186G) resulted in growth inhibition of *E. coli* (Supplementary Fig. 1b). Likewise, co-expression of VbhT(FIC) and VbhA(E24G), as representatives for class I, caused *E. coli* growth defects (Supplementary Fig. 1a). We also included SoFic, a bacterial class II protein, in this analysis. Consistent with the effects of E to G single point mutants in NmFic and VbhA, mutant SoFic(E73G) revealed a negative effect on *E. coli* growth
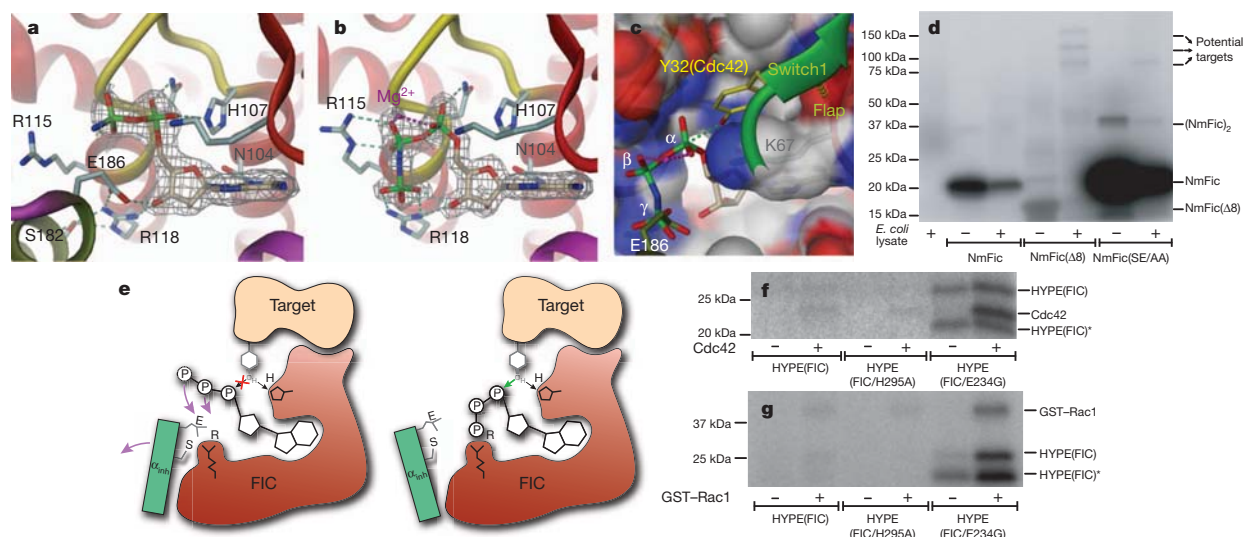
Figure 3 | Structure and function of wild-type and mutant NmFic reveals a general inhibition mechanism corroborated by HYPE protein analysis.
a, Active site of NmFic with bound ATP analogue AMPPNP. The γ-phosphate seems to be disordered and has not been modelled. **b**, Active site of NmFic(Δ8) with bound AMPPNP and Mg^{2+}. The γ-phosphate occupies the position taken by E186 of the SXXXEG motif in the wild-type (as shown in **a**). Also shown are $2Fo - Fc$ maps that are contoured at 1.2σ and cover the ligands. **c**, Surface representation of the active site of NmFic(Δ8) with modifiable Y32 (yellow) of the Cdc42 switch1 loop (green) in a position that is obtained from a superposition of the active site loops of NmFic over the IbpA(FIC2)–Cdc42 complex[5]. **d**, Autoradiography of an SDS gel after incubation of wild-type and mutant NmFic with [α-^{32}P]ATP in the presence or absence of *E. coli* lysate. The

mutants, but not the wild-type, catalyse AMP transfer onto *E. coli* target proteins. **e**, Scheme of a general inhibition mechanism for Fic proteins. The α_{inh} helix (green) with the (S/T)XXXE(G/N) motif prevents productive ATP binding. It is only after dissociation of the helix that the ATP γ-phosphate attains the position close to a conserved arginine (indicated by "R") of the FIC motif. This is accompanied by reorientation of the α-phosphate to allow in-line attack of the target hydroxyl side chain after proton transfer to the active histidine as proposed before[5]. **f**, **g**, Autoradiography after incubation of HYPE(FIC) with α-^{32}P-ATP in the presence or absence of Cdc42 (**f**) or GST-tagged Rac1 (GST–Rac1) (**g**). HYPE(FIC/E234G) shows enhanced auto-adenylylation and target adenylylation. HYPE(FIC)* denotes a degradation product of HYPE(FIC).

71

(Supplementary Fig. 1c) and strongly enhanced auto-adenylylation (Supplementary Fig. 12). A similar marked effect on auto-adenylylation was observed after mutation of the corresponding residue in human HYPE (HYPE(FIC/E234G)), another class II protein (Fig. 3f, g). It is known that HYPE catalyses Cdc42 and Rac1 adenylylation[1,4]. The rather low adenylylation activity of wild-type HYPE on these substrates is markedly enhanced in the HYPE(FIC/E234G) mutant (Fig. 3f, g), demonstrating that the relief of inhibition by α_{inh} enhances not only auto-adenylylation but also AMP transfer onto bona fide protein targets.

From this study a general mechanism for the inhibition of the FIC-domain-mediated adenylylation has emerged that invokes a glutamate finger from α_{inh} to inter- or intramolecularly block part of the ATP binding site (Fig. 3e). Dissociation of the toxin–antitoxin complex (class I) or intramolecular loosening of the contact between α_{inh} and the FIC domain active site (classes II and III) allows ordered binding of the entire ATP moiety with the α-phosphate in an orientation that is productive for accepting an incoming target hydroxyl side chain.

Class I proteins might exert functions similar to the classical bacterial toxin–antitoxin modules[12], whereas class II and III Fic proteins seem to regulate intrinsic cellular functions that are related to physiological adaptation and cell homeostasis. Conservation of class II Fic proteins in all three kingdoms of life (Supplementary Table 2) emphasizes the important role of the regulatory mechanisms described here. How the adenylylation activity is activated by weakening the interaction of α_{inh} with the FIC active site in class II and class III Fic proteins is one of many important questions for future research. We also anticipate our study to be a starting point for rational approaches to modulate the adenylylation activities of Fic proteins—approaches that should aid in elucidating the diverse biological functions of these widespread signalling proteins.

METHODS SUMMARY

Toxicity experiments were performed with *E. coli* strain MG1655, which encodes an IPTG (isopropyl-β-D-thiogalactoside)-inducible T7 polymerase (strain AB472). Protein expression was controlled by the addition of either $100\,\mu M$ IPTG (induction) or 1% glucose (repression). A PSI-BLAST search for homologues with VbhA of *B. schoenbuchensis* was conducted to identify Fic proteins belonging to class I. Class II and class III proteins were classified by structure predictions using the program HHpred[17]. We analysed all 2,189 proteins of the FIC PFAM family (pf02661, release 24). Adenylylation assays were carried out using bacterial crude cell lysates for VbhT constructs or purified proteins for NmFic, SoFic and HYPE constructs. For structure determination, VbhA and VbhT(FIC) were co-expressed in the BL21(DE3) strain of *E. coli*, and NmFic, SoFic(Δ8) and NmFic(SE/AA) in the BL21-AI strain of *E. coli*. VbhA–VbhT(FIC) and NmFic were purified by affinity chromatography followed by size exclusion chromatography. An additional anion exchange chromatography step was performed for NmFic(Δ8) and NmFic(SE/AA). Diffraction data were collected at beamline X06SA of the Swiss Light Source. For NmFic(Δ8) and NmFic(SE/AA), we obtained phases by molecular replacement using the NmFic structure (PDB code 2G03) as a search model. The VbhA–VbhT(FIC) complex was solved by molecular replacement using a fragment of BepA (PDB code 2JK8).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 July; accepted 24 November 2011.
Published online 22 January 2012.

1. Worby, C. A. *et al.* The fic domain: regulation of cell signaling by adenylylation. *Mol. Cell* **34**, 93–103 (2009).
2. Yarbrough, M. L. *et al.* AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* **323**, 269–272 (2009).
3. Roy, C. R. & Mukherjee, S. Bacterial FIC proteins AMP up infection. *Sci. Signal.* **2**, pe14 (2009).
4. Mattoo, S. *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).
5. Luong, P. *et al.* Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J. Biol. Chem.* **285**, 20155–20163 (2010).
6. Xiao, J., Worby, C. A., Mattoo, S., Sankaran, B. & Dixon, J. E. Structural basis of Fic-mediated adenylylation. *Nature Struct. Mol. Biol.* **17**, 1004–1010 (2010).
7. Dehio, C. *et al. Bartonella schoenbuchii* sp. nov., isolated from the blood of wild roe deer. *Int. J. Syst. Evol. Microbiol.* **51**, 1557–1565 (2001).
8. Engel, P. *et al.* Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen *Bartonella. PLoS Genet.* **7**, e1001296 (2011).
9. Schulein, R. *et al.* A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. *Proc. Natl Acad. Sci. USA* **102**, 856–861 (2005).
10. Palanivelu, D. V. *et al.* Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein Sci.* **20**, 492–499 (2011).
11. Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M. & Komano, T. Involvement of cyclic AMP and its receptor protein in filamentation of an *Escherichia coli* fic mutant. *J. Bacteriol.* **151**, 807–812 (1982).
12. Engelberg-Kulka, H., Amitai, S., Kolodkin-Gal, I. & Hazan, R. Bacterial programmed cell death and multicellular behavior in bacteria. *PLoS Genet.* **2**, e135 (2006).
13. Garcia-Pino, A. *et al.* Doc of prophage P1 is inhibited by its antitoxin partner Phd through fold complementation. *J. Biol. Chem.* **283**, 30821–30827 (2008).
14. Kinch, L. N., Yarbrough, M. L., Orth, K. & Grishin, N. V. Fido, a novel AMPylation domain common to Fic, Doc, and AvrB. *PLoS ONE* **4**, e5818 (2009).
15. Das, D. *et al.* Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gi|24375750) from *Shewanella oneidensis* MR-1 at 1.6 Å resolution. *Proteins* **75**, 264–271 (2009).
16. Mukherjee, S. *et al.* Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477**, 103–106 (2011).
17. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).

Author Contributions P.E., F.V.S. and A.H. cloned recombinant plasmids. P.E. discovered and physiologically characterized VbhT–VbhA as a toxin–antitoxin module and carried out the bioinformatic analysis. A.G. expressed, purified and crystallized VbhA–VbhT(FIC), NmFic(SE/AA) and NmFic(Δ8), and determined their structures. F.V.S. expressed, purified and crystallized NmFic with AMPPNP and determined the structure. A.G. and A.H. performed the adenylylation assays. A.H carried out the growth curve experiments. A.S. conducted the mass spectrometry analysis. All authors contributed to experimental design and data analysis. The manuscript was written by P.E., A.G., T.S. and C.D.

Author Information The atomic coordinates of VbhA–VbhT(FIC) and the complexes of NmFic, NmFic(SE/AA) and NmFic(Δ8) with AMPPNP have been deposited in the Protein Data Bank under accession codes 3SHG, 3S6A, 3SN9 and 3SE5, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.D. (Christoph.Dehio@unibas.ch) or T.S. (Tilman.Schirmer@unibas.ch).

72

METHODS

Identification of VbhA homologues. VbhA of *B. schoenbuchensis* was queried against a database composed of translated open reading frames (>10 amino acids) identified in the 500-bp upstream region of all *fic* loci (PFAM release 24, 2,189 proteins). Nine rounds of Psi-Blast were performed and hits with an *E*-value <1 manually validated.

E. coli toxicity tests and cell filamentation. *E. coli* AB472, a derivative of MG1655, was transformed with VbhT-expressing plasmids and always handled in LB containing 1% glucose. A single colony was picked, resuspended in 20 μl of LB and plated on LB plates containing 100 μM IPTG (induction) or 1% glucose (repression). Plates were incubated overnight at 37 °C. Growth curves were acquired by measuring optical density of liquid cultures in LB containing 100 μM IPTG that had been inoculated from overnight cultures of single colonies and were grown at 30 °C with continuous shaking. Cell filamentation was visualized by co-transformation of plasmid pC10E that constitutively expresses GFP and examined using fluorescence microscopy.

Conjugation experiments for VbhT toxicity tests in Bartonella. Plasmids were introduced into *Bartonella* strains by conjugation from *E. coli* using three-parental mating. *Bartonella* strains were grown for 36–48 h at 35 °C with 5% CO$_2$ on Columbia base agar plates supplemented with 5% defibrinated sheep blood and 100 μg ml^{-1} streptomycin. *E. coli* β2150 that harbours helper plasmid pRK2013, and *E. coli* β2150 that contains the VbhT-expressing plasmid were grown overnight at 37 °C in LB supplemented with 50 μg ml^{-1} kanamycin or 30 μg ml^{-1} chloramphenicol, respectively, and both media were also supplemented with diaminopimelic acid (DAP) and 1% glucose. After 16 h of incubation, *E. coli* strains were diluted 1:50 in fresh LB medium and grown to a optical density (OD) at 595 nm (OD$_{595nm}$) of 0.4–0.8. Subsequently, each *E. coli* strain was diluted to OD$_{595nm}$ of 0.25, washed once and resuspended in supplemented M199 medium (containing 10% FCS and 1% glucose). *Bartonella* strains were collected in 1 ml and resuspended in 60 μl of M199 (OD$_{595nm}$ = 1). Each *Bartonella* suspension was mixed with 20 μl of *E. coli* β2150 that harbours pRK2013 and 20 μl of *E. coli* β2150 that harbours the VbhT-expressing plasmid. The conjugation mix was distributed on a conjugation filter on a columbia blood agar (CBA) plate supplemented with 150 μl DAP and 150 μl 1% glucose. After 6 h of incubation under *Bartonella* growth conditions as described before, the bacteria were washed off the filter with 1 ml supplemented M199. Dilutions of bacterial suspensions were plated on lysogeny broth agar (LA) supplemented with DAP, 1% glucose and 30 μg ml^{-1} chloramphenicol for selecting donors, on CBA supplemented with 1.2 μg ml^{-1} chloramphenicol for selecting transconjugants, and on CBA supplemented with 100 μg ml^{-1} streptomycin for selecting recipients. Agar plates were incubated under *Bartonella* growth conditions and colony-forming units were counted after 1 day for donors and after 7 days for recipients and transconjugants.

In vitro adenylylation assay. Adenylylation activity of VbhT was assessed in an assay using lysates of ectopically expressing *E. coli*. Bacterial pellets were resuspended in reaction buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1 mM EGTA, 15 mM MgCl$_2$, 140 μg ml^{-1} RNase A and protease inhibitor cocktail (Roche)) and lysed by sonication. After clearing lysates by centrifugation, supernatants were used for experimentation or stored at −20 °C.

Adenylylation reactions were prepared by supplementing 15 μl supernatant from expression cultures with 10 μCi [α-^{32}P]ATP (Hartmann Analytic) and 25 μl blank *E. coli* supernatant. Adenylylation activity of NmFic and SoFic constructs was assessed by incubating 60 μg purified protein with 10 μCi [α-^{32}P]ATP (Hartmann Analytic) and 25 μl blank *E. coli* supernatant. Reactions were incubated for 1 h at 30 °C, resolved by SDS–PAGE, and adenylylation was probed by autoradiography. For HYPE(FIC) assays, 15 ng of pure protein was incubated with 1.6 μg of purified GTPases.

Protein expression and purification. pFVS0011 vector (encoding VbhA and VbhT(FIC)) was transformed into *E.coli* BL21 (DE3). *E. coli* cultures were grown at 37 °C in LB medium supplemented with 50 μg ml^{-1} of kanamycin to an OD$_{595nm}$ of 0.6 before induction with 0.3 mM IPTG for 16 h at 23 °C. Vectors pFVS0015 (carrying the *NmFic* gene), pFVS0016 (encoding NmFic(Δ8)), pFVS0037 (encoding NmFic(SE/AA)) were transformed into BL21-AI cells. Cells were incubated overnight in 750 ml LB medium that was supplemented with 50 μg ml^{-1} kanamycin and 1% glucose at 23 °C at 200 r.p.m. until OD$_{595nm}$ of 2 was reached. Bacterial pellets were resuspended in terrific broth medium containing 50 μg ml^{-1} kanamycin to obtain an OD$_{595nm}$ of approximately 1. Protein expression was induced at 23 °C with 0.1% arabinose and 0.5 mM IPTG for 23 h at 200 r.p.m. Plasmids harbouring HYPE(FIC) and SoFic constructs were transformed in *E. coli* Rosetta (DE3) cells and BL21-AI cells, respectively. The proteins were expressed as described for NmFic.

Cells containing overexpressed VbhA–VbhT(FIC) and NmFic were resuspended in lysis buffer containing 20 mM Tris (pH 7.5), 250 mM NaCl and 25 mM imidazole and disrupted using French press. Cell debris was pelleted by

ultracentrifugation and the supernatant was applied to a His-Trap column (GE Healthcare). The stable complex was eluted with a gradient of elution buffer containing 20 mM Tris (pH 7.5), 250 mM NaCl and 500 mM imidazole. The protein was then concentrated and injected on a Superdex 75 16/60 gel filtration column (GE Healthcare) equilibrated with 10 mM Tris (pH 7.5) and 100 mM NaCl. The pure proteins were concentrated to 6 mg ml^{-1} for VbhA–VbhT(FIC) and 53 mg ml^{-1} for NmFic.

The same purification protocol was used for NmFic(Δ8) and NmFic (SE/AA), but with an additional intermediate purification step. After affinity purification, the proteins were adjusted to 20 mM Tris (pH 8.5), 25 mM NaCl, applied to a Resource-Q anion exchange column (Amersham Biosciences) and eluted with a linear gradient of NaCl. Peak fractions were concentrated and further purified by gel filtration chromatography. Purified proteins in 10 mM Tris (pH 7.8), 100 mM NaCl were concentrated to 30 mg ml^{-1} for NmFic(Δ8) and 51 mg ml^{-1} for NmFic(SE/AA). SoFic and SoFic(E73G) were purified as described previously[15]. HYPE(FIC), HYPE(FIC/H295A) and HYPE(FIC/E234G) were purified in the same way as NmFic. GST-tagged Cdc42 and Rac1 were expressed and purified as described previously[18,19].

Crystallization. All crystals were obtained at 20 °C (except for NmFic crystals, which were obtained at 4 °C) using the hanging-drop vapour diffusion method after mixing 1 μl protein solution with 1 μl reservoir solution. VbhA–VbhT(FIC) and NmFic(SE/AA) crystallized in 23% (w/v) PEG 3350 and 0.2 M di-ammonium tartrate, and were cryoprotected with 25% (w/v) PEG 3350, 0.2 M di-ammonium tartrate and 10% glycerol. NmFic crystallized in 5% 2-propanol, 0.1 M MES pH 6.0 and 0.1 M Ca-acetate, and was transferred into 8% 2-propanol, 0.1 M MES pH 6.0, 0.1 M Mg-acetate and 15% glycerol, then into 30% glycerol for cryoprotection. NmFic(Δ8) crystallized with 44% (v/v) PEG 600, 0.1 M Na-citrate pH 5.6. No cryoprotection was needed for data collection. In each case, the substrate analogue complex was produced by crystal soaking for 2 h with 10 mM AMPPNP, 10 mM MgCl$_2$.

Structure determination. Statistics of data collection and refinement are given in Supplementary Table 1. Diffraction data were collected at beamline X06SA (PXIII) of the Swiss Light Source ($\lambda = 1.0$ Å) at 100 K on a MAR CCD detector, processed using MOSFLM[20], and scaled with SCALA[21]. The structures were determined by molecular replacement (PHASER[22]) using a BepA fragment (PDB code 2JK8[10], residues 30–194) or the uncomplexed NmFic structure (PDB code 2G03, unpublished, Midwest Center for Structural Genomics) as models for structure solution of VbhA–VbhT(FIC) and the different NmFic proteins, respectively. A structure solution of wild-type and mutant NmFic in complex with AMPPNP was straightforward, whereas a weak solution (RFZ = 5.1, TFZ = 3.6) with poor phasing power was obtained for VbhA–VbhT(FIC). The partial model lacking VbhA was refined by rigid-body refinement using REFMAC5 (ref. 23) with three bodies to an R_{free} of 52%. Model extension using the module AutoBuild of the PHENIX package[24] yielded an almost complete model ($R/R_{free} = 30.8\%/35.2\%$). The remainder of the molecule was traced manually with COOT[25] and then by full refinement using PHENIX[24]. The Ramachandran plot showed that more than 99% of the residues are in favoured regions of the four structures. The figures were generated with Dino (http://www.dino3d.org).

Prediction of inhibition motif in Fic proteins. All Fic proteins (PFAM release 24, 2,189 proteins) were subjected to a profile-to-profile comparison with sequences from the PDB using HHpred[17]. HHpred builds an alignment of homologues for each query sequence by using iterations of PSI-BLAST searches against the non-redundant database. Secondary structures are then predicted on the PSI-BLAST alignment using PSIPRED[26]. On the basis of this data, a profile Hidden Markov Model (HMM) is generated. Each query profile is compared with the precomputed HMMs of the proteins in the PDB to identify structural homologues. In terms of query profiles, the PDB profiles include secondary structure information derived from their three-dimensional structure. We analysed the pairwise profile alignments of each Fic protein with the eight different Fic family members deposited in the PDB. A Fic protein was predicted to belong to class II or class III if the templates' inhibitory motifs were aligned to a corresponding query sequence in the profile alignments (see Supplementary Information, section 1).

Phylogenetic analysis of Fic proteins. Phylogenetic trees of the Fic family were inferred with FastTree 2 (ref. 27) and RAxML 7.0.4 (ref. 28). Trees were built on the amino acid alignment provided by the PFAM database. Unaligned overhanging ends were trimmed off and identical sequences were reduced to one representative. We used local support values based on the Shimodaira–Hasegawa test to estimate the reliability of the tree inferred with FastTree2. The RaxML tree was inferred using the PROTMIXWAGF model and 25 rate categories.

Liquid chromatography–mass spectrometry analysis. 2.5 μM of purified NmFic, NmFic(Δ8) or NmFic(SE/AA) were incubated in reaction buffer (10 mM Tris, pH 8, 100 mM NaCl) in the presence or absence of 50 μM ATP and 50 μM MgCl$_2$ for 1 h. Proteins were reduced in 5 mM TCEP, alkylated in

10 mM iodacetamid and digested with sequencing grade trypsin (Promega). The generated peptides were purified with C18 Microspin columns (Harvard Apparatus) and analysed using liquid chromatography–mass spectrometry (LC–MS) or MS on an easy nano-LC system coupled to an LTQ-Orbitrap-Velos mass spectrometer (both from Thermo-Fisher Scientific), as recently described[29] using a linear gradient from 95% solvent A (0.15% formic acid, 2% acetonitrile) and 5% solvent B (98% acetonitrile, 0.15% formic acid) to 35% solvent B over 40 min. The data acquisition mode was set to obtain one high-resolution MS scan in the Fourier Transform (FT) part of the mass spectrometer at a resolution of 60,000 (full width at half maximum) and MS–MS scans in the linear ion trap of the 20 most intense ions. The resulting MS2 scans were searched against a *N. meningitidis* protein database containing the target protein sequence, including NmFic and NmFic(SE/AA) sequences, that was obtained from EBI (http://www.ebi.ac.uk) using the SEQUEST search algorithm provided in the Proteome Discoverer software package (Thermo-Fisher Scientific). *In silico* trypsin digestion was performed after lysine and arginine (unless followed by proline), with a tolerance of two missed cleavages in fully tryptic peptides. Database search parameters were set to allow phosphoadenosine modification (+329.05252 Da) of threonine and tyrosine residues as variable modification and carboxyamidomethylation (+57.021464 Da) of cysteine residues as fixed modification. The fragment mass tolerance was set to 0.8 Da and the precursor mass tolerance to 15 p.p.m.

Strain construction. For toxicity experiments, the *vbhT* wild-type gene (FN645515) from *B. schoenbuchensis* R1 was cloned into pRSF-Duet1 (pPE0017, His$_6$-tagged *vbhT*). VbhT(H136A) (pPE0034) was constructed by introducing a two-base-pair point mutation in the FIC motif of *vbhT* of pPE0017, as described elsewhere[30]. Plasmid co-expressing VbhT and VbhA (VbhT/VbhA) was constructed by cloning *vbhA* (FN645515) amplified from *B. schoenbuchensis* R1 into pRSF-Duet1 (pPE0020, HA-tagged *vbhA*). *vbhT* was then cloned into pPE0020, resulting in pPE0021. To construct VbhT-expressing plasmids for *Bartonella*, *vbhT* from *B. schoenbuchensis* R1 was cloned into vector pMMB206 (ref. 31) (pVbhT, HA-tagged *vbhT*). pVbhT(H136A) was constructed from pVbhT as described before.

The in-frame deletion of the complete *vbhA/vbhT* operon in *B. schoenbuchensis* (*Bsch ΔvbhA/vbhT*) was generated as described previously by a two-step gene replacement procedure[9]. The mutagenesis vector pPE3005 was constructed by ligating a cassette with the flanking regions of the in-frame deletion into pTR1000[9].

For protein purification, the full-length *vbhA* gene and part of the *vbhT* gene encoding the FIC domain (*vbhT(FIC)*, amino acid residues 1–198, His$_6$-tagged) were amplified from plasmid pPE0021 and cloned into the pRSF-Duet1 vector (pFVS0011). VbhA(E24G)/VbhT(FIC) expression plasmid (pFVS0065) was generated by introducing a two-base-pair mutation in pFVS0011. The *NmFic* gene was amplified with an N-terminal His$_6$-tag from *N. meningitidis* from coding region of amino acid residues 11–191 and from coding region of amino acid residues 11–167 to generate plasmids expressing NmFic (pFVS0015) and NmFic(Δ8) (pFVS0016), respectively. The S182A/E186A double mutant construct (NmFic(SE/AA)) was generated by introducing two subsequent point mutations in pFVS0015. The E186G mutant construct (NmFic(E186G)) was

generated by the same approach. The *SoFic* gene was amplified from plasmid (ASU biodesign institute, Clone ID SoCD00104192) and cloned with an N-terminal His$_6$-tag into pRSF-Duet1 (pFVS0040). The SoFic(E73G) plasmid (pFVS0058) was generated by introducing a two-base-pair point mutation in pFVS0040. GST–HYPE(E234G) (pFVS0064) was generated by introducing point mutations in the plasmid containing GST–HYPE. From these plasmids, shorter constructs (HYPE(FIC), HYPE(FIC/E234G) and HYPE(FIC/H295A)) only carrying the FIC domain of HYPE (from amino acids 187 to 437) were generated. For the expression of human Cdc42 and Rac1, the *Cdc42–Q61L* and *Rac1–Q61L* coding sequences were amplified from plasmid pRK5myc L61 Cdc42[32] and pRK5FLAG L61 Rac1 (ref. 32) and cloned into pGex6p1 with an N-terminal GST-tag, resulting in pAH088 and pAH060, respectively. The wild-type variants of Cdc42 and Rac1 were generated from the mutant constructs through polymerase chain reaction (PCR)-based site-directed mutagenesis[30] (resulting in pAH059 and pAH071). All primers and the resulting vectors are summarized in Supplementary Tables 5 and 6.

18. Self, A. J. & Hall, A. Purification of recombinant Rho/Rac/G25K from *Escherichia coli*. *Methods Enzymol.* **256**, 3–10 (1995).
19. Smith, S. J. & Rittinger, K. Preparation of GTPases for structural and biophysical analysis. *Methods Mol. Biol.* **189**, 13–24 (2002).
20. Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62**, 48–57 (2006).
21. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
22. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
23. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
24. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
25. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
26. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
27. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
28. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
29. Schmidt, A. *et al.* Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Mol. Syst. Biol.* **7**, 510 (2011).
30. Zheng, L., Baumann, U. & Reymond, J. L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, e115 (2004).
31. Dehio, C. & Meyer, M. Maintenance of broad-host-range incompatibility group P and group Q plasmids and transposition of Tn5 in *Bartonella henselae* following conjugal plasmid transfer from *Escherichia coli*. *J. Bacteriol.* **179**, 538–540 (1997).
32. Rhomberg, T. A., Truttmann, M. C., Guye, P., Ellner, Y. & Dehio, C. A translocated protein of *Bartonella henselae* interferes with endocytic uptake of individual bacteria and triggers uptake of large bacterial aggregates via the invasome. *Cell. Microbiol.* **11**, 927–945 (2009).

Content:

1. Supplementary Text

1.1 Defining an adenylylation competent FIC signature motif

Based on sequence conservation, previous mutagenesis results, and the structure of the competent NmFic(Δ8)/AMPPNP complex (Supplementary Fig. 10b), we have defined the HxFx[D/E]GNGRxxR profile as the crucial core motif for catalytically active Fic proteins with respect to adenylylation. The histidine residue (H) has been shown in several studies[1,2] to be essential for activity and probably acts as a general base for deprotonation of the target side-chain hydroxyl group[2,3]. The phenylalanine (F) is part of the hydrophobic core and important for correct positioning of the active loop. It remains unclear whether other hydrophobic residues could replace the phenylalanine. Strikingly, all Fic proteins so far shown to exhibit adenylylation activity harbour a phenylalanine at this position. The carboxylic side-chain of the aspartic/glutamic acid (D/E) probably coordinates the divalent cation (Mg^{2+} or Mn^{2+}) that is required for activity (data not shown). However, in the competent substrate complex structure (Supplementary Fig. 10b) it is partly disordered. As predicted before[2] and confirmed by the competent AMPPNP complex structure (Supplementary Fig. 10b), the three backbone amide nitrogens of GNG are involved in α- and β- phosphate binding. No other residue than a glycine is allowed for the two G positions, since side-chains would clash with the core of the protein or the ATP substrate, respectively. The asparagine side-chain is involved in positioning of the α-phosphate and establishes two H-bonds with the backbone amide and carbonyl group of the active site phenylalanine residue. Finally, the two arginines are forming a salt-bridge with the β- and γ-phosphate, respectively (Supplementary Fig. 10b). For all residues of the motif, single point mutants have been constructed except for the first glycine and first arginine that showed reduced or negligible auto-adenylylation and/or target adenylylation[1-5], thereby highlighting the

importance of these conserved residues. Thus, based on current structural and biochemical knowledge, we can assume that Fic proteins with a motif conforming to the consensus HxFx[D/E]GNGRxxR are likely to represent AMP transferases. In contrast, proteins divergent from this motif have most probably lost adenylylation activity.

1.2 Prediction and analysis of the three regulatory classes of Fic proteins

While class I proteins were identified by Psi-Blast for the detection of VbhA-homologous peptides encoded in the upstream region of *fic* loci, we used a structure based approach to detect class II and class III Fic proteins. The program HHpred predicts structures of distant homologs from proteins deposited in structure databases[6]. The analysis is based on a profile-to-profile comparison of Hidden Markov Models (HMMs) generated from Psi-Blast and secondary structure predictions (see Online Methods). We conducted an HHpred analysis for each protein in the FIC family (PFAM release 24, 2,189 proteins) using the protein structure database PDB. The PDB contains structures of FIC domains from eight different proteins: (i) 3dd7, DocH66Y of prophage P1 in complex with the C-terminal domain of Phd; (ii) 2jk8, BepA of *Bartonella henselae,* a type IV secretion system effector protein; (iii) 3n3u, IbpA of *Histophilus somni*, a protein secreted into host cells; (iv) 3let, VopS of *Vibrio parahaemolyticus*, a Type III secretion system effector protein; (v) 2f6s, Fic protein of *Helicobacter pylori* (HpFic); (vi) 2g03, Fic protein of *Neisseria meningitidis* (NmFic); (vii) 3eqx, Fic protein of *Shewanella oneidensis* (SoFic); (viii) 3cuc, Fic protein of *Bacteroides thetaiotaomicron* (BtFic)*.* For all Fic proteins analyzed with HHpred, we detected significant structural homology with these eight FIC domain structures in the PDB. Based on the corresponding HHpred alignments, we validated each Fic protein for the presence of an N-terminal (class II) or C-terminal (class III) inhibition motif. Fic proteins were predicted to fall into one of these two classes, if the structural alignments with the two structural prototypes of

a class (for class II, SoFic and BtFic; for class III, NmFic and HpFic) included the region comprising their inhibition motif. The segment of the query sequence aligning with the inhibition motifs of a given class was then identified as the putative inhibition motif of the query protein. Only proteins for which both alignments predicted the same inhibition motif were categorized into the corresponding class.

Supplementary Tables 4a-e show that for most proteins, predicted to fall into one of the two classes, the corresponding template structures resulted in the highest homology scores among the eight Fic proteins in our HHpred analysis. Further, the inhibition motifs independently predicted from the two template structures, turned out to be the same for most Fic proteins. Finally, only four proteins were identified to belong to more than one of the three classes. Altogether, this demonstrates the high specificity of our analysis for the prediction of these regulatory features.

Strikingly, the consensus sequence of all motifs from class I, II, and III converges towards the general inhibition motif [S/T]xxxE[G/N]. The glutamate (E24 of VbhA), thereby, displays the most conserved amino acid, which is consistent with the presented mechanism of inhibition (Fig. 3).

As mentioned in the main text, in comparison to the entire family, the total number of proteins predicted to fall into one of the three regulatory classes is significantly increased in the subset of Fic proteins harbouring the adenylylation-competent FIC motif HxFx[D/E]GNGRxxR (see Supplementary Text 1.1). Thus, the inhibition motif seems to be a specific feature of proteins with adenylylation activity and indicates co-evolution of these two functional features. In contrast to this general trend, the relative number of class I proteins (regulated by an antitoxin-like protein) decreases in the subset of Fic proteins with a conserved FIC motif. This can be explained by the fact that a significant fraction of class I comprises proteins from *E. coli* and related species which harbour a serine instead of an

asparagine and a leucine instead of arginine (1st position) in the active loop. It remains to be elucidated whether HxFx[D/E]GSGLxxR either can still mediate adenylylation or has evolved another activity which is regulated by a similar mechanism as adenylylation.

Interestingly, class I and III are restricted to bacteria (which present 90% of all Fic proteins), while in archaea and eukaryotes only class II inhibition motifs were predicted (Supplementary Table 2). Most of the metazoan Fic proteins including the human homolog HYPE as well as some of the fungal Fic proteins are predicted to contain a class II regulatory motif suggesting that this regulatory mechanism of adenylylation activity is conserved in all domains of life.

1.3 Prediction of regulatory classes for selected Fic proteins

The classification based on the presence/absence of an inhibition motif is listed in Supplementary Table 3 for a subset of Fic proteins which were previously studied. For VopS, the effector protein of *Vibrio parahaemolyticus*, our analysis did not predict an inhibition motif. This is in agreement with the 3D structure of VopS[3], which does not reveal any regulatory helix positioned next to the FIC active site.

In case of the second bacterial virulence factor for which adenylylation activity in eukaryotic host cells was described, IbpA[4], our analysis predicted an N-terminal inhibition motif (class II). However, the 3D structure of IbpA does not show this predicted inhibition motif to be located in the active center of the FIC domain[1]. Thus, our prediction for IbpA may be a false positive. Both VopS and IbpA are translocated into eukaryotic host cells, where they adenylylate proteins of the Rho family GTPases. This leads to the collapse of the actin cytoskeleton and subsequently to cell death[4,5]. In general, effector proteins of bacterial pathogens that aim at killing or dramatically subverting the cells of their eukaryotic hosts (e.g.

immune cells) by adenylylation of cellular targets might not need to be regulated. Further, due to their specificity for eukaryotic targets, they may not be toxic for the bacterium before translocation and would thus not require an inhibited state.

However, for effector proteins of *Bartonella,* our Psi-Blast analysis identified homologs of VbhA located upstream of the effector gene loci. In contrast to VopS and IbpA, the effector proteins of *Bartonella* seem to have a much more subtle effect on host cells[7]. It remains to be elucidated whether for *Bartonella* effector proteins co-translocation of a VbhA-homolog into host cells allows regulation of adenylylation activity.

Like the bacterial virulence factors IbpA or VopS, the human Fic protein HYPE, was shown to adenylylate the GTPases RhoA, Rac1, and Cdc42[4]. The conserved N-terminal inhibition motif identified for HYPE and other mammalian Fic proteins might explain why HYPE, in contrast to IbpA or VopS, did not show any cytotoxic effect on mammalian cells[4]. Thus, the here-identified regulatory mechanism might be essential for the tight regulation of mammalian Fic proteins and their integration into key signaling pathways of mammals.

Doc proteins share the same structural fold as Fic proteins and are thus grouped into the FIC PFAM family (which is also referred to as Fic/Doc family). They constitute the toxins of toxin-antitoxin modules homologous to the post-segregational killing system Doc/Phd of Bacteriophage P1[8]. In our phylogenetic analysis (Supplementary Fig. 7), Doc proteins form a separate cluster clearly distinct from the remaining Fic proteins. Furthermore, their core motif (Supplementary Table 3) is highly divergent from the adenylylation competent motif as defined above (see Supplementary Text 1.1). Consistently, our HHpred and Psi-Blast analyses did not identify any adenylylation-inhibiting motif for Doc-related proteins. Nevertheless, Doc activity is inhibited by the antitoxin Phd which binds in a similar way as VbhA to VbhT (Supplementary Fig. 5). Doc-mediated toxicity on bacteria results from binding to the 30S ribosomal subunit, but the exact underlying mechanism of Doc-mediated toxicity remains

elusive and adenylylation activity has not been reported[9]. Still, inhibition of Doc toxicity by mutation of the histidine in the central motif of the FIC domain indicates its involvement in mediating the toxic effect[10]. For Doc as well as for other Fic proteins carrying a motif divergent from the adenylylation-competent consensus profile, the FIC domain could mediate other enzymatic activities[11]. Catalytic versatility of the FIC domain fold would further explain why this protein family is so abundant in a wide range of organisms.

1.4 Mass spectrometry (MS) results for the identification of the auto-adenylylation site of NmFic

In order to specifically pinpoint the location of the adenylylation in the FIC domain and to verify the results obtained for the adenylylation assay (Fig. 3d), a mass spectrometric analysis of the trypsinized NmFic protein samples was performed. As apparent from Supplementary Fig. 8a, the majority of the adenylylation could be identified at position Y6 in the displayed peptide sequence. Additional conformation was obtained by the identification of the same adenylylation site on the same peptide containing the C-terminal glycine as consequence of incomplete proteolysis (Supplementary Fig. 8b). However, due to the close proximity of the tyrosine residues, it cannot be excluded that Y7 or Y8 are also modified. Comparative analysis of the NmFic and NmFic(Δ8) mutant clearly indicates a strong increase in adenylylation for this domain (Supplementary Fig. 8c). Notably, we covered the whole protein sequence by MS analysis and could only identify adenylylation at the C-terminal tyrosine residues of this protein, indicating exclusive adenylylation of this protein at these sites and confirming the results obtained in Fig. 3d.

MS analysis of the NmFic(SE/AA) mutant revealed a higher degree of adenylylation, which is indicated by the identification of more and multiple adenylylated peptide species (Supplementary Fig. 9), again, confirming the observations made in Fig. 3d.

2. Supplementary Figures
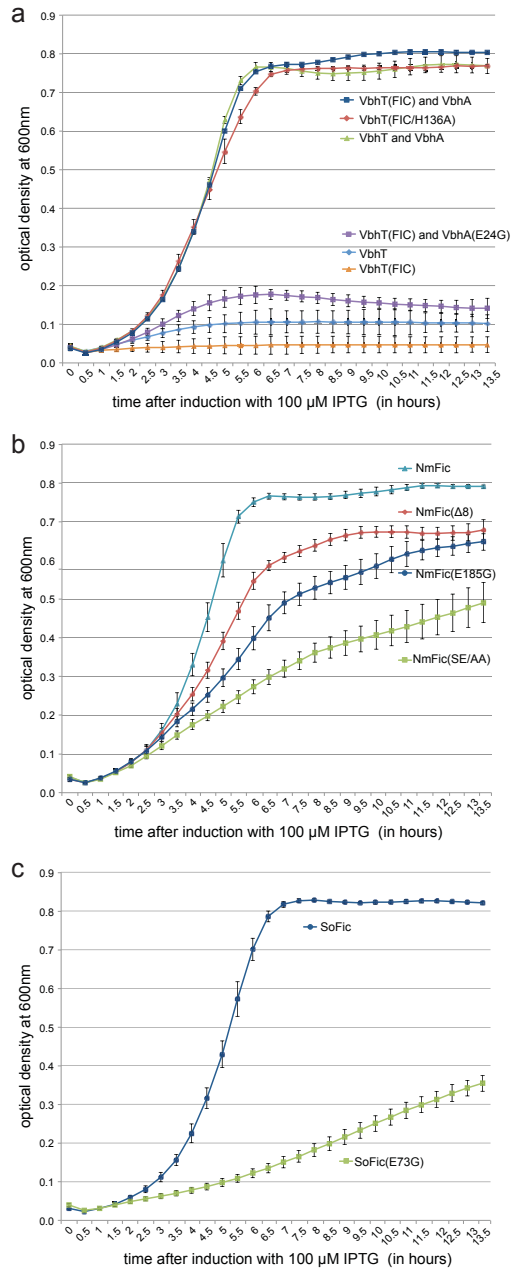
Supplementary Figure 1. Growth curves of *E. coli* MG1655 expressing different constructs of (a) VbhA-VbhT, (b) NmFic, or (c) SoFic. Cultures were inoculated with 5 µl

of an overnight culture and induced with 100 μM IPTG. Bacterial growth was measured by monitoring the optical density at 600 nm every 30 minutes. All experiments were carried out in quadruplicates.

| empty vector | VbhT | VbhA-VbhT | VbhT(H136A) |
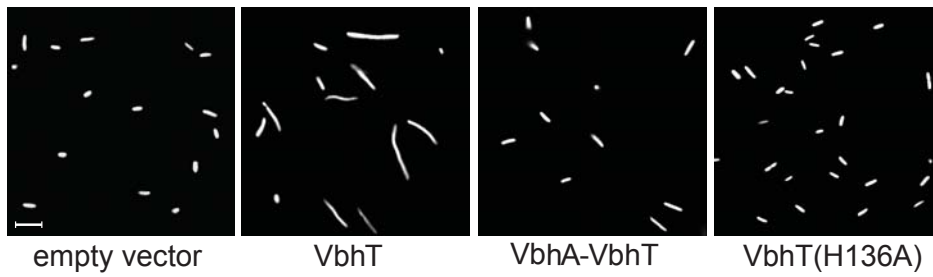
Supplementary Figure 2. Cell filamentation of *E. coli* upon expression of VbhT is repressed by co-expression of VbhA. Fluorescence microscopy of *E. coli* harbouring a GFP-reporter plasmid (pC10E) in addition to the same VbhA-VbhT-expression plasmids as presented in Fig. 1c (scale bar, 5 μm). Cell filamentation of *E. coli* observed upon VbhT expression is not evident for VbhT(H136A) or VbhA-VbhT.
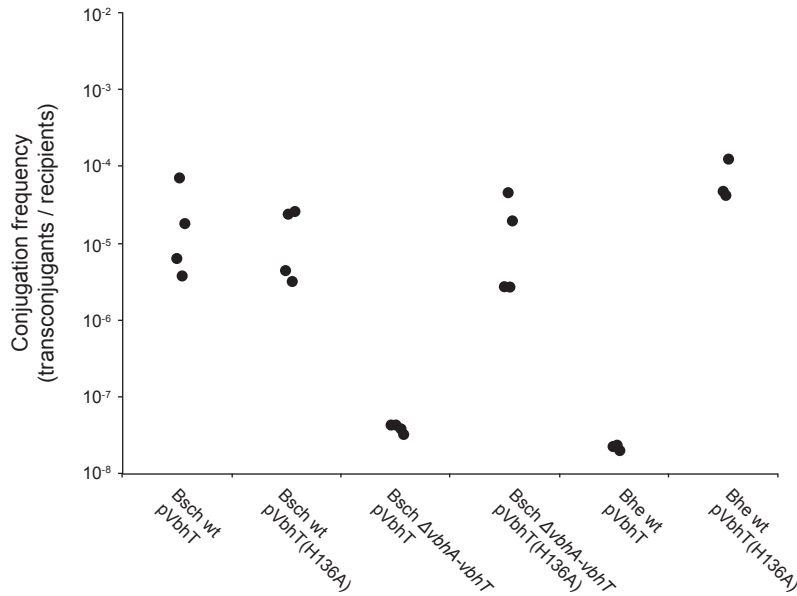
Supplementary Figure 3. Conjugation of VbhT expression plasmids into *Bartonella schoenbuchensis* and *Bartonella henselae* reveals toxicity of VbhT in the absence of vbhA. Conjugation frequencies are given by the number of transconjugants per recipient as determined by dilution plating and counting of colony forming units on media selective for either transconjugants or recipients. Conjugation of a plasmid expressing VbhT wild-type (pVbhT) into *B. schoenbuchensis* wild-type (Bsch wt) resulted in a conjugation frequency of about 10^{-5}. In contrast, conjugation of the same plasmid into *B. schoenbuchensis* harbouring a deletion of the complete *vbhA-vbhT* operon (Bsch Δ*vbhA-vbhT*) or into *B. henselae* wild-type (Bhe wt), which does not encode *vbhA,* gave no transconjugants in four and three independent experiments, respectively. This resulted in conjugation frequencies $< 10^{-7}$ (when taking the detection threshold in our experiments into account). The inability to conjugate the plasmid into these strains was dependent on a functional FIC motif: a plasmid expressing VbhT with an adenylylation-deficient FIC motif (pVbhT(H136A)) could be conjugated into Bsch Δ*vbhA-vbhT* and Bhe wt with a similar frequency as into Bsch wt.

Supplementary Figure 4. Protein alignment of translated ORFs with homology to VbhA encoded in the upstream region *of fic* gene loci. We used Psi-Blast to query the VbhA protein sequence against all translated ORFs (>10 aa) found in the 500 bp upstream of loci

encoding Fic proteins. In total, this analysis identified 158 ORFs with homology to VbhA. 127 ORFs are overlapping with the corresponding *fic* gene, together constituting putative operons. From the remaining 31 ORFs, 20 are less than 100 bp away from the annotated start codon of the *fic* gene. In the depicted alignment, sequences identical to each other were reduced to one representative resulting in 87 aligned sequences including VbhA. Overhanging ends were trimmed. Start and end positions of the aligned sequences are indicated. The consensus sequence logo is depicted on top of the alignment indicating the conserved central motif. The alignment was generated with MUSCLE[12] (implemented in Geneious v5.3.6) and manually curated. Shading of amino acids indicates degree of conservation. Sequences used in the alignment of Fig. 1b are highlighted in red color.
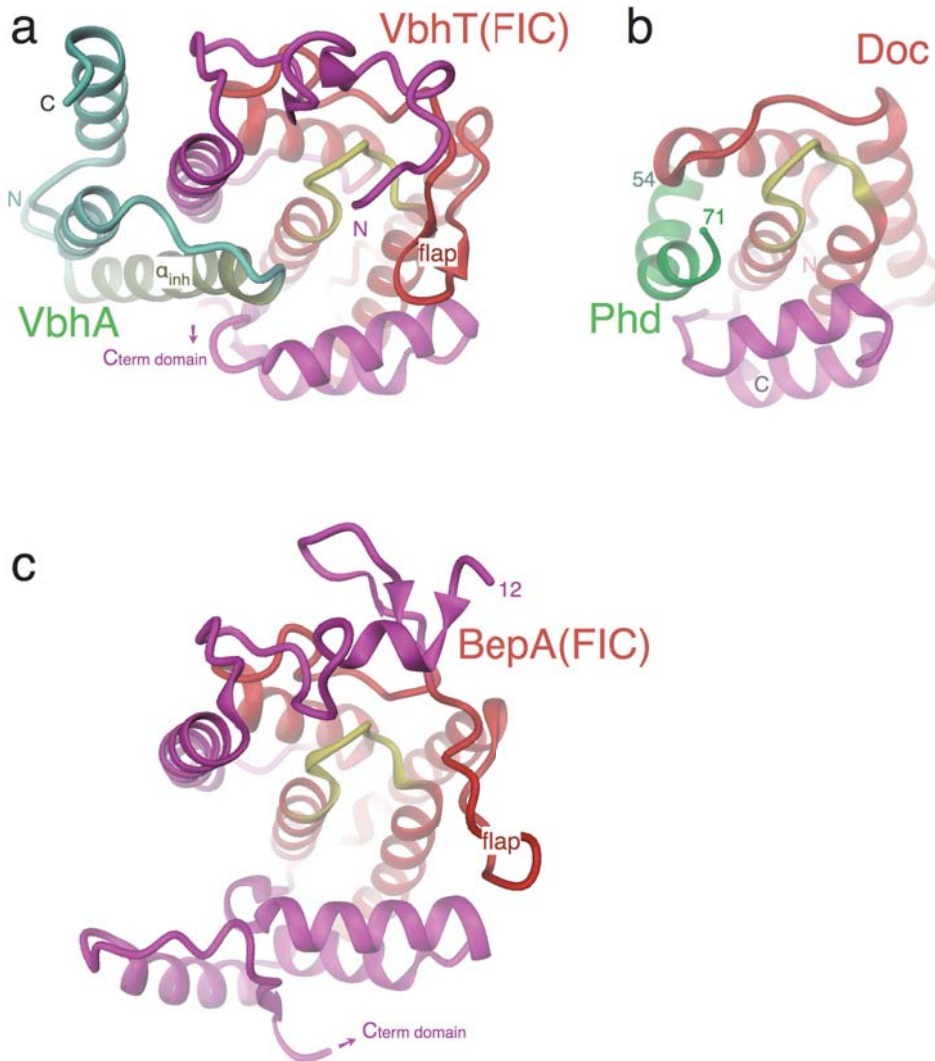
Supplementary Figure 5. Comparison of FIC domain folds and FIC/antitoxin interactions. a, VbhA-VbhT(FIC) from *B. schoenbuchensis*. The VbhT(FIC) structure is highly similar to that of BepA(FIC)[2] (rmsd of 1.6 Å for 169 C_α-positions, panel c). **b**, Doc/Phd from Bacteriophage P1[13], PDB code 3dd7). **c**, BepA(FIC) from *B. henselae*[2]. Representations of the structures as in Fig. 2.
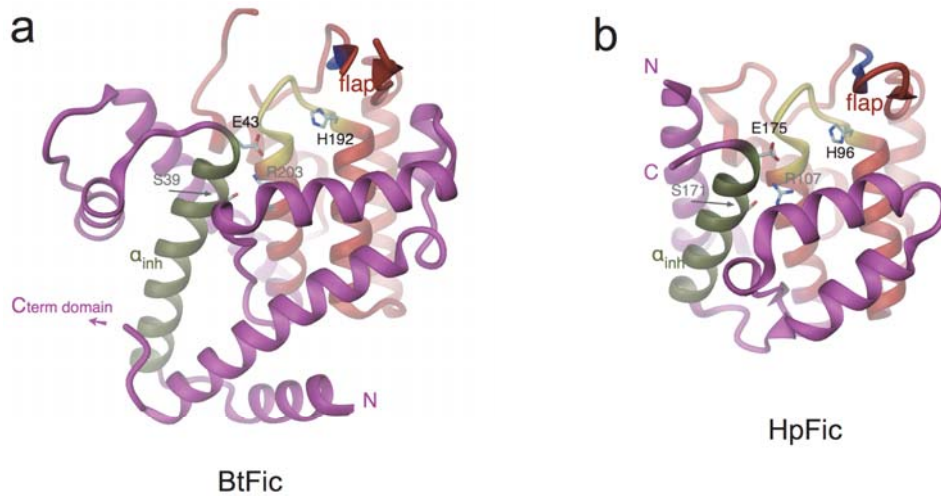
Supplementary Figure 6. Fic protein structures BtFic (BT2513 from *Bacteroides thetaiotaomicron*, PDB code 3CUC) (a) and HpFic (HP1159 from *Helicobacter pylori*, PDB code 2F6S) (b). Structures are shown in cartoon representation as in Fig. 2 (Fic core as defined by PFAM, red; active site loop with histidine, yellow; inhibitory helix (α_{inh}) with C-terminal SxxxEG motif, green; termini labeled with N and C) with important residues in full. Note that BtFic and HpFic are class II and III Fic proteins, respectively, and are close structural homologs of SoFic(Fig. 2b) and NmFic (Fig. 2c).
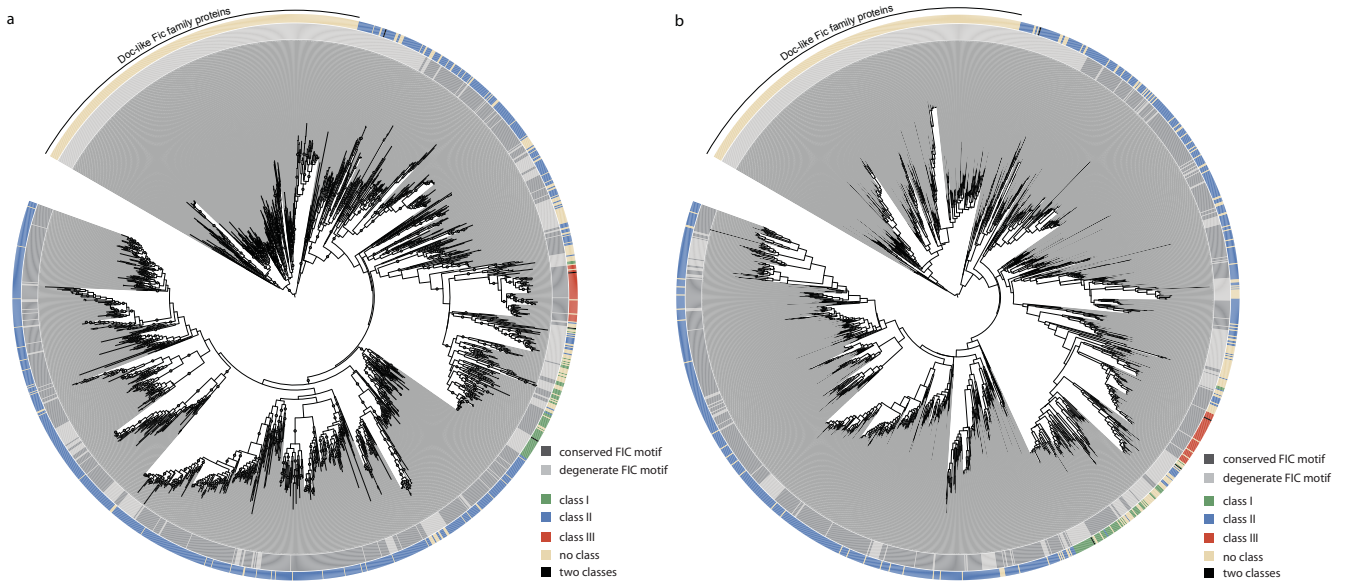
Supplementary Figure 7. **Phylogenetic trees of Fic proteins.** a, Approximately-maximum-likelihood tree inferred with FastTree2.

Local support values were calculated as implemented in FastTree2[14] using the Shimodaira-Hasegawa test. Circles indicate branches

with local support values > 0.9. b, Maximum-likelihood tree inferred with RAxML[15] using the PROTMIXWAG model. Dark and light

grey coloring on the inner circle indicates the presence (conserved) or absence (degenerate) of the adenylylation competent motif,

respectively. The outer circle depicts the class assigned to each Fic protein conforming the coloring of Fig. 2e. The large cluster of Doc-

like proteins lacking the conserved FIC motif is indicated. An interactive version of the tree with species and accession number labels

for each leaf can be accessed via the following link http://itol.embl.de/shared/Engel2011.
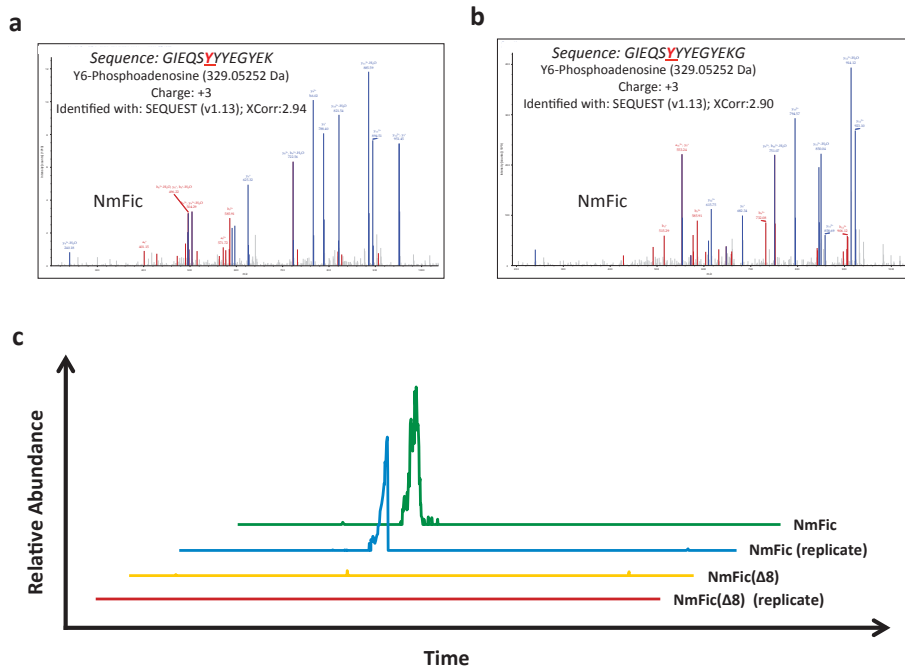
Supplementary Figure 8. LC-MS/MS analysis of adenylylated NmFic protein. a, MS/MS-spectrum of the fully cleaved peptide GIEQSYYYEGYEK at a charge state of 3 using CID. The identified adenylylation sites are underlined in the peptide sequence and indicated in red. All correctly assigned y- and b- fragment ions are indicated in blue and red, respectively, unassigned fragment ions are shown in gray. **b,** MS/MS-spectrum of the incompletely cleaved peptide GIEQSYYYEGYEKG showing the same modification. **c,** Extracted ion chromatograms of the fully cleaved peptide displayed in panel a for the different samples analyzed.
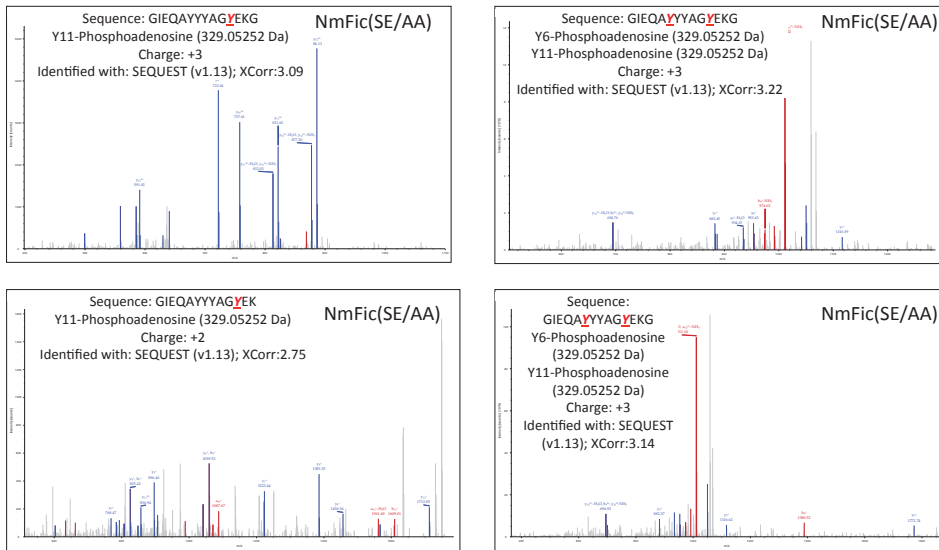
Supplementary Figure 9. LC-MS/MS analysis of adenylylated NmFic(SE/AA) protein.
Four MS/MS-spectra are displayed showing the most significant adenylylated peptide
sequences identified in the NmFic(SE/AA) mutant. The identified adenylylation sites are
underlined in the peptide sequence and indicated in red. All correctly assigned y- and b-
fragment ions are indicated in blue and red, respectively, unassigned fragment ions are shown
in gray.
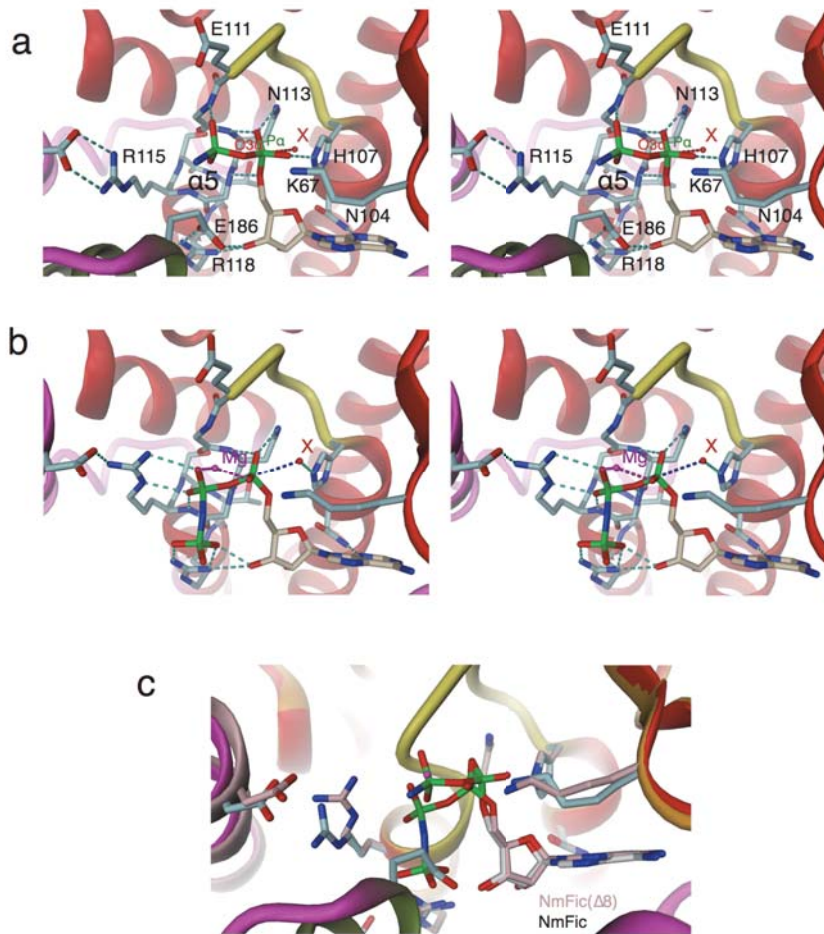
Supplementary Figure 10. AMPPNP binding to NmFic and NmFic(Δ8). The stereoviews of NmFic (a) and NmFic(Δ8) (b) show the detailed interactions of the nucleotide with the Fic active site with H-bonds marked by stippled lines in aquamarine. The structure represenations show similar views as in Figs 3a and b, but with the N-terminal end of helix α5 shown in full to visualize the capping interactions of the α- and β-phosphate with the helix. Additionally, the position that would have to be adopted by a target hydroxyl group for productive nucleophilic attack inline with the scissile Pα-O3α bond is indicated by a small red sphere (marked X) with the X - Pα distance set to 3.0 Å. (c) Superposition of NmFic (gray carbons) and NmFic(Δ8) (orange carbons).
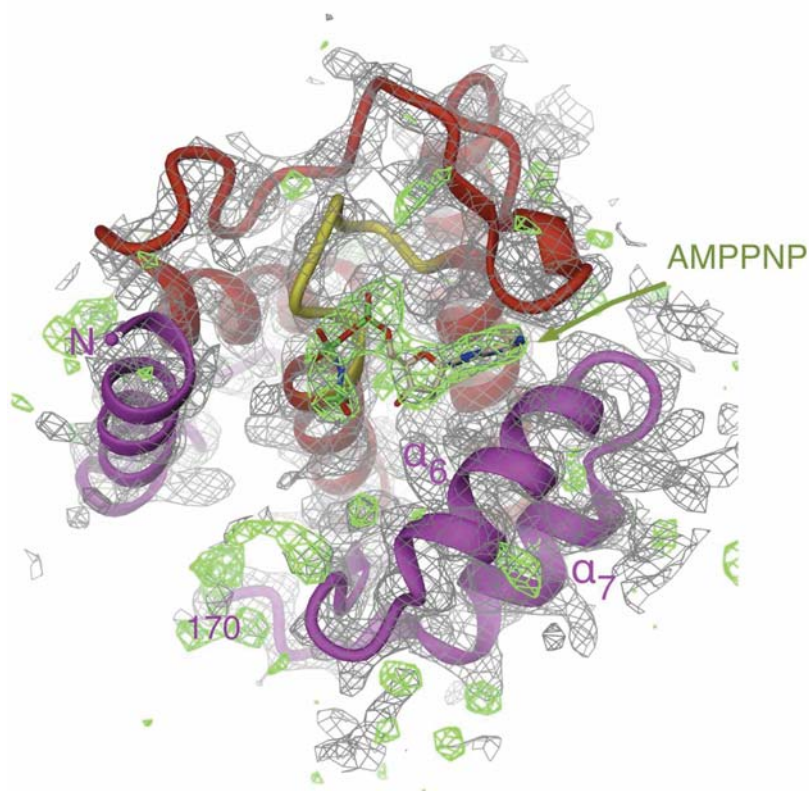
Supplementary Figure 11. Crystal structure of NmFic(SE/AA). The schematic cartoon of

chain A is shown in the same representation as in Fig. 2a. Note that (as in all other 15

molecules of the asymmetric unit) helix α_{inh} which follows α7 is disordered as evidenced by

missing electron density. The AMPPNP substrate analog, however, is fully resolved and

found in the same conformation and orientation as in the corresponding

NmFic(Δ8)/AMPPNP structure (Fig. 3b). The ligand has been omitted from phasing. The

2Fo-Fc density map is shown in grey (1.2 σ), the Fo-Fc map in green (3.0 σ).
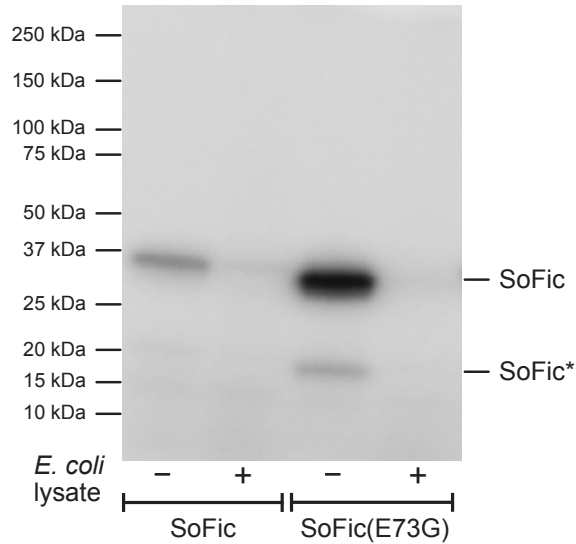
Supplementary Figure 12. Autoradiography of an SDS-gel after incubation of SoFic and SoFic(E73G) with α-^{32}P-ATP in the presence or absence of *E. coli* lysate. Auto-adenylylation is drastically increased in the SoFic(E73G) mutant. SoFic* indicates a truncated version of SoFic as evidenced by mass spectrometry. No target bands are revealed upon incubation with *E. coli* lysate. It is conceivable that the target is a small, unresolved peptide or DNA.

3. Supplementary Tables

Supplementary Table 1. Data collection and refinement statistics.

	VbhA–VbhT(FIC)	NmFic	NmFic(Δ8)	NmFic(SE/AA)
Data collection				
Space group	C2	P6$_4$22	P2$_1$	P2$_1$
Cell dimensions				
a, b, c (Å)	106.34, 40.37, 73.79	148.96, 148.96, 75.80	73.94, 65.03, 75.99	110.31, 136.92, 114.66
α, β, γ (°)	90.00, 121.40, 90.00	90.00, 90.00, 120.00	90.00, 107.08, 90.00	90.00, 100.26, 90.00
Resolution (Å)	45.21-1.50(1.58-1.50) *	65.35-2.15(2.27-2.15) *	37.03-1.7(1.79-1.70) *	87.07-3.02(3.19-3.02) *
R_{sym} or R_{merge}	6.3(23.8)	9.3(46.9)	6.9(37.3)	11.2(35.2)
$I/\sigma I$	9.1(2.0)	18.5(4.0)	9.8(2.6)	6.9(2.2)
Completeness (%)	96.2(74.8)	99.1(93.9)	97.7(95.4)	89.0(43.1)
Redundancy	3.8(2.0)	13.6(8.8)	4.0(3.5)	3.8(2.1)
Refinement				
Resolution (Å)	15.0-1.50	15.0-2.15	15.0-1.7	15.0-3.02
No. reflections	41,211(4,622)	27,218(3,678)	73,865(10,470)	58,488(4,120)
R_{work}/R_{free}	18.0/21.5	18.8/21.3	15.9/19.7	22.2/24.8
No. atoms				
Protein	2173	1681	5076	21120
Ligand/ion	2 TAR	1 ANP	4 ANP, 4 MG, 4 P6G	16 ANP
Water	255	189	424	-
B-factors				
Protein	22.1	37.2	18.4	66.4
Ligand/ion	22.2	44.8	34.6	36.5
Water	31.1	47.8	24.4	-
R.m.s deviations				
Bond lengths (Å)	0.012	0.012	0.010	0.009
Bond angles (°)	1.3	1.2	1.3	1.1

*Highest resolution shell is shown in parenthesis.

Supplementary Table 2. Distribution of Fic proteins from different domains of life in the three regulatory classes I to III.

Domain		Data analyzed				Distribution of Fic proteins into the different regulatory classes I to III														
				class I[d]		class II[e]		class III[f]		no class		class I and II		class I and III		class II and III				
Domain	Dataset[a]	Proteins[b]	Domains[c]	#	%	#	%	#	%	#	%	#	%	#	%	#	%			
All	FIC family	2189	2191	156	7.12	1217	55.55	73	3.33	741	33.82	1	0.05	1	0.05	2	0.09			
	conserved FIC motif	1129	1131	50	4.42	898	79.40	72	6.37	110	9.73	0	0.00	0	0.00	1	0.09			
Bacteria	FIC family	2072	2074	154	7.03	1146	52.30	73	3.33	697	31.81	1	0.05	1	0.05	2	0.09			
	conserved FIC motif	1058	1059	50	4.42	834	73.74	72	6.37	102	9.02	0	0.00	0	0.00	1	0.09			
Archaea	FIC family	49	49	0	0.00	30	1.37	0	0.00	19	0.87	0	0.00	0	0.00	0	0.00			
	conserved FIC motif	26	26	0	0.00	26	2.30	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00			
Eukaryota	FIC family	59	59	0	0.00	38	1.73	0	0.00	21	0.96	0	0.00	0	0.00	0	0.00			
	conserved FIC motif	40	40	0	0.00	35	3.09	0	0.00	5	0.44	0	0.00	0	0.00	0	0.00			
Viruses	FIC family	5	5	0	0.00	2	0.09	0	0.00	3	0.14	0	0.00	0	0.00	0	0.00			
	conserved FIC motif	4	4	0	0.00	2	0.18	0	0.00	2	0.18	0	0.00	0	0.00	0	0.00			
Unclass[g]	FIC family	1	1	0	0.00	1	0.05	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00			
	conserved FIC motif	1	1	0	0.00	1	0.09	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00			
Other[g]	FIC family	3	3	2	0.09	0	0.00	0	0.00	1	0.05	0	0.00	0	0.00	0	0.00			
	conserved FIC motif	1	1	0	0.00	0	0.00	0	0.00	1	0.09	0	0.00	0	0.00	0	0.00			

[a] we analyzed the entire FIC PFAM family (PFAM release 24) as well as a subset of Fic proteins harbouring an adenylylation-competent core motif (conserved FIC motif, HxFx[D/E]GNGRxxR),
[b] number of proteins analyzed,
[c] number of domains analyzed; two proteins harbour two FIC domains,
[d] class I, antitoxin-like ORF harbouring inhibition motif,
[e] class II, N-terminal inhibition motif,
[f] class III, C-terminal inhibition motif,
[g] "Unclass" stands for unclassified Fic proteins, "Other" comprises three Fic proteins from plasmid sequences sampled from the environment

Supplementary Table 3. Prediction of class I, class II, and class III inhibition motifs for selected Fic proteins.

Fic protein[a]	FIC motif[b]	predicted regulatory features (class I-III)[c]				HHpred results[d]																			
		class I	class II	class III	none	class II structures						class III structures						Structures without class II or III regulatory features							
						SoFic (3eqx)			BtFic (3cuc)			NmFic (2g03)			HpFic (2f6s)			Doc (3dd7)		BepA (2jk8)		VopS (3let)		IbpA (3n3u)	
						rank#	score	motif	rank#	score	motif	rank#	score	motif	rank#	score	motif	rank#	score	rank#	score	rank#	score	rank#	score
BepA (CAD89506)	HPFREGNGRTQR	yes[e]	no	no	no	7	84	no	4	130	no	3	226	no	2	240	no	8	41	1	686	6	92	5	105
IbpA_FIC1 (BAC78649)	HPFAEGNGRMAR	no	yes	no	no	3	237	GSAVDD	2	250	GSAVDD	5	162	no	6	155	no	8	59	7	137	4	207	1	427
IbpA_FIC2 (BAC78649)	HPFAEGNGRMAR	no	yes[f]	no	no	4	210	TKVIED	3	228	TKVIED	6	163	no	7	149	no	9	57	8	129	5	194	1	488
VopS (BAC59949)	HGFTDGNGRMGR	no	no	no	no	4	48	no	3	50	no	2	52	no	6	43	no	8	24	7	41	1	426	5	47
Doc (CAA66833)	HIFNDANKRTAL	no	no	no	no	6	72	no	3	79	no	4	75	no	5	73	no	1	279	2	86	8	71	7	72
E.coli_Fic (AAA23773)	HPFRVGSGLAQR	yes[g]	no	no	no	6	170	no	4	197	no	3	233	no	2	240	no	8	60	1	326	7	141	5	178
Hype/FicD (AAQ89351)	HPFIDGNGRTSR	no	yes	no	no	4	267	TVAIEG	1	379	TVAIEG	5	226	no	6	209	no	8	81	7	193	2	366	3	292
SoFic (AAN57237)	HPFIDGNGRTGR	no	yes	no	no	1	616	SSEIEN	2	286	SSEIEN	4	171	no	6	156	no	8	58	7	150	5	169	3	270
BtFic (AAO77620)	HPFEDGNGRIAR	no	yes	no	no	2	283	SNHLEG	1	492	SNHLEG	5	216	no	6	193	no	8	77	7	180	4	242	3	269
NmFic (NP_273311)	HPFLEGNGRSTR	no	no	yes	no	6	150	no	4	193	no	1	363	SYYYEG	2	338	SYYYEG	8	61	3	218	7	143	5	171
HpFic (ACI27853)	HPFLEGNGRATR	no	no	yes	no	7	145	no	4	187	no	2	316	SYYYEG	1	348	SYYYEG	8	63	3	208	6	148	5	171

[a] protein name and genbank accession
[b] grey shading indicates conserved FIC motif (HxFx[D/E]GNGRxxR)
[c] based on Psi-Blast and HHpred analysis, class I: antitoxin-like ORF harbouring inhibition motif; class II: N-terminal inhibition motif; class III: C-terminal inhibition motif
[d] results of the HHpred alignments with each of the eight template structures from the FIC family, the rank of each alignment, the alignment score, and the presence of an inhibition motif for class II and class III are indicated.
[e] accession of the antitoxin-like ORF: YP_034061
[f] in the IbpA xtal structure (3n3u) the identified motif locates to a loop preceding a helix α6 that is roughly in the position of a canonical α_{inh} (see Supplementary Text)
[g] accession of the antitoxin-like ORF: NP_417821

Supplementary Table 4. Classification of all PFAM FIC domain-containing proteins according to the presence of an antitoxin or an intrinsic inhibition motif. Due to its bulkiness, the table is provided as an extra file in Excel format on http://www.nature.com/nature.

Supplementary Table 5. List of primers used in this study.

Primer name	Sequence (5'-3')
prPE484	CCGCTCGAGGTGAGGAAATATGAGGGTAGC
prPE485	CCGCTCGAGTTACCTTGTAATTCCCTTTGAAG
prPE500	CCGGAATTCAAGAAGGAGATATACATGAGACCATGGCCTACCCATAC
prPE501	CCGGAATTCTTACCTTGTAATTCCCTTTGAAG
prPE519	GGAAGATCTTCATATTTCCTCACGTTTTATCCG
prPE520	GGGAATTCCATATGGCCTACCCATACGATGTTCCAGATTACGCGGCCGCGATGTTGAGCGAGGAAGAAATC
prPE526	CGGGATCCAGCTGCACTTTATAATGTTCCTC
prPE527	ATTACATCTCCTTCAATTACCTA
prPE530	TAGGTAATTGAAGGAGATGTAATTAATTGCAATTATATTATTTCTTGAC
prPE517	CGGGATCCCATCAATGCTTGAAGGAATATGG
prAH095	CGGGATCCATGCAGACAATTAAGTGTGTTG
prAH097	CCGCTCGAGTTAGAATATACAGCACTTCCT
prAH106	CGGGATCCATGCAGGCCATCAAGTGTGT
prAH107	CCGCTCGAGTTACAACAGCAGGCATTTTC
prAH116	GAATGCGTTGGCCCCGTTTCGAGAAGGTAATGGACGTAC
prAH127	CCTTCTCGAAACGGGGCCAACGCATTCAATTCGCCCATG
prAH138	GGGATACAGCTGGACAGGAAGATTATGACAGATTACGCCCCC
prAH139	CTGTCATAATCTTCCTGTCCAGCTGTATCCCATAAGCCCAG
prAH181	CTTTTTGATACTGCAGGGCAAGAGGATTATGACAGATTACG
prAH182	CATAATCCTCTTGCCCTGCAGTATCAAAAAGTCCAAGAGTATATGG
prAG041	GCGCCCATGGTGTTGAGCGAGGAAGAAATCG
prAG042	CGCGGATCCTCATATTTCCTCACGTTTTATCC
prFVS037	GGGAATTCCATATGGTGAGGAAATATGAGGGTAGC
prFVS001	CGACCTCGAGTTAGTGATGGTGATGGTGATGTGTAAATTCAGTGAGGTTTCTAC
prFVS007	GGGAATTCCATATGCATCACCATCACCATCACATGAAATCCATAGACGAACAAAG
prFVS008	CGACCTCGAGTCAGCCTTTTTCATACCCTTCG
prFVS009	CGACCTCGAGTTAGTCAGTCAGGTTGTCCTTTAAC
prFVS010	CAGGCCTATTATTACGCAGGGTATGAAAAAG
prFVS011	CCTGCGTAATAATAGGCCTGCTCGATAC
prFVS026	GTATCGAGCAGGCGTATTATTACGCAGGGTATG
prFVS027	GTAATAATACGCCTGCTCGATACCTTTAAAGATG
prFVS032	GGGAATTCCATATGCATCACCATCACCATCACGAATGGCAAGCTGAACAAGC
prFVS033	CGACCTCGAGTTAAAGCGCATAACGGCTGAAC
prFVS063	CGTTTAGGCGGTTTGGAACCAGATCCACAAG
prFVS064	CTGGTTCCAAACCGCCTAAACGTTGACTAGC
prFVS067	CCAAGGCAGTAGCGAGATTGGCAATATCGTTACCACCAC
prFVS068	GATATTGCCAATCTCGCTACTGCCTTGGGCTTCGAGCAG
prFVS076	GTGGCCATCGGCGGCAACACCCTCACCCTCTC
prFVS077	GTGTTGCCGCCGATGGCCACTGTGTGGTAG
prFVS082	GTATTATTACGGCGGGTATGAAAAAGGCTGAC
prFVS083	CATACCCGCCGTAATAATACGACTGCTCG
prFVS088	GGGAATTCCATATGCATCACCATCACCATCACGTGGAAGAGATCGACCAGAG
prFVS089	CCGCTCGAGTCACGAGTACTCAGTTGTGGCAAAAAG

Supplementary Table 6. List of plasmids constructed in this study.

Plasmid name	Description	Primers used	
		fwd	rev
pPE0017	VbhT (MCS1) in pRSFDuet1	prPE484	prPE485
pPE0020	VbhA (MCS2) in pRSFDuet1	prPE520	prPE519
pPE0021	VbhT (MCS1) + VbhA (MCS2) in pRSFDuet1	-	-
pPE0034	VbhT His136Ala (MCS1) in pRSFDuet1	prAH116	prAH127
pPE1018	HA-VbhT in pMMB206 (pVbhT)	prPE500	prPE501
pPE1031	HA-VbhT(H136A) in pMMB206 (pVbhT(H136A))	prAH116	prAH127
pPE3005	VbhA-VbhT deletion fragment in pTR1000	prPE526	prPE517
pFVS0011	VbhA (MCS1) + VbhT 1-198 (MCS2) in pRSFDuet1	prAG037	prFVS001
pFVS0012	VbhT 1-198 (MCS2) in pRSFDuet1	prAG037	prFVS001
pFVS0015	NmFic 11-191 (MCS2) in pRSFDuet1	prFVS007	prFVS008
pFVS0016	NmFic 11-167 Δ8 (MCS2) in pRSFDuet1	prFVS007	prFVS009
pFVS0023	NmFic 11-191 Glu186Ala (MCS2) in pRSFDuet1	prFVS010	prFVS011
pFVS0037	NmFic 11-191 Ser182Ala Glu186Ala (MCS2) in pRSFDuet1	prFVS026	prFVS027
pFVS0040	SoFic 1-372 (MCS2) in pRSFDuet1	prFVS032	prFVS033
pFVS0058	SoFic 1-372 Glu73Gly (MCS2) in pRSFDuet1	prFVS067	prFVS068
pFVS0059	NmFic 11-191 Glu186Gly (MCS2) in pRSFDuet1	prFVS082	prFVS083
pFVS0065	VbhA Glu24Gly (MCS1) + VbhT 1-198 (MCS2) in pRSFDuet1	prFVS063	prFVS064
pFVS0080	VbhA Glu24Gly (MCS1) + VbhT 1-198 His136Ala (MCS2) in pRSFDuet1	prAH116	prAH127
GST-HypE	GST-HypE in pET-GSTx (kindly provided by Jack Dixon)	-	-
GST-HypE$_{H/A}$	GST-HypE His295Ala in pET-GSTx (kindly provided by Jack Dixon)	-	-
pFVS0064	GST-HypE Glu234Gly in pET-GSTx	prFVS076	prFVS077
pFVS0085	HypE(FIC) 187-437 (MCS2) in pRSFDuet1	prFVS088	prFVS089
pFVS0087	HypE(FIC) 187-437 His295Ala (MCS2) in pRSFDuet1	prFVS088	prFVS089
pFVS0091	HypE(FIC) 187-437 Glu234Gly (MCS2) in pRSFDuet1	prFVS088	prFVS089
pFVS0058	SoFic 1-372 Glu73Gly (MCS2) in pRSFDuet1	prFVS067	prFVS068
pAH059	GST-Cdc42 Gln61Leu		
pAH088	GST-Cdc42		
pAH060	GST-Rac1 Gln61Leu		
pAH071	GST-Rac1		

4. References

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

2   Palanivelu, D. V. *et al.* Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein Sci* **20**, 492-499 (2011).

3   Luong, P. *et al.* Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J Biol Chem* **285**, 20155-20163 (2010).

4   Worby, C. A. *et al.* The fic domain: regulation of cell signaling by adenylylation. *Mol Cell* **34**, 93-103 (2009).

5   Yarbrough, M. L. *et al.* AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* **323**, 269-272 (2009).

6   Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244-248 (2005).

7   Schmid, M. C. *et al.* A translocated bacterial protein protects vascular endothelial cells from apoptosis. *PLoS Pathog* **2**, e115 (2006).

8   Lehnherr, H., Maguin, E., Jafri, S. & Yarmolinsky, M. B. Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J Mol Biol* **233**, 414-428 (1993).

9   Liu, M., Zhang, Y., Inouye, M. & Woychik, N. A. Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. *Proc Natl Acad Sci U S A* **105**, 5885-5890 (2008).

10   Magnuson, R. & Yarmolinsky, M. B. Corepression of the P1 addiction operon by Phd and Doc. *Journal of bacteriology* **180**, 6342-6351 (1998).

11   Mukherjee, S. *et al.* Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477**, 103-106 (2011).

100

12   Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and

space complexity. *BMC Bioinformatics* **5**, 113 (2004).

13   Garcia-Pino, A. *et al.* Doc of prophage P1 is inhibited by its antitoxin partner Phd

through fold complementation. *J Biol Chem* **283**, 30821-30827 (2008).

14   Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood

trees for large alignments. *PLoS One* **5**, e9490 (2010).

15   Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with

thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690 (2006).

3.3 Research article III

Conserved inhibitory mechanism and competent ATP binding mode for adenylyltransferases with Fic fold

Arnaud Goepfert, Frédéric V. Stanger, Christoph Dehio, and Tilman Schirmer

Statement of the own participation

I contributed to this publication by expressing, purifying and crystallizing VbhA−VbhT(FIC), VbhA$_{E24G}$−VbhT(FIC), SoFic, SoFic$_{E73G}$, and NmFic$_{E186G}$, and by determining their structures. I also performed *in vitro* adenylylation assays. Frédéric Stanger constructed expression plasmids. The manuscript was written by me, Christoph Dehio, and Tilman Schirmer.

Conserved Inhibitory Mechanism and Competent ATP Binding Mode for Adenylyltransferases with Fic Fold

Arnaud Goepfert[1,2], Frédéric V. Stanger[1,2], Christoph Dehio[1]*, Tilman Schirmer[2]*

1 Focal Area Infection Biology, Biozentrum, University of Basel, Basel, Switzerland, 2 Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, Basel, Switzerland

Abstract

The ubiquitous FIC domain is evolutionarily conserved from bacteria to human and has been shown to catalyze AMP transfer onto protein side-chain hydroxyl groups. Recently, it was predicted that most catalytically competent Fic proteins are inhibited by the presence of an inhibitory helix α_{inh} that is provided by a cognate anti-toxin (class I), or is part of the N- or C-terminal part of the Fic protein itself (classes II and III). *In vitro*, inhibition is relieved by mutation of a conserved glutamate of α_{inh} to glycine. For the class III bacterial Fic protein NmFic from *Neisseria meningitidis*, the inhibitory mechanism has been elucidated. Here, we extend above study by including bacterial class I and II Fic proteins VbhT from *Bartonella schoenbuchensis* and SoFic from *Shewanella oneidensis*, respectively, and the respective E->G mutants. Comparative enzymatic and crystallographic analyses show that, in all three classes, the ATP substrate binds to the wild-type FIC domains, but with the α-phosphate in disparate and non-competent orientations. In the E->G mutants, however, the tri-phosphate moiety is found reorganized to the same tightly bound structure through a unique set of hydrogen bonds with Fic signature motif residues. The γ-phosphate adopts the location that is taken by the inhibitory glutamate in wild-type resulting in an α-phosphate orientation that can be attacked in-line by a target side-chain hydroxyl group. The latter is properly registered to the Fic active center by main-chain β-interactions with the β-hairpin flap. These data indicate that the active site motif and the exposed edge of the flap are both required to form an adenylylation-competent Fic protein.

Introduction

Adenylyl transferases (ATases) utilize adenosine triphosphate (ATP) to covalently modify proteins, nucleic acids, or small molecules with adenosine monophosphate (AMP), a reaction known as adenylylation or AMPylation. The ubiquitous FIC domain (pfam 02661) found in proteins of all domains of life and viruses has only recently been shown to confer ATase activity. Thus, the bacterial T3SS effector protein VopS from *Vibrio parahaemolyticus* and the surface antigen IbpA from *Histophilus somni* covalently attach the bulky AMP moiety onto a specific threonine or tyrosine, respectively, of the switch I region of Rho family GTPases [1,2]. This abrogates binding of downstream effectors and results in actin cytoskeleton collapse and concomitant cell detachment and death. Mutational and bioinformatics analysis indicated that Fic proteins containing a strictly conserved HxFx(D/E)GNGRxxR signature motif in the active center typically display adenylylation activity [1,2,3,4,5], while Fic proteins with an active center deviating from this consensus are considered to have adopted different activities. Indeed, the host-targeted effector protein AnkX of *Legionella pneumophila* exhibiting an HxFxDANGRxxV signature motif displays phosphocholination activity towards the GTPase Rab1 [6].

The FIC domain is structurally characterized by a conserved central core of four helices (α2 to α5) that is flanked by three helices (α1, α6 and α7) found in diverse dispositions in different Fic proteins [3,7]. Helices α4 and α5 are joined by a loop that together with the N-terminal cap of helix α5 forms the active center represented by a signature motif with the consensus sequence HxFx(D/E)GNGRxxR. The catalytic mechanism of adenylylation was deduced from the crystal structure of the second FIC domain of IbpA in complex with the adenylylated Cdc42 target [4] and from biochemical studies [5] and shown to involve nucleophilic attack of the target side-chain hydroxyl onto the ATP α-phosphate. The triphosphate binding site at the anionic nest at the N-terminus of helix α5 was characterized by the crystal structure of BepA from *Bartonella henselae* in complex with pyrophosphate, the side product of the reaction [3]. An ATP substrate complex structure was obtained recently for the Fic protein of *Neisseria meningitidis* [8] corroborating the catalytic mechanism. The histidine of the signature motif is critical for deprotonation of the incoming target hydroxyl group [5], whereas the phenylalanine is part of the hydrophobic core of the domain. The remaining residues of the motif are involved in ATP/Mg^{2+} binding and loop stabilization [3,8].

We recently demonstrated that the Fic protein VbhT from *Bartonella schoenbuchensis* causes bacterial growth arrest when overexpressed in *Bartonella* or *E. coli* and that this effect can be repressed by co-expression with the anti-toxin VbhA, a small protein encoded upstream of VbhT [8]. As shown by structure

105

analysis, VbhA forms a tight complex with the FIC domain of VbhT with the conserved glutamate (E_{inh}) from the inhibitory helix α_{inh} partly obstructing the ATP binding site, which gave a first clue regarding the inhibitory mechanism mediated by VbhA binding.

Exhaustive bioinformatic analysis coupled with homology modeling revealed that the (S/T)xxxE(G/N) signature motif of α_{inh} is not only found in several other putative anti-toxin sequences coded immediately upstream of Fic proteins, but is often part of the FIC domain itself either preceding helix $\alpha 1$ or immediately following helix $\alpha 7$ [8]. Thus, a classification system was introduced grouping the Fic proteins for which an anti-toxin with an inhibitory helix α_{inh} had been found into class I and those with an equivalent of α_{inh} in the N- or C-terminal part of the Fic protein into classes II and III, respectively. Indeed, 90% of the Fic proteins with the canonical FIC signature motif could be classified accordingly, suggesting that all these enzymes are inhibited in their enzymatic activity.

The physiological stimulus or condition for relief of α_{inh}-mediated inhibition is not yet known. For T4SS Fic proteins of class I (such as VbhT or BepA [9]), however, it appears likely that, for injection into host cells, the Fic protein has to unfold and will be translocated without the antitoxin. For class II and III proteins, detachment, unfolding, or proteolytic cleavage of the α_{inh} helix may cause relief of inhibition. In fact, a truncation mutant of the class III Fic protein from *N. meningitidis* (NmFic) lacking the entire C-terminal α_{inh} helix showed strong ATase activity and allowed to study the catalytic and inhibitory mechanism in detail [8]. A more subtle means to relieve inhibition, which is applicable to Fic proteins of all three classes, is the replacement of the inhibitory glutamate by glycine. *In vivo*, such E->G mutations showed a detrimental effect on bacterial growth [8]. For the human HYPE protein (class II), the corresponding mutant protein catalyzed *in vitro* AMP transfer to the small GTPases Rac1 and Cdc42, whereas only marginal effect was seen with the wild-type proteins [8].

Here, we assayed in a systematic approach Fic representatives of the three Fic classes and their E->G mutants for *in vitro* adenylylation showing that the mutation causes inhibition relief across the Fic classes. Binding of ATP substrate or AMPPNP substrate analog to the wild-type and the E->G mutant proteins was studied by protein crystallography to reveal the inhibitory mechanism and to get further insight into catalysis. This yielded a consistent molecular mechanism that most likely applies to most adenylylation competent Fic proteins irrespective of class.

Materials and Methods

Cloning

The full-length *vbhA* gene and part of the *vbhT* gene (amino acid residues 1–248, His$_6$-tagged) were amplified from plasmid pPE0021 and cloned into the pRSF-Duet1 vector leading to plasmid pAG0077 (VbhA/VbhT(FIC)). The full-length *vbhA* gene and part of the *vbhT* gene encoding the FIC domain (amino acid residues 1–198, His$_6$-tagged) were PCR-amplified from plasmid pPE0021 and cloned into the pRSF-Duet1 vector (pFVS0011). A two-base pair mutation is then introduced in pFVS0011 to obtain plasmid pFVS0065 (VbhA$_{E24G}$/VbhT(FIC)). The *fic* gene of *Neisseria meningitidis* was PCR-amplified with an N-terminal His$_6$-tag from *Neisseria meningitidis* from coding region of amino acid residues 11–191 to generate plasmid expressing NmFic (pFVS0015). The E186G mutant construct (NmFic$_{E186G}$; pFVS0059) was generated by introducing a two-base pair mutation in pFVS0015. The *fic* gene of *Shewanella oneidensis* was

PCR-amplified from plasmid (ASU biodesign institute, Clone ID SoCD00104192) and cloned with an N-terminal His$_6$-tag into pRSF-Duet1 (pFVS0040). The SoFic$_{E73G}$ plasmid (pFVS0058) was generated by introducing a two-base pair point mutation in pFVS0040.

Protein Expression and Purification

Vectors pAG0077 (VbhA/VbhT(FIC)), pFVS0040 (SoFic) and pFVS0015 (NmFic) were transformed into *E.coli* BL21 (DE3). *E. coli* cultures were grown at 37°C in LB medium supplemented with 50 µg/ml of kanamycin to an OD_{595} of 0.6 before induction with 0.3 mM IPTG for 16 h at 23°C. Vectors pFVS0065 (VbhA$_{E24G}$/VbhT(FIC)), pFVS0059 (NmFic$_{E186G}$), pFVS0058 (SoFic$_{E73G}$) were transformed into BL21-AI cells. Cells were incubated in 750 ml LB medium supplemented with 50 µg/ml kanamycin and 1% glucose at 37°C at 200 rpm until an OD_{595} value of 1.5 was reached. Bacterial pellets were resuspended in 1 L of Terrific Broth media containing 50 µg/ml^{-1} kanamycin. Protein expression was induced at 23°C with 0.1% arabinose and 0.1 mM IPTG for 23 h at 200 rpm.

Cells containing overexpressed VbhA/VbhT(FIC) and NmFic were resuspended in lysis buffer containing 20 mM Tris (pH 7.5), 250 mM NaCl, and 25 mM imidazole and disrupted using French press. Cell debris were pelleted by ultracentrifugation and the supernatant was applied to a His-Trap column (GE Healthcare). The proteins were eluted with a gradient of elution buffer containing 20 mM Tris (pH 7.5), 250 mM NaCl, and 500 mM imidazole. The proteins were then concentrated and injected on a Superdex 75 16/60 gel filtration column (GE Healthcare) equilibrated with 10 mM Tris (pH 7.6) and 100 mM NaCl. The pure proteins were concentrated to 3.7 mg/ml for VbhA/VbhT(FIC) and 30 mg/ml for NmFic.

The same purification protocol as described above was used for VbhA$_{E24G}$/VbhT(FIC) and NmFic$_{E186G}$ with an additional intermediate purification step. After affinity purification, the proteins were adjusted to 20 mM Tris (pH 8.5), 25 mM NaCl, applied to a Resource-Q anion exchange column (Amersham Biosciences), and eluted with a linear gradient of 1 M NaCl. Peak fractions were concentrated and further purified by gel filtration chromatography. Purified proteins in 10 mM Tris (pH 7.6), 100 mM NaCl were concentrated to 4.1 mg/ml for VbhA$_{E24G}$/VbhT(FIC) and 33 mg/ml for NmFic$_{E186G}$. Cells containing overexpressed SoFic and SoFic$_{E73G}$ were resuspended in lysis buffer containing 50 mM HEPES (pH 8.0), 50 mM NaCl, 1 mM TCEP, 10% glycerol and 10 mM Imidazole and disrupted using French press. Cell debris were pelleted by ultracentrifugation and the supernatant was applied to a His-Trap column (GE Healthcare). The proteins were eluted with a gradient of elution buffer containing 50 mM HEPES (pH 8.0), 50 mM NaCl, 1 mM TCEP, 10% glycerol and 300 mM imidazole. The proteins were then concentrated and injected on a Superdex 75 16/60 gel filtration column (GE Healthcare) equilibrated with 20 mM HEPES (pH 8.0), 200 mM NaCl and 1 mM TCEP. The pure proteins were concentrated to 21.8 mg/ml for SoFic and 12 mg/ml for SoFic$_{E73G}$.

Protein Crystallization

For crystallization, the hanging-drop vapor diffusion method was used with 1 µl protein solution mixed with 1 µl reservoir solution. The VbhA/VbhT(FIC) and VbhA$_{E24G}$/VbhT(FIC) complexes were concentrated to 3.7 mg/ml and 4.1 mg/ml, respectively, and crystallized at 20°C using a reservoir solution composed of 15% (w/v) PEG 4000, 0.1 M MES pH 6.5. Whereas, the wild-type crystal was soaked with 5 mM ATP, and 5 mM

MgCl$_2$, the mutant was co-crystallized with 10 mM ATP, and 10 mM MgCl$_2$. For data collection, crystals were transferred to reservoir solutions supplemented with 20% glycerol and flash frozen in liquid nitrogen. SoFic and SoFic$_{E73G}$ were concentrated to 21.8 mg/ml and 12 mg/ml, respectively, and co-crystallized with either 5 mM ATP or 5 mM AMPPNP and supplemented with 5 mM MgCl$_2$ in a solution composed of 21% (w/v) PEG 3350 and 0.2 M NaF pH 7.1 at 4°C. For data collection, crystals of the protein-ligand complex were cryoprotected by transfer to a reservoir solution supplemented with 15% (v/v) PEG 200 and flash cooled in liquid nitrogen. For crystallization of NmFic$_{E186G}$ (33 mg/ml), a reservoir solution composed of 4 M potassium formate, 0.1 M Bis-Tris propane pH 9.0, 2% (w/v) PEG MME 2000 was used. Crystals were soaked with 5 mM AMPPNP and 5 mM MgCl$_2$ and then cryoprotected with 20% glycerol prior flash-cooling in liquid nitrogen.

Data Collection, Structure Determination, and Refinement

Diffraction data were collected at the Swiss Light Source at 100 K and processed using XDS [10]. The structures were solved by molecular replacement using the apo structures of VbhA/VbhT(FIC) (PDB code 3SHG), SoFic (PDB code 3EQX) or NmFic (PDB code 2G03) as search models using Phaser [11]. Several rounds of iterative model building and refinement were performed using Coot [12] and PHENIX [13] or REFMAC5 [14], respectively. 5% of the data were excluded from refinement and used for cross-validation. The geometry of the final model was assessed using MolProbity [15] showing >99% of the residues in the core and allowed regions of the Ramachandran plot. Data collections and refinement statistics are summarized in Table 1. The atomic coordinates and structure factors of VbhA/VbhT(FIC)/ATP, VbhA$_{E24G}$/VbhT(FIC)/ATP, SoFic/ATP, So-Fic$_{E73G}$/AMPPNP, and NmFic$_{E186G}$/AMPPNP have been deposited in the Protein Data Bank under accession codes 3ZC7, 3ZCB, 3ZCN, 3ZEC and 3ZLM, respectively. The figures were generated with Dino (A. Philippsen unpublished, http://www.dino3d.org).

In vitro Adenylylation Assay

Adenylylation activity of VbhA/VbhT(FIC), SoFic and NmFic constructs was assessed by incubating 125 ng, 1.25 µg and 2.5 µg of purified protein, respectively, with 10 µCi α-^{32}P-ATP (Hartmann Analytic) in a buffer containing 50 mM Tris pH 8.0, 150 mM NaCl, 0.1 mM EGTA, 15 mM MgCl$_2$, and protease inhibitor cocktail (Roche). Reactions were incubated for 1 h at 30°C, resolved by SDS–PAGE, and subjected to autoradiography.

Results

Constitutive Inhibition is Relieved by Truncation of the Inhibitory Glutamate in All Three Fic Classes

For the comparative structure/function study on the inhibitory mechanism of Fic proteins from the various classes we chose as representatives the FIC domain of VbhT (residues 1 to 198) from *Bartonella schoenbuchensis* in complex with its cognate antitoxin VbhA (VbhA/VbhT(FIC); class I), Fic protein SO_4266 *from Shewanella oneidensis* (SoFic; class II) and Fic protein NMB0255 from *Neisseria meningitidis* (NmFic; class III).

Auto-adenylylation is a convenient read-out to assess adenylylation activity of Fic proteins. It does not require the presence of a physiological protein target that may, in fact, not yet been known as in the case of SoFic. Autoradiographies of SDS-PAGE gels after incubation with α-^{32}P-ATP (Fig. 1) show that auto-adenylylation is virtually absent in the wild-type Fic proteins of all three classes, i.e. for VbhA/VbhT(FIC), SoFic, and NmFic (see also ref. 8), but is drastically boosted in the respective E->G mutants suggesting a common inhibitory mechanism.

ATP Binds to Wild-type Fic Proteins in Disparate and Catalytically Incompetent Conformations

Fig. 2 shows the high-resolution structures of VbhA/VbhT(FIC) (class I) and SoFic (class II), both in complex with ATP. Whereas VbhA/VbhT(FIC) crystallized isomorphously to the unliganded wild-type crystals ([8], PDB code 3SHG), SoFic yielded crystals of monoclinic space group, i.e. distinct to the orthorhombic form of the apo structure ([16], PDB code 3EQX). In the two structures the nucleotide is clearly visible, albeit with elevated B-factors (40 Å^2) in VbhA/VbhT(FIC). Only marginal structural changes are induced upon substrate binding (rms deviations between the Cα-positions of apo and complex form of 0.4 Å and 0.8 Å for VbhA/VbhT(FIC) and SoFic, respectively).

In both structures the ATP substrate is found at analogous sites (Fig. 2) with the base filling a pocket formed by α4, α6, and the β-hairpin flap, the ribose 3′-hydroxyl H-bonded to the conserved glutamate of α$_{inh}$, and the triphosphate moiety interacting with the anionic nest formed by the N-terminus of α5. The same binding mode has been observed for class III NmFic [8]. In all three structures, also the ribose 2′-hydroxyl is forming an H-bond, but to non-homologous protein side-chains. Similarly, the binding subsite for the base is not conserved on the residue level. However, in each case, hydrophobic residues are contributed by helix α6 and by the flap. A weak H-bond is formed between the adenine N3 and N133 in VbhT(FIC). A homologous interaction (with N104) is found in NmFic [8].

Most relevant for catalysis is the orientation of the α-phosphate that has to be accessible to nucleophilic attack by the target side-chain hydroxyl group. In VbhA/VbhT(FIC) and SoFic, as in NmFic [8], the position that is in-line with the scissile Pα-O3α bond is not accessible for an attacking group (Fig. 2). Such a group positioned there would severely clash with atoms of the enzyme. Thus, in Fic proteins of all three classes, catalytically non-competent orientation of the α-phosphate appears to be the reason for the lack of adenylylation activity.

Interestingly, while the α-phosphate is locked in a secured position in each of the structures, it shows distinct orientations among the three proteins that can be traced back to differences in the binding mode of the β- and γ-phosphates (Fig. 2C). Though interacting with the same protein groups (anionic nest; histidine, asparagine, and first arginine of the signature motif), the detailed H-bonding patterns are different (e.g. the main chain amide of the second glycine of the motif interacts with the bridging O3β in VbhA/VbhT(FIC), and with the non-bridging O1β in SoFic).

It seems that during convergent evolution of α$_{inh}$-mediated adenylylation inhibition in the different Fic protein classes no strict constraints for the ATP binding mode were operational apart from the requirement for a non-competent orientation for the reacting phosphate.

Truncation of the Inhibitory Glutamate Allows the ATP Substrate to Bind in a Catalysis Competent Conformation

Relief of Fic protein inhibition was achieved previously by expression of VbhT without its cognate antitoxin VbhA or by replacing in NmFic the SxxxE inhibition motif by AxxxA or – most drastically - by deleting the entire α$_{inh}$ [8]. The conserved glutamate of α$_{inh}$, E$_{inh}$, was identified to be crucial for the inhibitory effect, since mere truncation of its side-chain (E->G

Table 1. Data collection and refinement statistics.

Protein	VbhA/VbhT(FIC)	VbhA$_{E24G}$/VbhT(FIC)	SoFic	SoFic$_{E73G}$	NmFic$_{E186G}$
Ligand	ATP	ATP	ATP	AMPPNP	AMPPNP
PDB code	3ZC7	3ZCB	3ZCN	3ZEC	3ZLM
Data collection					
Wavelength (Å)	1.000	0.979	0.979	0.979	1.000
Detector	MAR225 CCD	PILATUS 2M	MAR225 CCD	MAR225 CCD	PILATUS 2M
Space group	C2	C2	P2$_1$	P2$_1$2$_1$2$_1$	P6$_4$22
Cell dimensions					
a, b, c (Å)	106.5, 40.6, 73.7	106.5, 40.3, 73.9	37.8, 164.9, 70.2	71.3, 80.6, 141.8	149.1, 149.1, 76.4
$\alpha\ \beta\ \gamma$ (°)	90.0, 121.6, 90.0	90.0, 121.4, 90.0	90.0, 94.4, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 120.0
Resolution (Å)	45.4−2.1 (2.2−2.1)	45.4−1.9 (2.1−1.9)	35.5−1.7 (1.8−1.7)	42.7−2.2 (2.3−2.2)	49.3−2.0 (2.1−2.0)
R_{sym} or R_{merge} (%)	8.5 (33.8)	5.7 (33.4)	4.4 (41.3)	10.5 (52.8)	6.3 (72.7)
CC(1/2) (%)	99.8 (93.6)	99.9 (93.1)	99.9 (87.1)	99.8 (90.3)	100.0 (97.9)
I/σ	18.9 (5.7)	14.5 (3.4)	22.9 (3.4)	17.6 (4.5)	31.1 (5.7)
Completeness (%)	99.2 (92.3)	99.2 (97.0)	99.5 (96.7)	100.0 (100.0)	99.9 (100.0)
Multiplicity	5.4 (4.9)	3.6 (3.6)	3.9 (3.4)	7.4 (7.5)	21.4 (22.6)
Refinement					
Resolution (Å)	15.0−2.10	30.0−1.94	15.0−1.70	15.0−2.20	30.0−2.00
No. reflections	15,769 (2,342)	18,923 (1,355)	93,100 (2,837)	42,085 (3,956)	32,490 (2,338)
R_{work}/R_{free} [%]	16.6/23.0	19.4/23.4	16.7/20.2	16.5/21.1	18.2/19.9
Mol./a.u	1	1	2	2	1
No. atoms					
Protein	2172	2011	5961	5984	1458
Ligand/ion	1 ATP	1 ATP, 1 MG	2 ATP	2 ANP, 1 MG	1 ANP, 1 MG
Water	226	134	981	695	135
Average B (Å^2)					
Protein	22.0	25.0	21.9	20.7	45.9
Ligand/ion	39.7	27.5/11.4	21.9	12.9/28.9	40.1/64.8
Water	28.4	32.1	33.7	28.0	49.3
R.m.s deviations					
Bond lengths (Å)	0.007	0.011	0.008	0.009	0.010
Bond angles (°)	1.0	1.3	1.2	1.2	1.2

Values for the highest resolution shell are shown in brackets.
doi:10.1371/journal.pone.0064901.t001

mutation) rendered recombinantly overexpressed Fic proteins of all three classes toxic to *E. coli* [8].
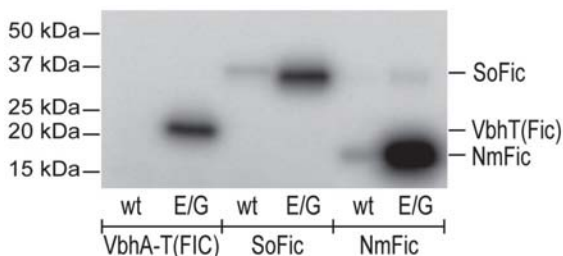
Figure 1. AMP transfer catalyzed by Fic proteins and their inhibition-relieved variants. Autoradiography of VbhA/VbhT(FIC), SoFic and NmFic (wt, wild type; E/G, E->G mutant) after incubation with radioactively labeled α-^{32}P-ATP.
doi:10.1371/journal.pone.0064901.g001

In vitro, the mutation has a drastic effect in that auto-adenylylation is boosted in all three representative Fic proteins (Fig. 1). This opens the door for studying the action of any active Fic protein *in vivo*, even without knowing the physiological stimulus for inhibition relief.

To reveal the underlying inhibition relief mechanism, crystal structures of the three mutant proteins in complex with ATP or AMPPNP were determined to high resolution. Although, in solution the mutants show auto-adenylylation, no such modification is observed in the crystal structures. For NmFic$_{E186G}$ this is not surprising, since the complex structure has been obtained by soaking and auto-adenylylation would require partial unfolding of α$_{inh}$ carrying the modifiable tyrosine (Y183) [8]. The VbhA$_{E24G}$/VbhT(FIC) and SoFic$_{E73G}$ complexes were co-crystallized. Since we do not see adenylylated residues, the extend of modifications may be either minor, locate to flexible loops or only the unmodified fraction may have crystallized. Figs. 3A–C show that in all three cases, the nucleotide is well resolved and, in contrast to
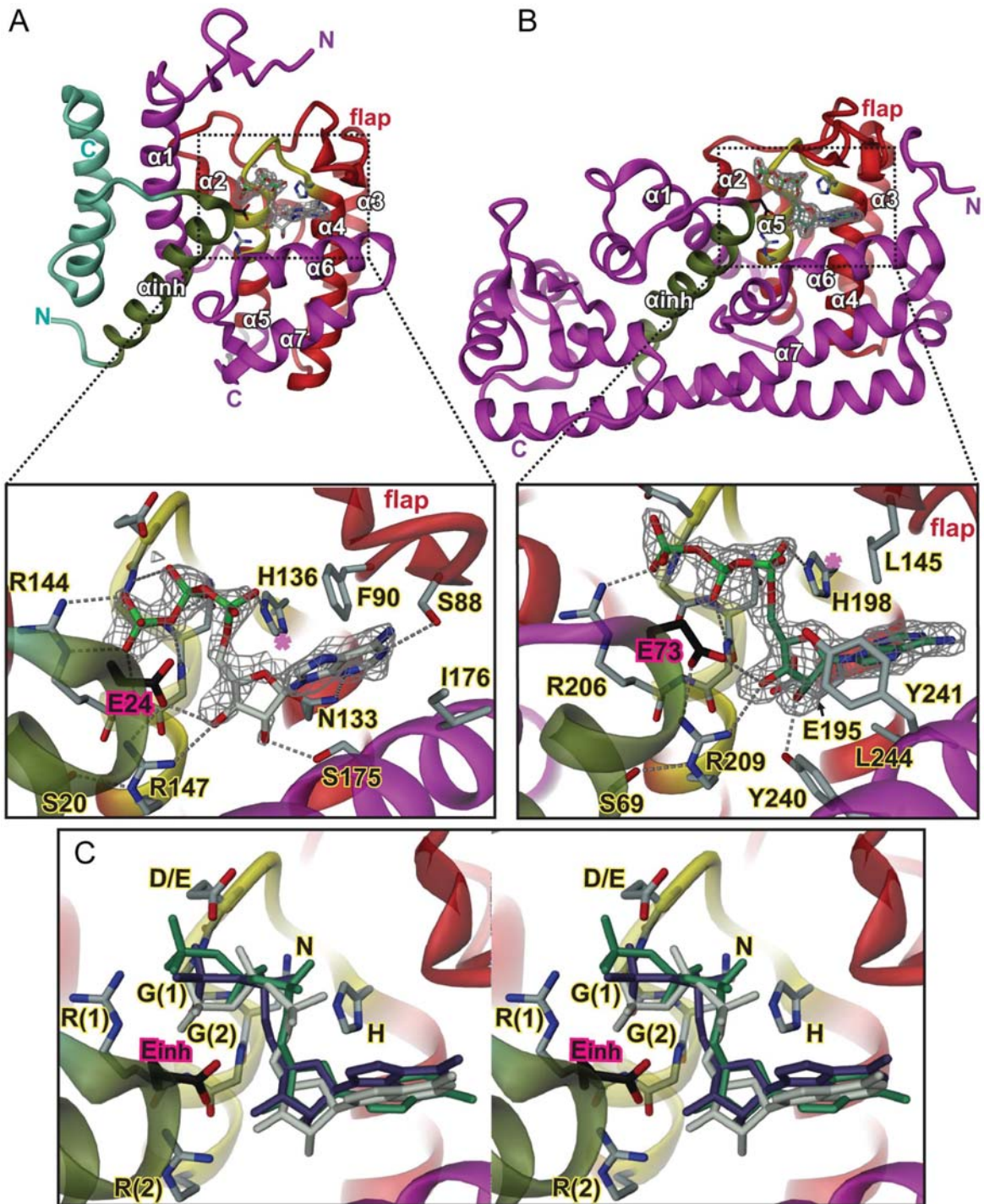
Figure 2. Crystal structures of wild-type Fic proteins representing classes I to II in complex with ATP substrate. (A) VbhA/VbhT(FIC), (B) SoFic. Structures are shown in cartoon representation (red, FIC core as defined by PFAM; yellow, active site loop and N-terminal end of helix α5; dark-green, inhibitory helix α$_{inh}$). In (A), the fold of the antitoxin is shown in dark-green and steel-blue. Selected residues are shown in full with the inhibitory glutamate (E24 or E73, respectively) colored in dark. The 2Fo-Fc simulated annealing omit maps covering the ligand are contoured at 1.1 σ. In both structures, the orientation of the α-phosphate prevents nucleophilic attack of a putative target side-chain hydroxyl onto the ATP substrate, since the position inline with the scissile Pα-O3α bond (magenta star) is unattainable. C) Stereo view of the superposition of the ATP nucleotides

shown in panel A and B with AMPPNP from the complex structure of the class III NmFic protein (PDB 3S6A [8]) within the active site of the VbhA/VbhT(FIC) complex (same as in panel A). The nucleotides of the various complexes are distinguished by their colors (white for the ATP bound to VbhA/VbhT(FIC), green for the ATP bound to SoFic, and blue for the AMPPNP of the NmFic complex. Note that the AMPPNP γ-phosphate in NmFic is found disordered [8] and therefore not shown. The residues of the HxFx(D/E)GNGRxxR Fic signature motif are labeled, the two glycine and the two arginine residues are distinguished by a "1" or "2" in brackets. The phenylalanine (not shown) is part of the hydrophobic core. The inhibitory glutamate from α_{inh} is labeled as E_{inh}.

doi:10.1371/journal.pone.0064901.g002

the wild-type complexes, shows a unique conformation and relative position within the binding site (Fig. 3D).

While the base and ribose moieties interact with the mutant in the same way as with the wild-type proteins (compare with Fig. 2, see also Fig. 3a in [8] for NmFic), the triphosphate has adopted a strongly curved conformation with the terminal γ-phosphate approaching closely the ribose moiety and forming a tight salt-bridge with the second arginine of the FIC motif (R(2): R147, R209, or R118, respectively).

The position and orientation of the triphosphate is defined by a multitude of specific interactions (Fig. 3A–C). In all structures, the α- and β-phosphate moieties form four H-bonds with the four exposed backbone amide groups of the compound anion binding nest [17] at the N-terminal end of helix α5. In addition, the first arginine of the signature motif (R(1): R144, R206, or R115, respectively) forms a salt-bridge with the β-phosphate involving two H-bonds and the asparagine of the motif (N142, N204, or N113, respectively) interacts with a non-bridging oxygen of the α-phosphate.

In all the structures, a magnesium ion is present albeit with high B-factor for NmFic$_{E186G}$ (63 Å^2). The metal bridges the α- and β-

phosphate and is coordinated in addition by the conserved D/E residue of the Fic signature motif in VbhA$_{E24G}$/VbhT(FIC) and SoFic$_{E73G}$ (E140, D202, respectively). It is particularly well resolved in the former structure where three well-defined water molecules complete its octahedral coordination shell (Fig. 3A). Interestingly, the divalent cation is observed only in the adenylylation competent complexes, but not in the wild-type complexes. Indeed, magnesium is indispensable for Fic mediated ATase activity (data not shown) and is probably important for fine-tuning of the α- and β-phosphate orientation within the compound anion binding nest and for stabilization of the transition state.

Overall, the three structures display a unique mode of ATP binding that can be attained only in the mutants, since the γ-phosphate effectively adopts the position that is taken by the inhibitory glutamate in the wild-type proteins (Fig. 4). Most relevantly, the reorganization of the triphosphate in the binding site results in a α-phosphate orientation that is now prone for in-line attack by an incoming target side-chain (Fig. 3D). Clearly, the conservation of this binding mode across the FIC classes shows that it is essential for FIC function.
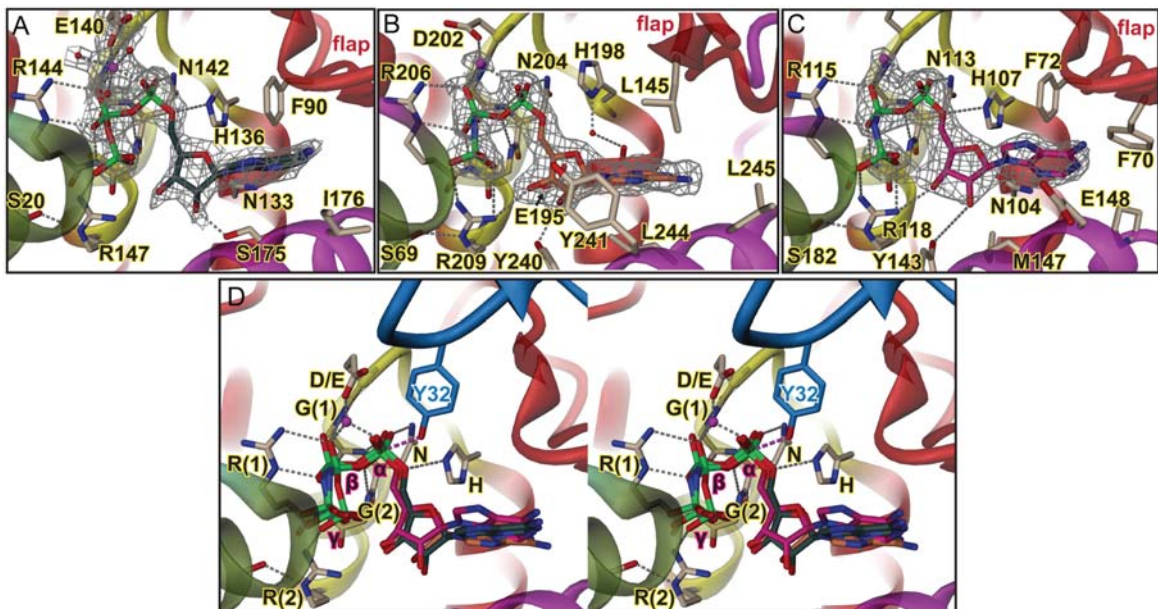
Figure 3. Crystal structures of E->G mutated Fic proteins representing classes I to III in complex with substrate or substrate analog. A, VbhA$_{E24G}$/VbhT(FIC) in complex with ATP/Mg^{2+}; B, SoFic$_{E73G}$, C, NmFic$_{E186G}$, both in complex with AMPPNP/Mg^{2+}. Representation as in Fig. 2 with magnesium ions shown as magenta spheres. The 2Fo-Fc simulated annealing omit maps covering the nucleotide/Mg^{2+} ligands are contoured at 1.1 σ. D, Stereo view of the superposition of the ligand structures shown in panels B and C onto the VbhA$_{E24G}$/VbhT(FIC) complex (same as in panel A). Note that the nucleotides of the various complexes are distinguished by their carbon color (VbhA$_{E24G}$/VbhT(FIC) ATP in green, SoFic$_{E73G}$ AMPPNP in orange and NmFic$_{E186G}$ AMPPNP in pink). The residues of the HxFx(D/E)GNGRxxR signature motif are labeled as in Fig. 2C with the phenylalanine not shown. Also shown is the modifiable hydroxyl side-chain Y32 of Cdc42 (blue) after superposition of the IbpA(FIC2)/Cdc42 complex [4] onto VbhA$_{E24G}$/VbhT(FIC). For the superposition, only the Fic active site loops were used. The α-phosphate moieties appear well-suited for in-line attack of the target hydroxyl group (broken line in magenta).

doi:10.1371/journal.pone.0064901.g003

110
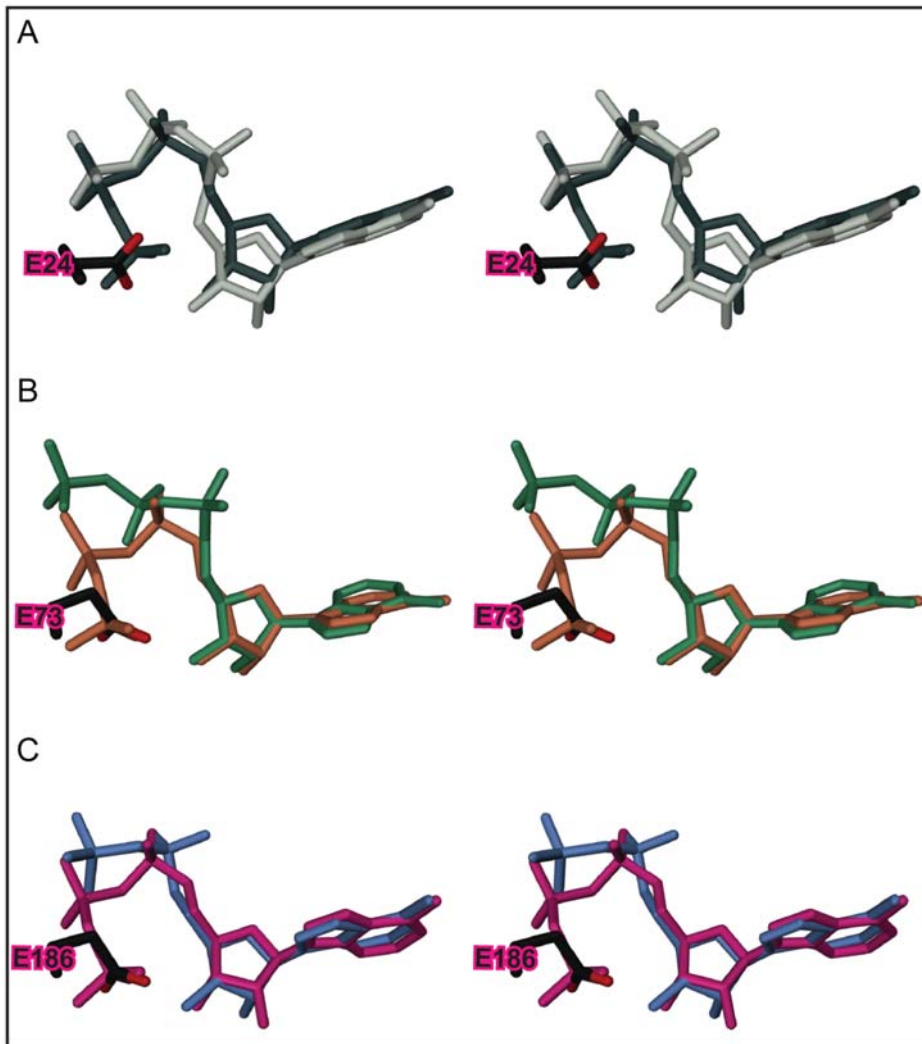
Figure 4. Comparison of triphosphate nucleotide structures as bound to wild-type and E->G mutated Fic proteins from class I to III. Stereo views of the ligand structures after superposition of the FIC domains (not shown). Also shown is the inhibitory glutamate of the wild-type structures. A, ATP as bound to VbhA/VbhT wild-type (white) and the E24G mutant (dark green). B, ATP and AMPPNP as bound to SoFic wild-type (green) and the E73G mutant (orange), respectively. C, AMPPNP as bound to NmFic (blue) and the E186G mutant (pink). Note that the AMPPNP γ-phosphate in NmFic is found disordered [8] and therefore not shown.
doi:10.1371/journal.pone.0064901.g004

Target Registration to the FIC Active Site

The conservation of the FIC active site and the ATP substrate binding mode prompts for a precise alignment of the incoming side-chain hydroxyl with the scissile Pα-O3α bond. The beta-hairpin flap partly covering the active site appears to represent a "target dock" that ensures this precise positioning of the target backbone stretch immediately following the modifiable hydroxyl side-chain and thus registers the side-chain to the active site as has been proposed before (2). This was deduced mainly from the only known Fic protein/target complex structure IbpA(FIC)/Cdc42 [4] where the AMPylated Y32 of Cdc42 is part of a segment (switch 1 loop) in extended conformation and complements inter-molecularly the β-hairpin of the flap (Fig. 5A).

This notion is further corroborated by the structure of the wild-type VbhA/VbhT(FIC) complex presented here that revealed additional density close to the flap above the active site (Fig. 5B). This was interpreted as a four residue peptide in extended conformation that is associated antiparallely to the edge of the two-stranded β-hairpin of the flap via three main chain-main chain H-bonds. Location and side-chain densities are consistent with the peptide representing residues 203 to 206 of a symmetry mate (note that the ordered part of the VbhT(FIC) construct ends with residue F197). Very similarly, peptide density is present at an equivalent location in the A-chain of SoFic$_{E73G}$ and could be attributed to the N-terminus (residues 0 to 3) of a symmetry related B-chain as also reported for the isomorphous crystal structure of wild-type SoFic (Fig. 5C) [16].
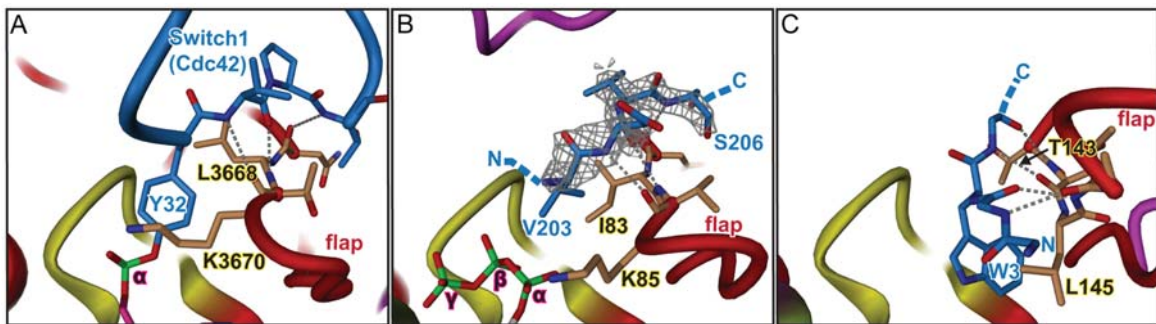
111

Figure 5. Sequence independent registration of peptide or target protein to the FIC flap. The bound peptide/protein segment (blue) and the target dock (brown) are shown in full. Main chain-main chain H-bonds are depicted as stippled lines. A, Product complex of IbpA(Fic2) with Cdc42 target [4]. Tyrosine 32 from the switch1 region of Cdc42 is adenylylated. B, VbhA/VbhT(FIC) complexed with residues 203 to 206 of a symmetry related molecule. The 2Fo-Fc simulated annealing omit map covering the residues 203 to 206 is contoured at 1.1 σ. Note that the preceding 7 residues are disordered and not shown. C, SoFic complexed with residues 0 to 4 of a symmetry related molecule (PDB 3EQX) [16]. The side-chains of residues 0, 1 and 4 are disordered and not displayed for clarity reason. Note that Y32 in panel A, V203 in panel B and W3 in panel C are in equivalent positions.

doi:10.1371/journal.pone.0064901.g005

Comparison of Figures 5A–C suggests that a tyrosine instead of the valine in position 203 of the symmetry related VbhT(FIC) chain or of the tryptophan in position 3 of the symmetry related SoFic chain would indeed be well poised to attack the ATP α-phosphate. Furthermore, it has been shown for IbpA that the side-chains of the target dock residues Leu3668 and Lys3670 form a hydrophobic clamp that fix the target tyrosine side-chain (Fig. 5C) [4]. Side-chains of residues I83 and K85 in VbhT(FIC) and residues T143 and L145 in SoFic, that hold the valine and tryptophan, respectively, may in a similar way clamp down the modifiable side-chain (Fig. 5A–B). Taken together, these observations show that the flap has propensity for peptide binding as it is well known for exposed beta-sheet edges in other proteins [18,19,20] and ensures productive alignment of the target hydroxyl side-chain with the bound ATP substrate.

Probably, sequence independent positioning of the backbone flanking the modifiable target residue confers an evolutionary advantage. While exposed loops in extended conformation of many proteins may easily dock to the flap, other parts of the enzyme would confer target affinity and specificity (as seen in the IbpA(FIC)/Cdc42 complex [4]) that were free to adopt during evolution without compromising on the catalytic mechanism. Notably, peptide registration to the active site via main-chain interactions is known also for serine proteases [21] and protein kinases [22,23].

Discussion

The vast majority of Fic proteins are characterized by a conserved HxF[D/E]GNGRxxR active site motif and catalyses adenylylation, an enzymatic activity that involves nucleophilic attack of a target hydroxyl group onto the α-phosphate of ATP. Productive AMP transfer therefore relies on the proper juxtaposition of the reaction partners. The inhibition-relieved (E->G) mutant structures of Fic proteins from the three distinct classes shed light on the importance of the active site [D/E]GNGRxxR residues to enable catalytically competent ATP substrate binding. Indeed, in the three classes, these residues, by way of a large hydrogen-bonding network, enable a unique mode of ATP binding to orientate favorably the α-phosphate relative to the target side-chain hydroxyl group (Fig. 3). The latter is registered to

the FIC active site in-line with the scissile Pα-O3α bond *via* sequence-independent main chain-main chain interactions with the target dock at the edge of the FIC flap (Fig. 5). Thus, the FIC active site and the target dock are two indivisible structural elements that have been exposed to high functional constraints to ensure productive catalysis. Fic proteins with degenerated active site signature motifs and/or devoid of a flap-like structure are likely to have adopted new functions.

In Fic proteins of the three inhibition classes, the inhibitory glutamate plays the same role. It out-competes the γ-phosphate for binding to arginine R(2) of the FIC signature motif (Fig. 4). This results in an α-phosphate orientation that does not permit an attack of the incoming target side-chain hydroxyl group. Interestingly, though the active sites are structurally well conserved, the nucleotide triphosphates show variation in their binding to the Fic proteins of the three classes (Fig. 2). This is in contrast to the uniform binding mode found in the inhibition relieved mutants (Fig. 3) and shows that the mode of ATP binding to the inhibited enzyme *per se* was not under evolutionary constraints.

Knowledge of the universal catalytic and inhibitory mechanism of Fic mediated AMP transfer will now pave the way for further studies towards the physiological roles of Fic proteins and particularly the identification of their protein targets. It may also prompt rational structure based design of small molecule inhibitors targeting the ATP binding pocket or novel peptides that mimic the inhibitory helix to neutralize bacterial virulence factors which kill their host *via* uncontrolled Fic-mediated adenylylation activity.

Author Contributions

Conceived and designed the experiments: AG CD TS. Performed the experiments: AG FVS. Analyzed the data: AG TS. Wrote the paper: AG CD TS.

112

References

1. Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, et al. (2009) The fic domain: regulation of cell signaling by adenylylation. Molecular cell 34: 93–103.
2. Yarbrough ML, Li Y, Kinch LN, Grishin NV, Ball HL, et al. (2009) AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. Science 323: 269–272.
3. Palanivelu DV, Goepfert A, Meury M, Guye P, Dehio C, et al. (2011) Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. Protein science : a publication of the Protein Society 20: 492–499.
4. Xiao J, Worby CA, Mattoo S, Sankaran B, Dixon JE (2010) Structural basis of Fic-mediated adenylylation. Nature structural & molecular biology 17: 1004–1010.
5. Luong P, Kinch LN, Brautigam CA, Grishin NV, Tomchick DR, et al. (2010) Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. The Journal of biological chemistry 285: 20155–20163.
6. Mukherjee S, Liu X, Arasaki K, McDonough J, Galan JE, et al. (2011) Modulation of Rab GTPase function by a protein phosphocholine transferase. Nature 477: 103–106.
7. Kinch LN, Yarbrough ML, Orth K, Grishin NV (2009) Fido, a novel AMPylation domain common to fic, doc, and AvrB. PloS one 4: e5818.
8. Engel P, Goepfert A, Stanger FV, Harms A, Schmidt A, et al. (2012) Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. Nature 482: 107–110.
9. Schulein R, Guye P, Rhomberg TA, Schmid MC, Schroder G, et al. (2005) A bipartite signal mediates the transfer of type IV secretion substrates of Bartonella henselae into human cells. Proceedings of the National Academy of Sciences of the United States of America 102: 856–861.
10. Kabsch W (2010) Xds. Acta crystallographica Section D, Biological crystallography 66: 125–132.
11. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. Journal of applied crystallography 40: 658–674.
12. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta crystallographica Section D, Biological crystallography 66: 486–501.
13. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure

solution. Acta crystallographica Section D, Biological crystallography 66: 213–221.
14. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica Section D, Biological crystallography 53: 240–255.
15. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta crystallographica Section D, Biological crystallography 66: 12–21.
16. Das D, Krishna SS, McMullan D, Miller MD, Xu Q, et al. (2009) Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gi|24375750) from Shewanella oneidensis MR-1 at 1.6 A resolution. Proteins 75: 264–271.
17. Watson JD, Milner-White EJ (2002) A novel main-chain anion-binding site in proteins: the nest. A particular combination of phi,psi values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions. Journal of molecular biology 315: 171–182.
18. Hill CP, Yee J, Selsted ME, Eisenberg D (1991) Crystal structure of defensin HNP-3, an amphiphilic dimer: mechanisms of membrane permeabilization. Science 251: 1481–1485.
19. Nassar N, Horn G, Herrmann C, Scherer A, McCormick F, et al. (1995) The 2.2 A crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue. Nature 375: 554–560.
20. Doyle DA, Lee A, Lewis J, Kim E, Sheng M, et al. (1996) Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. Cell 85: 1067–1076.
21. Wilmouth RC, Clifton IJ, Robinson CV, Roach PL, Aplin RT, et al. (1997) Structure of a specific acyl-enzyme complex formed between beta-casomorphin-7 and porcine pancreatic elastase. Nature structural biology 4: 456–462.
22. Hubbard SR (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. The EMBO journal 16: 5572–5581.
23. Yang J, Cron P, Good VM, Thompson V, Hemmings BA, et al. (2002) Crystal structure of an activated Akt/protein kinase B ternary complex with GSK3-peptide and AMP-PNP. Nature structural biology 9: 940–944.

113

3.4   Unpublished results

Further structural investigations on the *B. henselae* effector BepA

3.4.1 Further structural investigations on the *B. henselae* effector BepA

3.4.1.1 Introduction

In Chapter 3.1[140] (Research article I), we investigated the mechanism of Fic-catalyzed adenylylation based on the complex structure of IbpA-Fic2/Cdc42 and the structure of a tailored construct of *Bartonella henselae* BepA (FIC-OB). It turned out that this construct had a pyrophosphate, the side product of the adenylylation reaction bound to the FIC active site cavity. From this data, we proposed a catalytic mechanism for the AMP transfer, where we modeled the ATP substrate based on the pyrophosphate location and the structure of the adenylylated IbpA-Fic2/Cdc42 complex[88]. But for a complete and unbiased understanding of the mechanism the structure of the FIC-ATP complex would be most desirable. To this end, at the beginning of my PhD, I extensively attempted to obtain BepA in complex with ATP by following experimental procedures developed by Dinesh Palanivelu (Chapter I (Research article I)). Soaking and Co-crystallization trials with ATP and non-hydrolyzable ATP analogue did not succeed and putative sulphate ions were constantly found within the anion binding nest. This is most likely due to the high salt concentration of the crystallization condition (1.5M $(NH_4)_2SO_4$) that prevented nucleotide binding. Therefore, I decided to slightly vary the protein buffer and went for a new round of crystallization condition screening to hopefully obtain crystals of a BepA/ATP complex.

3.4.1.2 Materials and Methods

A BepA$_{H159A}$ construct comprising residues 10-303 (Fic-OB) and a C-terminal His$_6$ tag was expressed and purified as described in Chapter 3.1[140]. At the end of the purification step, the protein was eluted in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl and 2 mM β-mercaptoethanol. The protein was then incubated with 5 mM of AMP-CPP (α,β-Methyleneadenosine 5′-triphosphate, a non-hydrolyzable ATP analog) and 10 mM MgCl2 for 1hour.

Crystals were obtained at 20°C using the hanging-drop vapor diffusion method upon mixing 1 μl protein solution with 1 μl reservoir solution. The reservoir solution was composed of 25% (w/v) PEG 3350 and 0.2 M ammonium sulfate (non- buffered). For data collection, crystals were transferred to 25% (w/v) PEG 3350, 0.2 M ammonium sulfate, and 5% v/v PEG 400 and frozen in liquid nitrogen. Diffraction data were collected on beam-line X06SA (PXIII) of the Swiss Light Source (λ = 1.0 Å) at 100 K on a MAR CCD detector, processed

using MOSFLM[141], and scaled with SCALA[142]. The structure was solved by molecular replacement using Phaser[143] with the BepA fragment (PDB code 2VZA) as the search model. Several rounds of iterative model building and refinement were performed using Coot[144] and PHENIX[145] respectively. 5% of the data were excluded from refinement and used for cross-validation. The geometry of the final model was assessed using MolProbity[146] with >99% of the residues in the core and allowed regions of the ramachandran plot. Data collections and refinement statistics are summarized in Table 1. The figure was generated with Dino (A. Philippsen unpublished, http://www.dino3d.org).

Table 1: Data collection and refinement statistics for BepA$_{H159A}$ (10-303).

Protein	BepA$_{H159A}$ (10-303)
Ligand	AMP-CPP
Data collection	
Space group	P2$_1$
Cell dimensions	
a, *b*, *c* (Å)	42.3, 50.8, 66.0
α, β, γ (°)	90.0, 96.6, 90.0
Resolution (Å)	28.0 - 1.8(1.9 - 1.8) *
R_{sym} or R_{merge}	5.7(30.6)
$I/\sigma I$	13.7(4.0)
Completeness (%)	99.6(99.0)
Redundancy	3.8(3.8)
Refinement	
Resolution (Å)	15-1.70
No. reflections	58,005(5,445)
R_{work}/ R_{free}	24.5/28.8
Average B (Å^2)	30.8
No. atoms	
Protein	2334
Ligand/ion	-
Water	266
R.m.s deviations	
Bond lengths (Å)	0.020
Bond angles (°)	1.6

* Values for the highest resolution shell are shown in parenthesis.

3.4.1.3 Results

The crystal structure of the BepA construct comprising the FIC-OB module was solved at 1.7 Å and yielded a new monoclinic ($P2_1$) crystal form compared to the orthorombic ($P2_12_12_1$) of the structure described in Chapter 3.1[140] (Research article I) (PDB 2JK8). Comparison of both structures does not reveal major changes and both proteins display the same overall topology with a Cα-RMS deviation (RMSD) of 1.1 Å (DaliLite). Whilst electron density at the FIC active site is present, neither AMP-CPP nor any obvious molecule from the crystallization condition could be unambiguously assigned to it (Fig. 1). At the time of writing, the refinement was not completed. The high free R-factor (28.8%) is explained by unassigned and yet unidentified strong positive density peaks located in close proximity to histidine and glutamate side-chains. Whether these electron density blops represent ions or modified histidine is still unclear and will need further investigations.
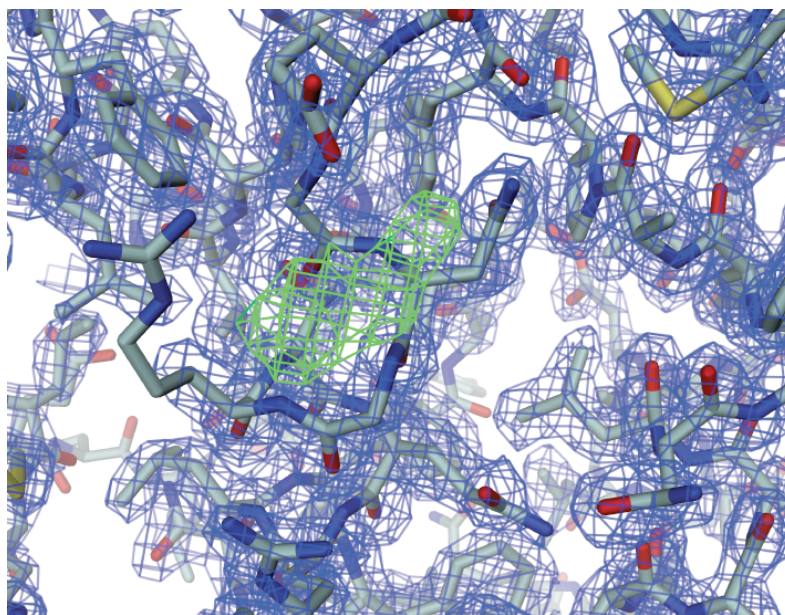
Figure 1: Active site of BepA. The 2Fo-Fc map calculated at 1.7 Å resolution and contoured at 1.2 σ is shown. The unassigned density is represented by green lines.

3.4.1.4 Conclusion

Although this new structure did not give any further valuable information on the putative mechanism of Fic-mediated adenylylation, the higher resolution might be pertinent for future structural investigations on Fic BepA and its interactions with the yet unknown target protein.

The FIC domain of *B. rochalimae* Bep1 effector protein adenylylates Rac1 and is inhibited by BiaA

3.4.2 The FIC domain of *B. rochalimae* Bep1 effector protein adenylylates Rac1 and is inhibited by BiaA

3.4.2.1 Introduction

Bartonella rochalimae belongs to lineage 3 of the *Bartonella* genus and has been shown to infect dogs and foxes. As the other *Bartonella* species, *B. rochalimae* utilizes a T4SS to translocate an arsenal of effectors (Beps) into host cells and manipulate host cell signaling pathways during infection. *B. rochalimae* species has not been deeply investigated and no phenotype has yet been associated with any of the effectors. Nevertheless, Alexander Harms (Biozentrum) unambiguously demonstrated that the effector Bep1, which possess the typical FIC-BID domain architecture, is capable of transferring AMP onto tyrosine32 from the switch I region of Rac1 GTPase that is likely to be the physiological target in the eukaryotic host cell[147]. Surprisingly, Bep1 showed high specificity towards its target since no adenylylation has been detected for the closely related Rho subfamily proteins Cdc42 and RhoA. Additionally, he could show that BiaA (bep-interacting antitoxin A), the protein encoded upstream of Bep1, inhibits Bep1-catalysed adenylylation of Rac1[147]. Here, I performed structural investigations on Bep1 in complex with BiaA to unravel the mechanism of BiaA-mediated inhibition of Bep1 and initiated biochemical assays that will help in elucidating the mode of recognition of Rac1 target.

3.4.2.2 Materials and Methods

The full-length biaA gene that codes for the small ORF directly upstream of bep1 gene and part of the bep1 gene from Bartonella rochalimae encoding the FIC domain (amino acid residues 13-229) were PCR-amplified from genomic DNA. The PCR products for biaA and the fragment of bep1 were cloned into the pRSF-Duet1 vector. The biaA gene was cloned into the multiple cloning site 1 (MCS1) and the vbhT construct with an additional C-terminal His6 affinity tag in MCS2.

The pRSF-Duet1 vector containing *biaA* and the *bep1* construct (denominated Bep1(FIC)) was transformed into *E.coli* BL21 (DE3). The construct was expressed and purified as described in Chapter 3.2 (Research article II) for VbhA/VbhT(FIC) 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.

A crystal was obtained at 4°C using the hanging-drop vapor diffusion method upon 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% v/v PEG 400. For data collection, crystal was frozen in liquid nitrogen without cryoprotectant. Diffraction data were collected on beam-line X06SA (PXIII) of the Swiss Light Source (λ = 1.0 Å) at 100 K on a MAR CCD detector. Data were processed and the structure solved as described for BepA$_{H159A}$ (10-303) in Chapter 3.4.1, using the VbhA/VbhT(FIC) structure (PDB 3SHG) as the search model. Data collections and refinement statistics are summarized in Table 1

Adenylylation assays were performed as described for HYPE$_{FIC}$ in Chapter 3.2 (Research article II) by mixing 300 ng of pure Bep1 protein with 1.2 μg GTPases.

Table 2. Data collection and refinement statistics for BiaA/Bep1(FIC).

Protein	BiaA/Bep1(FIC)
Data collection	
Space group	$P4_3 2_1 2$
Cell dimensions	
a, *b*, *c* (Å)	73.1, 73.1, 130.0
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	51.7 - 1.7(1.8 - 1.7) *
R_{sym} or R_{merge}	11.7(47.6)
$I/\sigma I$	21.0(3.0)
Completeness (%)	99.5(96.4)
Redundancy	18.9(4.7)
Refinement	
Resolution (Å)	15-1.70
No. reflections	38,495(3,165)
R_{work}/R_{free}	16.0/21.0
Average B (Å^2)	25.3
No. atoms	
Protein	2164
Ligand/ion	-
Water	266
R.m.s deviations	
Bond lengths (Å)	0.010
Bond angles (°)	0.9

* Values for the highest resolution shell are shown in parenthesis.

3.4.2.3 Results

The FIC domain is sufficient to adenylylate Rac1. Alexander Harms (Biozentrum, Basel) showed that full-length Bep1 is able to specifically adenylylate Rac1. To test whether the FIC domain on its own has the potential to adenylylate this particular GTPase, the enzymatic activity of Bep1(FIC) was assayed *in-vitro* using α-^{32}P-ATP and Rac1 or Cdc42 as substrates (Fig. 1). Bep1(FIC) affords efficient transfer of AMP onto Rac1 and the activity is repressed by mutation of histidine to alanine (Bep1(FIC/H158A)) in the conserved FIC motif HxFx[D/E]GNGRxxR. Interestingly, the constitutively active form (Rac1 Q61L) does not get adenylylated and suggests that Bep1(FIC) exclusively interacts with wild-type GDP-bound form of the GTPase. However, it cannot be excluded that the Q61L point mutant that prevents

endogenous and GAP-stimulated GTPase activity of Rac1 might artificially impede Bep1(FIC) binding. Finally, the absence of radioactive signal on Cdc42 infers that structural element(s) responsible for the specificity towards Rac1 are confined to the FIC domain.
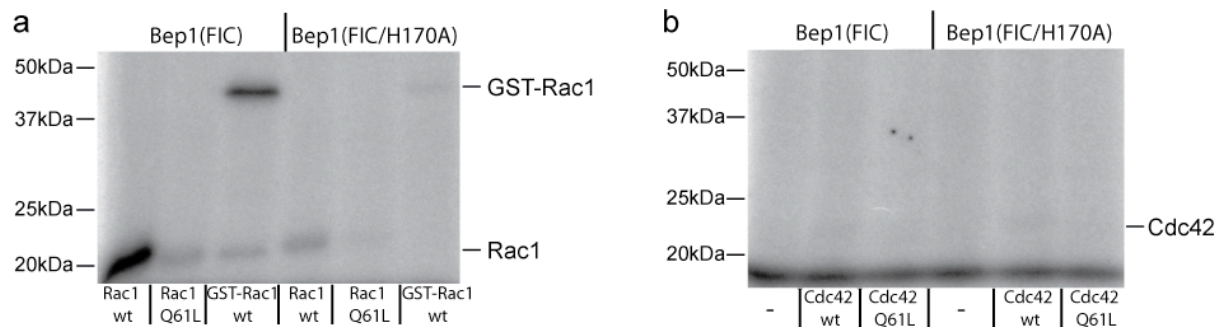
Figure 1: Bep1(FIC) specifically adenylylates Rac1. Autoradiography of an SDS-gel after incubation of wild-type and catalytic histidine mutant of Bep1(FIC) with α-^{32}P-ATP in presence of wild-type or constitutively active (Q61L) Rac1 (**a**) or Cdc42 (**b**). Wild-type Bep1(FIC) catalyses adenylylation of Rac1 but not Cdc42. Only weak adenylylation is observed for the constitutively active variants of Rac1 (Rac1Q61L).

Comparison of BiaA/Bep1(FIC) and VbhA/VbhT(FIC) reveals high structural and functional similarities. The high resolution structure (1.7 Å) of BiaA/Bep1(FIC) from *B. rochalimae* shows high level of structural similarity (RMSD of 1.8 Å) and 30% sequence identity with VbhA/VbhT(FIC) (PDB code 3SHG, see Chapter 3.2 (Research article II)) from *B. schoenbuchensis* (Fig. 2). Bep1(FIC) contains all elements characteristic for an adenylylation-competent Fic protein. Besides the typical α-helical Fic topology, the protein displays the β1-β2 hairpin feature for target positioning along with the highly conserved active site motif essential to mediate adenylylation. Noteworthy, superposition of Bep1(FIC) with VbhT(FIC) indicates an almost strict conservation of residues involved in ATP interaction, suggesting an equivalent substrate binding mode (Fig. 2a and 2d). Notwithstanding the similar overall topology, a major structural deviation is evident. Bep1(FIC) presents an insertion in the loop preceeding the β1-β2 hairpin. Moreover, the hairpin turns out to be elongated (Fig. 2a and 2c).

Analogously to VbhA, the BiaA antitoxin is folded to three antiparallel helices that wind around helix α1 of Bep1(FIC) with the N-terminal helix inserted in the Fic protein core

(Fig. 2c). This helix (α_{inh}) harbors the conserved inhibitory motif [S/T]xxxE[G/N] with Thr29 and Glu33 forming H-bonds with the Bep1(FIC) residue Arg166 (Fig. 2b-d). Thus, Glu33 intrudes in the Fic active site and may compete for ATP binding.

a

b

c

BiaA/Bep1(FIC)   VbhA/VbhT(FIC)

d

Figure 2: (*legend on next page*)

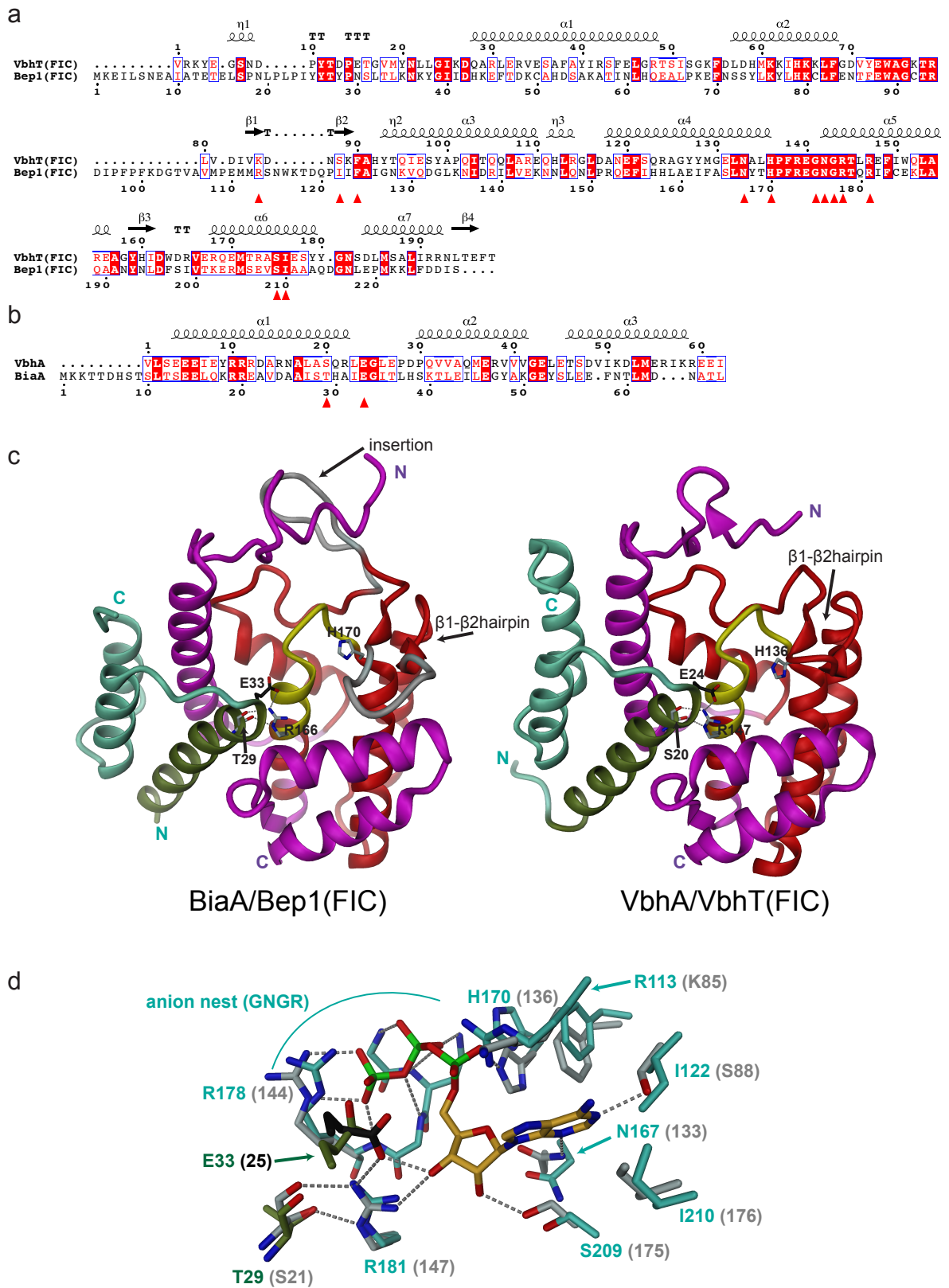**Figure 2**: Structure based sequence alignments of VbhT(FIC) with Bep1(FIC) (**a**) and VbhA with BiaA (**b**). (**a**) The two proteins share 30% sequence identity and a strictly conserved Fic motif located between helices α4 and α5. The major difference resides between helices α2 and α3 where Bep1(FIC) harbors loop extensions (see also panel C). The red triangles highlights residues conserved between the two proteins that are involved in ATP binding (see panel d). (**b**) The two proteins share 30% identity and possess the conserved inhibitory motif [S/T]xxE[G/N] with S/T and E highlighted with red triangles. (**c**) Structure comparison of BiaA/Bep1(FIC) (left panel) and VbhA/VbhT (PDB code 3SHG, right panel). Color code is the following: red, core of Fic domain as defined by Pfam[86]; yellow, HxFxxGNGRxxR signature motif; green, antitoxin; magenta, remainder of structure. BiaA/Bep1(FIC) is highly similar to VbhA/VbhT but with further loop extensions around the β-hairpin region situated between helices α2 and α3 (colored in grey). (**d**) Superposition of active-sites from proteins shown in panel C that highlights the conservation of residues involved in ATP binding. turquoise, Bep1(FIC); grey VbhT; green, BiaA; black, inhibitory glutamate from VbhA. H-bonds are marked by stippled grey lines. Residues corresponding to BiaA/Bep1(FIC) are written in turquoise with the equivalents from VbhA/VbhT(FIC) in brackets. The location of the glutamate from the inhibitory motif is similar in both structures supporting a conserved mechanism of active-site obstruction.

3.4.2.4 Conclusion

Biochemical analysis performed on Bep1 revealed that the FIC module is sufficient to specifically recognize and adenylylate the small Rho GTPase Rac1. Interestingly, Bep1(FIC) shows insertions in the vicinity of the β1-β2 hairpin. Expansion in this particular region has already been seen in AvrB where it mediates binding to the plant immune response regulator RIN4[98,101]. Therefore, these features may represent potential target recognition modules and enable Bep1(FIC) to discriminate between small Rho GTPases. Remarkably, these elements are also found in the structure of the FIC domain of the *B. henselae* effector BepA[140] (PDB 2JK8) and might be therefore characteristic of eukaryotic host-interacting *Bartonella* effectors. Furthermore, the elongated β-hairpin forms a cap on the Fic active site and may prevent futile ATP hydrolysis[140].

These biochemical findings provide groundwork for future structural investigations on the Bep1(FIC)/Rac1 complex structure aiming to decipher the molecular recognition and specific interactions of the toxin with its host cell target.

The complex structure of BiaA/Bep1(FIC) shows high structural homology with the putative VbhT/A toxin-antitoxin system of *B. schoenbuchensis* and suggests a conserved mechanism of FIC domain inhibition by active site obstruction. Whereas the plasmid encoded VbhA peptide might exert a function similar to the classical bacterial antitoxins, the biological role of the chromosomal BiaA is yet unknown and putative functions will be discussed in the general discussion Chapter.

The FIC domain of VbhT from *B. schoenbuchensis* might harbor both AMP transferase and kinase activity

3.4.3 The FIC domain of VbhT from *B. schoenbuchensis* might harbor both AMP transferase and kinase activity

3.4.3.1 Introduction

In Chapter 3.2 (Research article II), we showed that the Fic domain of VbhT from *B. schoenbuchensis* catalyses adenylylation of a 80kDa *E. coli* target protein[148]. This modification results in bacterial filamentation and growth arrest. The toxic activity of VbhT is repressed by the protein VbhA encoded directly upstream the locus for VbhT. VbhA possesses a conserved inhibitory motif [S/T]xxE[G/N] in which the glutamate impedes proper ATP binding to the FIC active site[148]. Furthermore, by solving the structure of VbhA$_{E24G}$/VbhT(FIC) in complex with ATP, we structurally demonstrated that mutation of this glutamate to glycine relieves inhibition in that it allows competent ATP binding[149] (Chapter 3.3, Research article III). Here, I present a new structure of VbhA$_{E24G}$/VbhT(FIC) in complex with ATP that shows a novel ATP binding mode.

3.4.3.2 Materials and Methods

The plasmid pFVS0065 containing VbhA$_{E24G}$/VbhT(FIC) has been cloned, expressed and purified as described in Chapter 3.3 (Research article III).

A crystal was obtained at 20 °C using the hanging-drop vapor diffusion method upon mixing 1 μl protein solution with 1 μl reservoir solution. The reservoir solution was composed of 12% w/v PEG 4000 and 0.1 M MES (pH 6.5). The protein was then incubated with 10 mM ATP and 10 mM MgCl2 for 1 hour. For data collection, the crystal was frozen in liquid nitrogen with reservoir buffer supplemented with 20% glycerol as cryoprotectant. Diffraction data were collected at beam-line X06SA (PXIII) of the Swiss Light Source ($\lambda = 1.0$ Å) at 100 K on a PILATUS 2M detector, and scaled and integrated with XDS/XSCALE[150]. The structure was refined with the program PHENIX[145] using starting phases from the VbhA$_{E24G}$/VbhT(FIC)$_{H136A}$/AMPPNP structure (see Chapter 3.3, Research article III), from which water molecules, magnesium ion and AMPPNP had been removed. The remainder of the molecule was traced manually with Coot[144] followed by full refinement using PHENIX[145].

Table 1: Data collection and refinement statistics for VbhA$_{E24G}$/VbhT(FIC).

Protein	VbhA$_{E24G}$/VbhT(FIC)
Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	106.4, 41.5, 74.3
α, β, γ (°)	90.0, 121.7, 90.0
Resolution (Å)	31.6 - 1.5(1.6 - 1.5) *
R_{sym} or R_{merge}	4.1(43.5)
$I/\sigma I$	18.2(2.7)
Completeness (%)	98.6(95.2)
Redundancy	3.6(3.6)
Refinement	
Resolution (Å)	15-1.5
No. reflections	44,097(4,087)
R_{work}/ R_{free}	15.1/20.8
Average B (Å^2)	21.6
No. atoms	
Protein	1821
Ligand/ion	1 ATP, 1MG
Water	291
R.m.s deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.2

• Values for the highest resolution shell are shown in parenthesis.

3.4.3.3 Results

The crystal structure of VbhA$_{E24G}$/VbhT(FIC) in complex with ATP/Mg^{2+} was solved at 1.5 Å resolution and is isomorphous to the wild-type VbhA/VbhT(FIC) structure[148] (Table S1, see Chapter 3.2 (Research article II)). Electron density corresponding to an ATP/Mg^{2+} complex is clearly visible in the active site (Fig. 1a). The divalent cation has a perfect octahedral coordination sphere mediated by three water molecules, the carboxyl side-chain of a conserved glutamate (E140) from the FIC active site motif and two non-bridging oxygens of the β- and γ-phosphates of ATP. Comparison with the previously reported VbhA$_{E24G}$/VbhT(FIC) structure in complex with ATP[149] (see Chapter 3.3 (Research article III)) reveals striking results (Fig. 1). Whereas the adenosine moiety adopts equivalent position in the two structures, the triphosphate is found reoriented (Fig. 1a). The position in the

compound nest adopted by the α-phosphate of ATP in the previously reported VbhA$_{E24G}$/VbhT(FIC) structure (Fig. 1b) is taken by the γ-phosphate of ATP in the new VbhA$_{E24G}$/VbhT(FIC) structure (Fig. 1a). In this particular conformation, the ATP α-phosphorous interacts with the second arginine of the FIC active site motif and is therefore not in an adenylylation-competent conformation. On the contrary, the γ-phosphate appears well suited to undergo in-line attack induced by the target hydroxyl group, resulting in phosphate transfer onto the target (Fig. 1a).
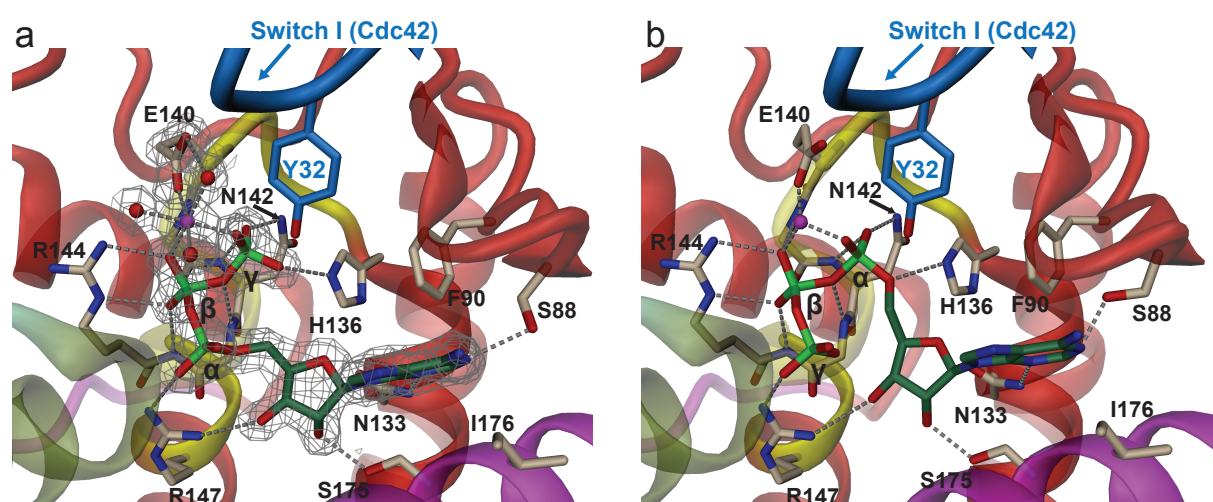
Figure 1: ATP can bind in two distinct orientations to the substrate binding site of VbhT(FIC) suggesting a dual AMP transferase and kinase activity for VbhT(FIC). ATP is represented in full and colored in dark green. Magnesium ion and water molecules are shown as magenta and red spheres, respectively. Residues important for ATP/Mg2+ binding are shown in full. The modifiable Y32 of the Cdc42 switch I-loop is shown in blue in a position obtained after superposition of the active site loops of VbhT(FIC) and the IbpA(FIC2)/Cdc42 complex[148]. (**a**) Active site of VbhA$_{E24G}$/VbhT(FIC) in complex with ATP/Mg^{2+}. Also shown is the 2Fo-Fc map contoured at 1.2 σ and covering the ligand. The ATP is found oriented such that the γ-phosphate is in an appropriate position allowing in-line attack of a target side-chain to accomplish phosphate transfer.

(**b**) Active site of VbhA$_{E24G}$/VbhT(FIC) in complex with ATP/Mg^{2+} (PDB code: 3ZCB) showing adenylylation-competent nucleotide binding[149] (see also Chapter 3.3 (Research article III)).

3.4.3.4 Conclusion

The two distinct ATP binding modes suggest that VbhT(FIC) might be a bifunctional enzyme that can utilize one active site to post-translationally modify target proteins by either adenylylation or phosphorylation. Further biochemical data are needed to support this hypothesis.

4 Summary

Throughout my PhD, I solved various structures of FIC-domain containing proteins that were essential for a thorough structural understanding of their adenylylation activity and regulation.

The vast majority of Fic proteins are characterized by a conserved HxF[D/E]GNGRxxR active site motif and catalyses adenylylation, an enzymatic activity that involves nucleophilic attack of a target hydroxyl group onto the α-phosphate of ATP. Productive AMP transfer therefore relies on the proper juxtaposition of the reaction partners. The active site [D/E]GNGRxxR residues are critical to enable catalytically competent ATP substrate binding. Indeed, these residues, by a way of a large hydrogen-bonding network, enable a unique mode of ATP binding to orientate favorably the α-phosphate relative to the target side-chain hydroxyl group. The latter is registered to the FIC active site in-line with the scissile Pα-O3α bond via sequence-independent main chain-main chain interactions with the target dock at the edge of the FIC flap. Thus, the FIC active site and the target dock are two indivisible structural elements that have been exposed to high functional constraints to ensure productive catalysis. We believe that this catalytic mechanism (summarized in Figure 1) can be generalized and extrapolated to all adenylylation-competent Fic proteins.

The structure of VbhA/VbhT(FIC) of *Bartonella schoenbuchensis* brought into light an unexpected and widely spread mode of regulation used by adenylylation-competent Fic proteins to regulate intrinsic signaling processes. The proteins present an alpha helix (α_{inh}) characterized by a conserved [S/T]xxxE[G/N] motif to the FIC active site and perturb ATP binding. Intriguingly, α_{inh} can be part of the Fic protein itself as a N- or C-terminal extension or provided by an additional antitoxin. Thus, we suggested a tripartite classification scheme to group Fic proteins, according to the position of the inhibitory helix along the polypeptide chain. Moreover, a detailed analysis of competent and inhibition-relieved states of VbhA/VbhT(FIC), *Shewanella oneidensis*, and *Neisseria meningitidis* Fic protein structures allowed us to propose a general mechanism for inhibition of FIC domain adenylylation activity, in which the strictly conserved inhibitory glutamate finger located in helix α_{inh} prevents competent ATP binding. Finally, structural homology modeling and biochemical assays indicate that this mechanism is conserved from bacteria to higher eukaryotes.
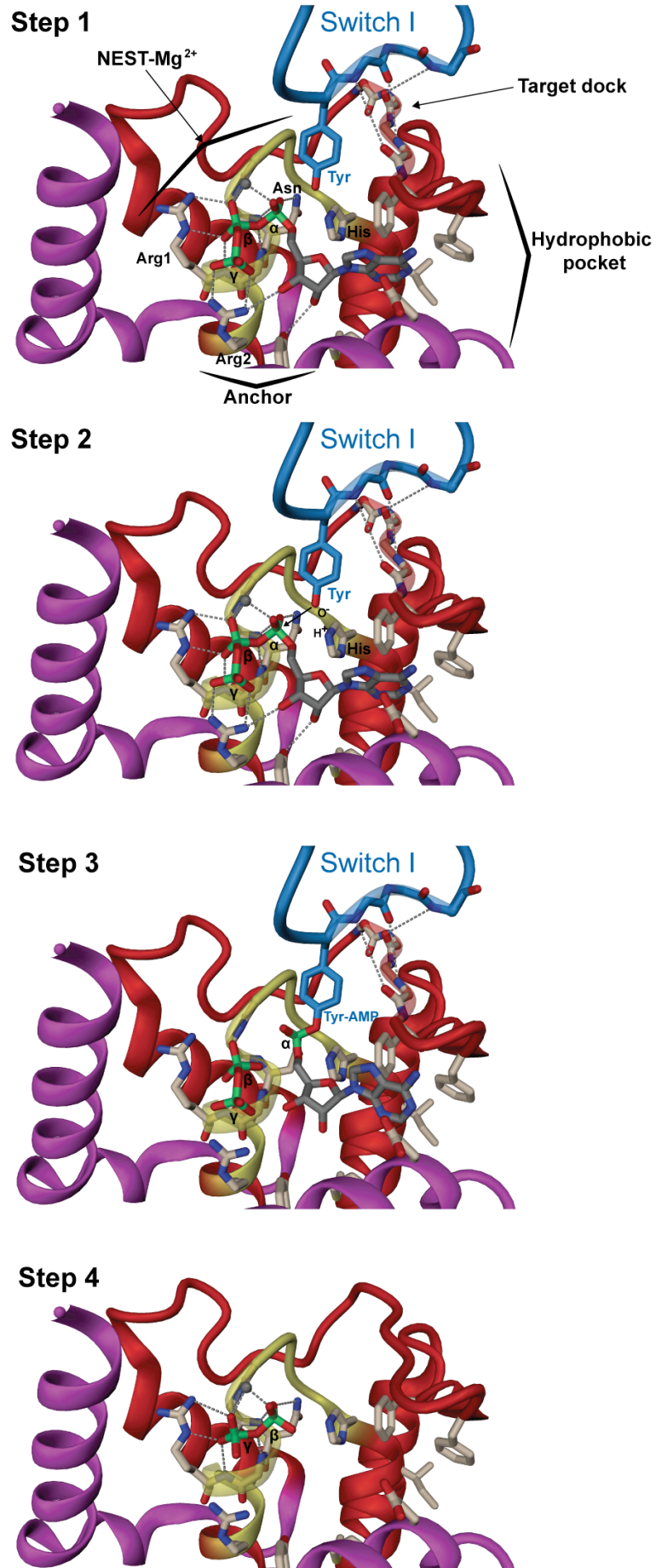
Figure 1: (*legend on next page*)

Figure 1: 3D scenario highlighting the common steps of FIC-catalyzed adenylylation. The magnesium ion is displayed with a grey ball. Important residues and the ATP are shown in full. The structure of NmFic(Δ8) from *Neisseria meningitidis* is used to illustrate the FIC backbone. The target is represented by the switch I region of Cdc42. Its position has been obtained after superposition of the active site loops of NmFic(Δ8) and IbpA-Fic2 from the IbpA-Fic2/Cdc42 complex (PDB code 3N3V). Side-chains of the residues involved in FIC-target β-β interactions are omitted.

Step 1: A ternary complex forms between the Fic protein and its two substrates. The ATP binding site is composed of: 1) hydrophobic pocket enclosing the purine ring, 2) The second arginine (Arg2) of the fic active site motif and a side-chain hydroxyl group coming from helix α6 that serves as a anchor for the nucleotide in the middle of the active site. 3) The compound anion binding nest formed by the conserved GNGR residues from the Fic active site motif together with Arg2 and the magnesium ion that provide extensive hydrogen-bonding network to interact with the triphosphate moiety and orientate the α-phosphate for catalysis. The target dock, through sequence-independent main chain-main chain hydrogen bonds with the target protein, favorably position the side-chain OH group of the target residue (Tyr) with respect to the α-phosphate and the catalytic histidine.

Step 2: The catalytic histidine function as a general base to deprotonate the hydroxyl group of the target residue side-chain and hence activates it for nucleophilic attack on the α-phosphate of the ATP.

Step 3: The AMP moiety is transferred onto the target.

Step 4: The adenylylated target is released with concomitant re-positionning of the pyrophosphate leaving group in the high-affinity binding nest. The pyrophosphate may stay in the Fic active site cavity until a new cycle of adenylylation starts.

136

5 Discussion

5.1   FIC: one fold with several functions

The ubiquitous FIC fold is a rigid and versatile α-helical protein scaffold on which various extensions and insertions give rise to different proteins and provide a broad range of target specificities[89]. Furthermore, proteins with a FIC fold often have additional domains that precede or follow the FIC domain generating additional variability (see Pfam database (PF02661)). The fold also displays a high degree of active site flexibility and plasticity. Plasticity, i.e., amino acid substitutions, is an essential feature to adopt new functions[151]. The mutated residues are usually found at the wall or vicinity of active sites being part of the substrate-binding pocket, and exhibiting high conformational flexibility[151,152].

Active site plasticity is Fic proteins has been nicely exemplified by the work on the effector protein AnkX from *Legionella pneumophila[18]*. The protein harbors an active site (HxFxDANGRxxV) slightly diverging from the consensus HxFx[D/E]GNGRxxR motif that is characteristic of adenylylation-competent Fic enzymes. Indeed, substitutions of the first glycine and second arginine into alanine and valine, respectively, enable AnkX to bind CDP-choline at the active site and catalyse phosphocholination. Apart form these two mutations, proteins bearing AMP or phosphocholine transferase activity share the same active site residues. They most likely utilize a similar anion binding hole made up by residues GNGR and ANGR, respectively, to accomodate the phosphorus groups of their respective ATP and CDP-choline substrates. In both cases, the phosphate moieties are probably further stabilize by a divalent cation coordinated by the common [D/E] amino acid. Remarkably, the two reactions occur via different kinetic mechanisms involving the conserved histidine for catalysis. Whereas a ternary complex forms between the enzyme and both substrates during adenylylation with the histidine serving as a general base, phospho-choline transfer requires a ping-pong mechanism where the histidine is first modified before subsequent transfer on the target protein[87,153]. The fact that proteins with activities of AMP and P-choline transfer are likely to share the main active site features of phosphate-binding and catalytic histidine may suggest that they divergently evolved from a common Fic ancestor with promiscuous functions[154].

Doc and AvrB proteins are structurally related to Fic proteins in the sense that they harbor the same overall α-helical topology and active site location[89] (Chapter 1.2.7). While they have been biologically characterized, the molecular mechanisms underlying their functions remain unknown[95,97]. Even though adenylylation activity has not been disproven,

the structural variations observed in their active sites may imply that they catalyze other reactions or transform other substrates.

The preliminary structural result obtained in Chapter 3.4.3 on the capacity of the VbhT(FIC) protein to recognize the ATP ligand in two distinct orientations is also indicative of the prominent structural flexibility of the substrate-binding site. Flexibility in the vicinity of the active site may allow Fic proteins to accommodate the same substrate in various conformations or probably even different substrates.

This high active site plasticity observed in the several thousand FIC domains suggests that adenylylation and phosphocholination represent only a fraction of the spectrum of reactions they can catalyze. Future studies will considerably expand the functional repertoire of biochemical strategies employed by this ubiquitous protein to interplay with key signaling pathways.

5.2 The Rho GTPase master regulators: preferential targets of Fic-mediated adenylylation

Up to know, all characterized proteins targeted by adenylylation-competent FIC domain-containing effectors appear to belong to the Rho GTPases family. Rho proteins are master regulators of the cytoskeleton and control a remarkable diversity of cellular functions like motility, cell morphology, differentiation and cell division[155-157]. It is therefore not surprising that bacterial Fic-containing effectors, like many virulence factors[16,17,158] directly target this essential protein to establish successful infection.

Rho proteins act as molecular switches that cycle between an 'activated' GTP-bound state allowing binding of downstream effectors and an 'inactivated' GDP-bound state. This cycle is controlled by three types of proteins: the guanine nucleotide exchange factors (GEFs) that catalyze exchange of GDP for GTP to activate the switch[159]; the GTPase activating proteins (GAPs) that stimulate the intrinsic GTPase activity to inactivate the switch[160]; and the guanine nucleotide dissociation inhibitors (GDIs), which mediate the sequestration of Rho proteins from membranes[161]. Fic proteins have been shown to recognize different GTPase states. VopS and IbpA are capable of AMP transfer onto the constitutively active form of the GTPases thereby preventing interaction with downstream partners, but can also chemically

modify the inactive (GDP-bound) state of Rho-GTPases[84]. Rho proteins are activated or inactivated upon binding of GEFs or GAPs, respectively. These two families of Rho GTPase modulators mediate a series of interactions with the switch 1 region of Rho GTPases and in particular with Tyr32/34 and Thr35/37[162,163]. It is therefore conceivable that adenylylation of one of these two residues will modulate GEF or GAP activities towards GTPases. Interestingly, Bep1(FIC) displays high affinity for wild-type Rac1, that is most likely in a GDP-bound form, and only negligible interaction with the constitutively active variant (Chapter 3.4.2). This may indicate a disparate mode of GTPase regulation. Though we cannot exclude that the point mutant that lock GTPases into a constitutively active (Q61L) form, alters the conformation of the GTPase, such that it may not be effectively recognized by Bep1(FIC).

Moreover, as proposed by Mattoo et *al*[84], Fic proteins, by covalent modification with an AMP moiety, may also differently impair the distribution between the membrane and the cytosol of Rho GTPases by perturbing Rho GDIs binding. Such a disruption of complex formation between GDIs and Ras-like GTPases has been already observed in the context of *L. pneumophila* infection. The covalent coercion of Rab GTPase by AnkX and DrrA through phosphocholine and AMP transfer, respectively, lead to GDI displacement from Rab proteins[164].

Thus, besides the role in blocking the binding of downstream effectors, Fic proteins are likely to disrupt the spatio-temporal regulation of Rho GTPases. But this hypothesis needs further investigations.

Interestingly, in contrast to VopS, IbpA, and PfhB2 that are able to adenylylate all Rho family members[84], Bep1 from *Bartonella rochalimae* shows remarkable specificity towards Rac1. Thus, upon translocation into the host cell, Bep1 is likely to hijack specific signal transduction pathways that are under Rac1 control. The specific contribution of Bep1 in *Bartonella* infection and the associated phenotype(s) are currently under investigation in our laboratory. My contribution, as a structural biologist, will be the characterization of the structural features allowing Bep1 to distinguish between this proteins that share high structure and sequence similarities.

The tendency of Fic proteins to target GTPases may reflect a strong selective pressure on the Fic protein to adopt a stable and evolutionarily conserved architecture that specifically recognize the highly abundant GTPase fold. GTPases, also commonly named G proteins,

belong to the P-loop NTPase superfamily that is defined by the so-called P-loop NTPase fold[165]. This fold is characterized by a highly conserved globular shape, made up by at least 5 α/β units, with the strands forming a core parallel sheet, and is the most abundant fold in proteins encoded in the genomes of most cellular life-forms[166]. Moreover, the phosphate-binding loop and the two switch regions that carry the basic functions of nucleotide binding, hydrolysis and effector binding are extremely similar among proteins harboring this typical architecture. It is therefore well imaginable, that all protein targeted by FIC domain display the typical NTPase fold.

5.3   Emergence of new targets of Fic-mediated adenylylation

In Chapter 3.1[140] (Research article I), I could show that the FIC domain of *Bartonella henselae* T4SS effector BepA can adenylylate two eukaryotic targets in-vitro. The apparent size of the proteins detected by in-vitro adenylylation assay clearly excludes them to be small Rho GTPases. It therefore highlights for the first time target proteins distinct from this family and their detection is currently ongoing. BepA has been shown to inhibit apoptosis upon infection of human endothelial cell through elevation of cytosolic cAMP concentration[136]. This phenotype that is triggered by manipulation of host adenylylate cyclases and Gα-subunit of heterotrimeric G-proteins has been exclusively associated with the BID domain of BepA[167]. Even though the FIC domain seems not to be implicated in this process, it is tempting to speculate that it may contribute to it by slight modulation of G-protein activity. Interestingly, the size of the G_s alpha subunit (45 kDa) fits with one of the two bands revealed on the in-vitro adenylylation assay performed with the BepA FIC-OB domain (Chapter 3.1 (Research article I))[140].

The linker region between the FIC and the BID domain adopts an OB-fold like topology that was originally named for its oligonucleotide/oligosaccharide binding properties. This fold is quite frequent in DNA- and RNA-binding proteins but has also been observed at protein-protein interfaces and in some cases can even bear a catalytic function[168,169]. The role of the OB domain in *Bartonella* effector proteins is still unknown. A putative function in target recognition analogous to the arm domain of VopS and IbpA might be considered[88]. However, the fact that OB-fold appears to be also present on the relaxases of *B. grahamii* and *B.*

schoenbuchensis, from which the BID translocation signal originates, and that the OB-fold linker region of Bep1 *rochalimae* does not contribute to Rac1 binding (Chapter 3.4.2), suggest that it might rather function independently or in conjunction with the BID domain.

In Chapter 3.2[148] (Research article II), we performed *in vitro* adenylylation assays with the FIC domain-containing protein VbhT from *B. schoenbuchensis*. Incubation of the protein with *E. coli* lysate, revealed a band with an apparent size of 80 kDa that is likely to represent a bacterial target for the Fic protein. By a combination of pull-down and mass spectrometry experiments, Alexander Harms could demonstrate that this band corresponds to the subunit beta of the DNA gyrase (Unpublished results). Strikingly, the β-subunit of DNA gyrase harbors a P-loop NTPase domain supporting the hypothesis that Fic proteins are pre-adapted to recognize this typical fold[170]. DNA gyrase is an enzyme that catalyzes the introduction of negative supercoils in DNA, an activity that is essential in the process of cell replication[171]. Whereas the catalytic machinery is situated in the subunit alpha, the subunit beta bears an ATPase domain that is essential for gyrase function. Interestingly, blockade of the gyrase ATPase activity leads to cell filamentation in *E. coli*[172], supporting that the filamentous growth phenotype observed for VbhT might result from specific adenylylation of DNA gyrase[148]. Mass spectrometry has traced Y109 to be the modified residue (Alexander Harms, Biozentrum Basel, unpublished result). This amino acid is highly conserved among bacterial gyrases and is located within the ATP-binding pocket of the ATPase domain.

From a biological point of view the VbhT toxic activity probably plays a role during the conjugation process in Bartonella. Indeed, the VbhA-T operon is tightly associated with the plasmid-encoded Vbh T4SS that has been shown to be a functional conjugation machinery mediating plasmid transfer between different *Bartonella* species[139]. Moreover, adjacent to the FIC domain, VbhT harbors a functional BID translocation signal suggesting that it gets co-translocated with the plasmid DNA during conjugation[139]. Thus, VbhT may toxify the recipient cell during conjugation by adenylylating the gyrase and blocking the replication process leading to cell filamentation and growth inhibition. Once the conjugation process is completed, the antitoxin encoded on the plasmid will be expressed, detoxifying the recipient cell. In case of abortive conjugation, VbhT activity would not be counteracted resulting eventually to cell death. Furthermore a blockade of replication might increase the amount of freely available bacterial host DNA replication machinery and allow rapid replication of the plasmid. The ability of the system to kill or inhibit the growth of recipient cell that does not receive a copy of plasmid through gyrase inhibition displays analogy with some post-

segregational killing (PSK) systems that ensure stable inheritance of plasmids in a population[173,174].

Intriguingly, the Fic protein from *N. meningitidis* (NmFic) was also found to adenylylate DNA gyrase *in vitro*, further indicating that Fic proteins are conservative in their target specificity (Alexander Harms, Biozentrum Basel, unpublished). However, in vitro adenylylation assay revealed additional modified proteins in *E. coli*[148] (Chapter 3.2 (Research article II)). These proteins have not been yet identified and it is unclear whether they are specific or non-specific targets. But the modification of multiple proteins that may have counteracting activities could explain the less severe toxic effect observed in *E. coli*.

Finally, recent investigations performed by Alexander Harms (Biozentrum, Basel) on the Bep2 effector from *B. rochalimae* also revealed adenylylation of an unknown eukaryotic target protein with an apparent size of 50 kDa[147]. Studies aiming to identify the target and the corresponding phenotype are currently underway in our laboratory.

A vast majority of Fic proteins bear adenylylation function to regulate multiple cellular functions. However, the activity is hindered in a large portion of Fic proteins by the general inhibitory mechanism described in Chapter 3.2 and 3.3[148,149] (Research articles II and III). This regulation may largely explain the current lack of knowledge about the various biological functions mediated by adenylylation-competent Fic proteins. Knowing the easy genetic trick to relieve the inhibition opens new avenues for exploring the diverse biological roles of Fic-mediated adenylylation in all kingdoms of life.

5.4 Putative roles of Fic-antitoxins associated with host-interacting effectors.

Comparative sequence analysis of the Vbh T4SS of *B. schoenbuchensis* with the VirB loci of *Bartonella* lineages 3 and 4 showed that the small open reading frame encoded directly upstream of *vbht* and coding for the antitoxin VbhA in the Vbh locus is also present in front of fic genes in the *virB* loci. The exact biological function of these Fic-antitoxins that are genetically associated with eukaryotic host-interacting bep effectors remains elusive.

In Chapter 3.4.2, I corroborated previous finding made by Alexander Harms[147] and demonstrated that the *B. rochalimae* BiaA protein, encoded upstream of Bep1, can inhibit Bep1-mediated adenylylation of the eukaryotic Rac1 GTPase. Remarkably, comparison with the structure of the VbhA/VbhT(FIC) complex from *B.schoenbuchensis* shows high homology and probably an identical inhibitory mechanism via Fic active site obstruction[148]. The first obvious hypothesis that emerges from this functional and structural data is that BiaA tightly associates with the FIC domain and gets translocated into the eukaryotic host where it would prevent adenylylation of Rac1. Nevertheless, it remains to be elucidated whether Fic-antitoxin are injected into the host cell during *Bartonella* infection. Preliminary experiment done by Philipp Engel with a *B. henselae* mutant devoid of the antitoxin did not affect invasome formation during *in vitro* infection.

Furthermore, it is assumed that proteins need to unfold or partially unfold in order to be translocated through the T4S apparatus[108]. Protein unfolding would lead to the release of the antitoxin from the effector prior secretion, indicative of a role within the bacteria. Thus, the antitoxin may still have an ancestral function of FIC inhibition in the bacteria. However, no toxic effect of any *Bartonella* Fic effector on bacterial cells could be detected so far.

More cryptic functions might be also envisaged, but would require further experimental support. The small ORF may act as a chaperone for the FIC domain-containing effector proteins, preventing their aggregation or targeting them to the secretion machinery[175]. Interestingly, Fic-antitoxins are characterized by a small size, acidic pI, and amphipathic helices that are physical properties characteristics of chaperones[175].

Also, the functional analogy with classical toxin-antitoxin systems may indicate a function in transcriptional regulation of *fic* operons.

Finally, adenylylation involves consumption of ATP that is a potential danger to the bacterial cell. Even in the absence of target, ATP hydrolysis may happen in the course of auto-adenylylation. Thus, antitoxin might control Fic activity to prevent futile ATP hydrolysis until translocation into host cell.

All these various aspects are currently being investigated in our laboratory.

6   References

References

1   Black, D. L. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**, 291-336, (2003).
2   Ayoubi, T. A. & Van De Ven, W. J. Regulation of gene expression by alternative promoters. *Faseb J* **10**, 453-460, (1996).
3   Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666-672, (2010).
4   Geiss-Friedlander, R. & Melchior, F. Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol* **8**, 947-956, (2007).
5   Aicart-Ramos, C., Valero, R. A. & Rodriguez-Crespo, I. Protein palmitoylation and subcellular trafficking. *Biochim Biophys Acta* **1808**, 2981-2994, (2011).
6   de Launoit, Y., Degerny, C., Wohlkonig, A. & Baert, J. L. [Multiple post-translational modifications implied in the gene expression regulation: example of the sumoylation of Ets factors]. *Bull Mem Acad R Med Belg* **162**, 299-305; discussion 306, (2007).
7   Ribet, D. & Cossart, P. Post-translational modifications in host cells during bacterial infection. *FEBS Lett* **584**, 2748-2758, (2010).
8   Wiley, D. J., Nordfeldth, R., Rosenzweig, J., DaFonseca, C. J., Gustin, R., Wolf-Watz, H. & Schesser, K. The Ser/Thr kinase activity of the Yersinia protein kinase A (YpkA) is necessary for full virulence in the mouse, mollifying phagocytes, and disrupting the eukaryotic cytoskeleton. *Microb Pathog* **40**, 234-243, (2006).
9   Bliska, J. B., Guan, K. L., Dixon, J. E. & Falkow, S. Tyrosine phosphate hydrolysis of host proteins by an essential Yersinia virulence determinant. *Proc Natl Acad Sci U S A* **88**, 1187-1191, (1991).
10   Brennan, D. F. & Barford, D. Eliminylation: a post-translational modification catalyzed by phosphothreonine lyases. *Trends Biochem Sci* **34**, 108-114, (2009).
11   Trosky, J. E., Li, Y., Mukherjee, S., Keitany, G., Ball, H. & Orth, K. VopA inhibits ATP binding by acetylating the catalytic loop of MAPK kinases. *J Biol Chem* **282**, 34299-34305, (2007).
12   Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M. & Aktories, K. Glucosylation of Rho proteins by Clostridium difficile toxin B. *Nature* **375**, 500-503, (1995).
13   Just, I., Wilm, M., Selzer, J., Rex, G., von Eichel-Streiber, C., Mann, M. & Aktories, K. The enterotoxin from Clostridium difficile (ToxA) monoglucosylates the Rho proteins. *J Biol Chem* **270**, 13932-13936, (1995).
14   Richard, J. F., Petit, L., Gibert, M., Marvaud, J. C., Bouchaud, C. & Popoff, M. R. Bacterial toxins modifying the actin cytoskeleton. *Int Microbiol* **2**, 185-194, (1999).
15   Rosebrock, T. R., Zeng, L., Brady, J. J., Abramovitch, R. B., Xiao, F. & Martin, G. B. A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature* **448**, 370-374, (2007).
16   Masuda, M., Betancourt, L., Matsuzawa, T., Kashimoto, T., Takao, T., Shimonishi, Y. & Horiguchi, Y. Activation of rho through a cross-link with polyamines catalyzed by Bordetella dermonecrotizing toxin. *Embo J* **19**, 521-530, (2000).
17   Shao, F., Merritt, P. M., Bao, Z., Innes, R. W. & Dixon, J. E. A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. *Cell* **109**, 575-588, (2002).
18   Mukherjee, S., Liu, X., Arasaki, K., McDonough, J., Galan, J. E. & Roy, C. R. Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477**, 103-106, (2011).
19   Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., Zekarias, B., Lazar, C. & Dixon, J. E. The fic domain: regulation of cell signaling by adenylylation. *Mol Cell* **34**, 93-103, (2009).

146

20   Yarbrough, M. L., Li, Y., Kinch, L. N., Grishin, N. V., Ball, H. L. & Orth, K. AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science* **323**, 269-272, (2009).

21   Itzen, A., Blankenfeldt, W. & Goody, R. S. Adenylylation: renaissance of a forgotten post-translational modification. *Trends Biochem Sci* **36**, 221-228, (2011).

22   Anderson, W. B., Hennig, S. B., Ginsburg, A. & Stadtman, E. R. Association of ATP: glutamine synthetase adenylyltransferase activity with the P1 component of the glutamine synthetase deadenylylation system. *Proc Natl Acad Sci U S A* **67**, 1417-1424, (1970).

23   Schmelz, S. & Naismith, J. H. Adenylate-forming enzymes. *Curr Opin Struct Biol* **19**, 666-671, (2009).

24   Pesole, G., Bozzetti, M. P., Lanave, C., Preparata, G. & Saccone, C. Glutamine synthetase gene evolution: a good molecular clock. *Proc Natl Acad Sci U S A* **88**, 522-526, (1991).

25   Son, H. S. & Rhee, S. G. Cascade control of Escherichia coli glutamine synthetase. Purification and properties of PII protein and nucleotide sequence of its structural gene. *J Biol Chem* **262**, 8690-8695, (1987).

26   Jaggi, R., Ybarlucea, W., Cheah, E., Carr, P. D., Edwards, K. J., Ollis, D. L. & Vasudevan, S. G. The role of the T-loop of the signal transducing protein PII from Escherichia coli. *FEBS Lett* **391**, 223-228, (1996).

27   Jaggi, R., van Heeswijk, W. C., Westerhoff, H. V., Ollis, D. L. & Vasudevan, S. G. The two opposing activities of adenylyl transferase reside in distinct homologous domains, with intramolecular signal transduction. *Embo J* **16**, 5562-5571, (1997).

28   Pedersen, L. C., Benning, M. M. & Holden, H. M. Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry* **34**, 13305-13311, (1995).

29   Morar, M., Bhullar, K., Hughes, D. W., Junop, M. & Wright, G. D. Structure and mechanism of the lincosamide antibiotic adenylyltransferase LinB. *Structure* **17**, 1649-1659, (2009).

30   Azucena, E. & Mobashery, S. Aminoglycoside-modifying enzymes: mechanisms of catalytic processes and inhibition. *Drug Resist Updat* **4**, 106-117, (2001).

31   Delarue, M. Aminoacyl-tRNA synthetases. *Curr Opin Struct Biol* **5**, 48-55, (1995).

32   Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**, 159-180, (2006).

33   Maupin-Furlow, J. Proteasomes and protein conjugation across domains of life. *Nat Rev Microbiol* **10**, 100-111, (2012).

34   Haglund, K. & Dikic, I. The role of ubiquitylation in receptor endocytosis and endosomal sorting. *J Cell Sci* **125**, 265-275, (2012).

35   Haglund, K. & Dikic, I. Ubiquitylation and cell signaling. *Embo J* **24**, 3353-3359, (2005).

36   Pickart, C. M. & Fushman, D. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* **8**, 610-616, (2004).

37   Schulman, B. A. & Harper, J. W. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**, 319-331, (2009).

38   Soucy, T. A., Dick, L. R., Smith, P. G., Milhollen, M. A. & Brownell, J. E. The NEDD8 Conjugation Pathway and Its Relevance in Cancer Biology and Therapy. *Genes Cancer* **1**, 708-716, (2010).

39   Hay, R. T. SUMO: a history of modification. *Mol Cell* **18**, 1-12, (2005).

40   Preiss, J. & Romeo, T. Molecular biology and regulatory aspects of glycogen biosynthesis in bacteria. *Prog Nucleic Acid Res Mol Biol* **47**, 299-329, (1994).

41   Jin, X., Ballicora, M. A., Preiss, J. & Geiger, J. H. Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *Embo J* **24**, 694-704, (2005).

42   Chang, H. Y., Peng, H. L., Chao, Y. C. & Duggleby, R. G. The importance of conserved residues in human liver UDPglucose pyrophosphorylase. *Eur J Biochem* **236**, 723-728, (1996).

43   Sauve, A. A. NAD+ and vitamin B3: from metabolism to therapies. *J Pharmacol Exp Ther* **324**, 883-893, (2008).

44   Magni, G., Amici, A., Emanuelli, M., Raffaelli, N. & Ruggieri, S. Enzymology of NAD+ synthesis. *Adv Enzymol Relat Areas Mol Biol* **73**, 135-182, xi, (1999).

45   Magni, G., Amici, A., Emanuelli, M., Orsomando, G., Raffaelli, N. & Ruggieri, S. Structure and function of nicotinamide mononucleotide adenylyltransferase. *Curr Med Chem* **11**, 873-885, (2004).

46   Massey, V. The chemical and biological versatility of riboflavin. *Biochem Soc Trans* **28**, 283-296, (2000).

47   Manstein, D. J. & Pai, E. F. Purification and characterization of FAD synthetase from Brevibacterium ammoniagenes. *J Biol Chem* **261**, 16169-16173, (1986).

48   Izard, T. A novel adenylate binding site confers phosphopantetheine adenylyltransferase interactions with coenzyme A. *J Bacteriol* **185**, 4074-4080, (2003).

49   Wang, W., Kim, R., Yokota, H. & Kim, S. H. Crystal structure of flavin binding to FAD synthetase of Thermotoga maritima. *Proteins* **58**, 246-248, (2005).

50   Saridakis, V., Christendat, D., Kimber, M. S., Dharamsi, A., Edwards, A. M. & Pai, E. F. Insights into ligand binding and catalysis of a central step in NAD+ synthesis: structures of Methanobacterium thermoautotrophicum NMN adenylyltransferase complexes. *J Biol Chem* **276**, 7225-7232, (2001).

51   Izard, T. The crystal structures of phosphopantetheine adenylyltransferase with bound substrates reveal the enzyme's catalytic mechanism. *J Mol Biol* **315**, 487-495, (2002).

52   Perona, J. J., Rould, M. A. & Steitz, T. A. Structural basis for transfer RNA aminoacylation by Escherichia coli glutaminyl-tRNA synthetase. *Biochemistry* **32**, 8758-8771, (1993).

53   Sekine, S., Nureki, O., Dubois, D. Y., Bernier, S., Chenevert, R., Lapointe, J., Vassylyev, D. G. & Yokoyama, S. ATP binding by glutamyl-tRNA synthetase is switched to the productive mode by tRNA binding. *Embo J* **22**, 676-688, (2003).

54   Mendel, R. R. Molybdenum: biological activity and metabolism. *Dalton Trans* 3404-3409, (2005).

55   Leimkuhler, S., Wuebbens, M. M. & Rajagopalan, K. V. Characterization of Escherichia coli MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *J Biol Chem* **276**, 34695-34701, (2001).

56   Lake, M. W., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. *Nature* **414**, 325-329, (2001).

57   Rudolph, M. J., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. Crystal structure of molybdopterin synthase and its evolutionary relationship to ubiquitin activation. *Nat Struct Biol* **8**, 42-46, (2001).

58   Negishi, M., Pedersen, L. G., Petrotchenko, E., Shevtsov, S., Gorokhov, A., Kakuta, Y. & Pedersen, L. C. Structure and function of sulfotransferases. *Arch Biochem Biophys* **390**, 149-157, (2001).

59   Harjes, S., Bayer, P. & Scheidig, A. J. The crystal structure of human PAPS synthetase 1 reveals asymmetry in substrate binding. *J Mol Biol* **347**, 623-635, (2005).

60   Horwitz, M. A. The Legionnaires' disease bacterium (Legionella pneumophila) inhibits phagosome-lysosome fusion in human monocytes. *J Exp Med* **158**, 2108-2126, (1983).

61   Ge, J. & Shao, F. Manipulation of host vesicular trafficking and innate immune defence by Legionella Dot/Icm effectors. *Cell Microbiol* **13**, 1870-1880, (2011).

62   Kagan, J. C., Stein, M. P., Pypaert, M. & Roy, C. R. Legionella subvert the functions of Rab1 and Sec22b to create a replicative organelle. *J Exp Med* **199**, 1201-1211, (2004).

63   Muller, M. P., Peters, H., Blumer, J., Blankenfeldt, W., Goody, R. S. & Itzen, A. The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b. *Science* **329**, 946-949, (2010).

64   Tan, Y. & Luo, Z. Q. Legionella pneumophila SidD is a deAMPylase that modifies Rab1. *Nature* **475**, 506-509, (2011).

65   Neunuebel, M. R., Chen, Y., Gaspar, A. H., Backlund, P. S., Jr., Yergey, A. & Machner, M. P. De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila. *Science* **333**, 453-456, (2011).

66   Goody, R. S., Muller, M. P., Schoebel, S., Oesterlin, L. K., Blumer, J., Peters, H., Blankenfeldt, W. & Itzen, A. The versatile Legionella effector protein DrrA. *Commun Integr Biol* **4**, 72-74, (2011).

67   Kawamukai, M., Matsuda, H., Fujii, W., Nishida, T., Izumoto, Y., Himeno, M., Utsumi, R. & Komano, T. Cloning of the fic-1 gene involved in cell filamentation induced by cyclic AMP and construction of a delta fic Escherichia coli strain. *J Bacteriol* **170**, 3864-3869, (1988).

68   Kawamukai, M., Matsuda, H., Fujii, W., Utsumi, R. & Komano, T. Nucleotide sequences of fic and fic-1 genes involved in cell filamentation induced by cyclic AMP in Escherichia coli. *J Bacteriol* **171**, 4525-4529, (1989).

69   Utsumi, R., Noda, M., Kawamukai, M. & Komano, T. Control mechanism of the Escherichia coli K-12 cell cycle is triggered by the cyclic AMP-cyclic AMP receptor protein complex. *J Bacteriol* **171**, 2909-2912, (1989).

70   Komano, T., Utsumi, R. & Kawamukai, M. Functional analysis of the fic gene involved in regulation of cell division. *Res Microbiol* **142**, 269-277, (1991).

71   Utsumi, R., Kawamukai, M., Obata, K., Morita, J., Himeno, M. & Komano, T. Identification of a membrane protein induced concurrently with cell filamentation by cyclic AMP in an Escherichia coli K-12 fic mutant. *J Bacteriol* **155**, 398-401, (1983).

72   Utsumi, R., Tanabe, H., Nakamoto, Y., Kawamukai, M., Sakai, H., Himeno, M., Komano, T. & Hirota, Y. Inhibitory effect of adenosine 3',5'-phosphate on cell division of Escherichia coli K-12 mutant derivatives. *J Bacteriol* **147**, 1105-1109, (1981).

73   Utsumi, R., Nakamoto, Y., Kawamukai, M., Himeno, M. & Komano, T. Involvement of cyclic AMP and its receptor protein in filamentation of an Escherichia coli fic mutant. *J Bacteriol* **151**, 807-812, (1982).

74   Daniels, N. A., MacKinnon, L., Bishop, R., Altekruse, S., Ray, B., Hammond, R. M., Thompson, S., Wilson, S., Bean, N. H., Griffin, P. M. & Slutsker, L. Vibrio parahaemolyticus infections in the United States, 1973-1998. *J Infect Dis* **181**, 1661-1666, (2000).

75   Morris, J. G., Jr. & Black, R. E. Cholera and other vibrioses in the United States. *N Engl J Med* **312**, 343-350, (1985).

76   Park, K. S., Ono, T., Rokuda, M., Jang, M. H., Okada, K., Iida, T. & Honda, T. Functional characterization of two type III secretion systems of Vibrio parahaemolyticus. *Infect Immun* **72**, 6659-6665, (2004).

77   Burdette, D. L., Yarbrough, M. L., Orvedahl, A., Gilpin, C. J. & Orth, K. Vibrio parahaemolyticus orchestrates a multifaceted host cell infection by induction of autophagy, cell rounding, and then cell lysis. *Proc Natl Acad Sci U S A* **105**, 12497-12502, (2008).

78   Casselli, T., Lynch, T., Southward, C. M., Jones, B. W. & DeVinney, R. Vibrio parahaemolyticus inhibition of Rho family GTPase activation requires a functional chromosome I type III secretion system. *Infect Immun* **76**, 2202-2211, (2008).

79   Corbeil, L. B. Histophilus somni host-parasite relationships. *Anim Health Res Rev* **8**, 151-160, (2007).

80   Harris, F. W. & Janzen, E. D. The Haemophilus somnus disease complex (Hemophilosis): A review. *Can Vet J* **30**, 816-822, (1989).

81   O'Toole, D., Allen, T., Hunter, R. & Corbeil, L. B. Diagnostic exercise: Myocarditis due to Histophilus somni in feedlot and backgrounded cattle. *Vet Pathol* **46**, 1015-1017, (2009).

82   Widders, P. R., Paisley, L. G., Gogolewski, R. P., Evermann, J. F., Smith, J. W. & Corbeil, L. B. Experimental abortion and the systemic immune response to "Haemophilus somnus" in cattle. *Infect Immun* **54**, 555-560, (1986).

83   Zekarias, B., Mattoo, S., Worby, C., Lehmann, J., Rosenbusch, R. F. & Corbeil, L. B. Histophilus somni IbpA DR2/Fic in virulence and immunoprotection at the natural host alveolar epithelial barrier. *Infect Immun* **78**, 1850-1858, (2010).

84   Mattoo, S., Durrant, E., Chen, M. J., Xiao, J., Lazar, C. S., Manning, G., Dixon, J. E. & Worby, C. A. 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).

85   Faber, P. W., Barnes, G. T., Srinidhi, J., Chen, J., Gusella, J. F. & MacDonald, M. E. Huntingtin interacts with a family of WW domain proteins. *Hum Mol Genet* **7**, 1463-1474, (1998).

86   Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L., Eddy, S. R., Bateman, A. & Finn, R. D. The Pfam protein families database. *Nucleic Acids Res* **40**, D290-301, (2012).

87   Luong, P., Kinch, L. N., Brautigam, C. A., Grishin, N. V., Tomchick, D. R. & Orth, K. Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J Biol Chem* **285**, 20155-20163, (2010).

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

89   Kinch, L. N., Yarbrough, M. L., Orth, K. & Grishin, N. V. Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PLoS One* **4**, e5818, (2009).

90   Das, D., Krishna, S. S., McMullan, D., Miller, M. D., Xu, Q., Abdubek, P., Acosta, C., Astakhova, T., Axelrod, H. L., Burra, P., Carlton, D., Chiu, H. J., Clayton, T., Deller, M. C., Duan, L., Elias, Y., Elsliger, M. A., Ernst, D., Feuerhelm, J., Grzechnik, A., Grzechnik, S. K., Hale, J., Han, G. W., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Kozbial, P., Kumar, A., Marciano, D., Morse, A. T., Murphy, K. D., Nigoghossian, E., Okach, L., Oommachen, S., Paulsen, J., Reyes, R., Rife, C. L., Sefcovic, N., Tien, H., Trame, C. B., Trout, C. V., van den Bedem, H., Weekes, D., White, A., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A. & Wilson, I. A. Crystal structure of the Fic (Filamentation induced by cAMP) family protein SO4266 (gi|24375750) from Shewanella oneidensis MR-1 at 1.6 A resolution. *Proteins* **75**, 264-271, (2009).

91   Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V. & Notredame, C. Expresso: automatic incorporation of structural information in

multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res* **34**, W604-608, (2006).

92   Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-308, (1999).

93   Henikoff, J. G. & Henikoff, S. Blocks database and its applications. *Methods Enzymol* **266**, 88-105, (1996).

94   Lehnherr, H., Maguin, E., Jafri, S. & Yarmolinsky, M. B. Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *J Mol Biol* **233**, 414-428, (1993).

95   Liu, M., Zhang, Y., Inouye, M. & Woychik, N. A. Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. *Proc Natl Acad Sci U S A* **105**, 5885-5890, (2008).

96   Nimchuk, Z., Marois, E., Kjemtrup, S., Leister, R. T., Katagiri, F. & Dangl, J. L. Eukaryotic fatty acylation drives plasma membrane targeting and enhances function of several type III effector proteins from Pseudomonas syringae. *Cell* **101**, 353-363, (2000).

97   Mackey, D., Holt, B. F., 3rd, Wiig, A. & Dangl, J. L. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743-754, (2002).

98   Desveaux, D., Singer, A. U., Wu, A. J., McNulty, B. C., Musselwhite, L., Nimchuk, Z., Sondek, J. & Dangl, J. L. Type III effector activation via nucleotide binding, phosphorylation, and host target interaction. *PLoS Pathog* **3**, e48, (2007).

99   Chung, E. H., da Cunha, L., Wu, A. J., Gao, Z., Cherkis, K., Afzal, A. J., Mackey, D. & Dangl, J. L. Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. *Cell Host Microbe* **9**, 125-136, (2011).

100   Liu, J., Elmore, J. M., Lin, Z. J. & Coaker, G. A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, leading to the activation of a plant innate immune receptor. *Cell Host Microbe* **9**, 137-146, (2011).

101   Lee, C. C., Wood, M. D., Ng, K., Andersen, C. B., Liu, Y., Luginbuhl, P., Spraggon, G. & Katagiri, F. Crystal structure of the type III effector AvrB from Pseudomonas syringae. *Structure* **12**, 487-494, (2004).

102   Dehio, C. Bartonella-host-cell interactions and vascular tumour formation. *Nat Rev Microbiol* **3**, 621-631, (2005).

103   Vayssier-Taussat, M., Le Rhun, D., Bonnet, S. & Cotte, V. Insights in Bartonella host specificity. *Ann N Y Acad Sci* **1166**, 127-132, (2009).

104   Saenz, H. L., Engel, P., Stoeckli, M. C., Lanz, C., Raddatz, G., Vayssier-Taussat, M., Birtles, R., Schuster, S. C. & Dehio, C. Genomic analysis of Bartonella identifies type IV secretion systems as host adaptability factors. *Nat Genet* **39**, 1469-1476, (2007).

105   Engel, P., Salzburger, W., Liesch, M., Chang, C. C., Maruyama, S., Lanz, C., Calteau, A., Lajus, A., Medigue, C., Schuster, S. C. & Dehio, C. Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen Bartonella. *PLoS Genet* **7**, e1001296, (2011).

106   Schulein, R. & Dehio, C. The VirB/VirD4 type IV secretion system of Bartonella is essential for establishing intraerythrocytic infection. *Mol Microbiol* **46**, 1053-1067, (2002).

107   Harms, A. & Dehio, C. Intruders below the radar: molecular pathogenesis of Bartonella spp. *Clin Microbiol Rev* **25**, 42-78, (2012).

108   Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S. & Cascales, E. Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* **59**, 451-485, (2005).

109 Hamilton, H. L., Dominguez, N. M., Schwartz, K. J., Hackett, K. T. & Dillard, J. P. Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* **55**, 1704-1721, (2005).

110 Juhas, M., Crook, D. W. & Hood, D. W. Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol* **10**, 2377-2386, (2008).

111 Schmid, M. C., Schulein, R., Dehio, M., Denecker, G., Carena, I. & Dehio, C. The VirB type IV secretion system of Bartonella henselae mediates invasion, proinflammatory activation and antiapoptotic protection of endothelial cells. *Mol Microbiol* **52**, 81-92, (2004).

112 Schulein, R., Guye, P., Rhomberg, T. A., Schmid, M. C., Schroder, G., Vergunst, A. C., Carena, I. & Dehio, C. 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).

113 Segal, G., Feldman, M. & Zusman, T. The Icm/Dot type-IV secretion systems of Legionella pneumophila and Coxiella burnetii. *FEMS Microbiol Rev* **29**, 65-81, (2005).

114 Backert, S. & Meyer, T. F. Type IV secretion systems and their effectors in bacterial pathogenesis. *Curr Opin Microbiol* **9**, 207-217, (2006).

115 Fronzes, R., Schafer, E., Wang, L., Saibil, H. R., Orlova, E. V. & Waksman, G. Structure of a type IV secretion system core complex. *Science* **323**, 266-268, (2009).

116 Chandran, V., Fronzes, R., Duquerroy, S., Cronin, N., Navaza, J. & Waksman, G. Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**, 1011-1015, (2009).

117 Jakubowski, S. J., Kerr, J. E., Garza, I., Krishnamoorthy, V., Bayliss, R., Waksman, G. & Christie, P. J. Agrobacterium VirB10 domain requirements for type IV secretion and T pilus biogenesis. *Mol Microbiol* **71**, 779-794, (2009).

118 Cascales, E. & Christie, P. J. Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science* **304**, 1170-1173, (2004).

119 Krause, S., Pansegrau, W., Lurz, R., de la Cruz, F. & Lanka, E. Enzymology of type IV macromolecule secretion systems: the conjugative transfer regions of plasmids RP4 and R388 and the cag pathogenicity island of Helicobacter pylori encode structurally and functionally related nucleoside triphosphate hydrolases. *J Bacteriol* **182**, 2761-2770, (2000).

120 Rivas, S., Bolland, S., Cabezon, E., Goni, F. M. & de la Cruz, F. TrwD, a protein encoded by the IncW plasmid R388, displays an ATP hydrolase activity essential for bacterial conjugation. *J Biol Chem* **272**, 25583-25590, (1997).

121 Durand, E., Oomen, C. & Waksman, G. Biochemical dissection of the ATPase TraB, the VirB4 homologue of the Escherichia coli pKM101 conjugation machinery. *J Bacteriol* **192**, 2315-2323, (2010).

122 Yeo, H. J., Yuan, Q., Beck, M. R., Baron, C. & Waksman, G. Structural and functional characterization of the VirB5 protein from the type IV secretion system encoded by the conjugative plasmid pKM101. *Proc Natl Acad Sci U S A* **100**, 15947-15952, (2003).

123 Backert, S., Fronzes, R. & Waksman, G. VirB2 and VirB5 proteins: specialized adhesins in bacterial type-IV secretion systems? *Trends Microbiol* **16**, 409-413, (2008).

124 Zupan, J., Hackworth, C. A., Aguilar, J., Ward, D. & Zambryski, P. VirB1* promotes T-pilus formation in the vir-Type IV secretion system of Agrobacterium tumefaciens. *J Bacteriol* **189**, 6551-6563, (2007).

125 Mossey, P., Hudacek, A. & Das, A. Agrobacterium tumefaciens type IV secretion protein VirB3 is an inner membrane protein and requires VirB4, VirB7, and VirB8 for stabilization. *J Bacteriol* **192**, 2830-2838, (2010).

126 Waksman, G. & Fronzes, R. Molecular architecture of bacterial type IV secretion systems. *Trends Biochem Sci* **35**, 691-698, (2010).

127 Dehio, C. Infection-associated type IV secretion systems of Bartonella and their diverse roles in host cell interaction. *Cell Microbiol* **10**, 1591-1598, (2008).

128 Karem, K. L. Immune aspects of Bartonella. *Crit Rev Microbiol* **26**, 133-145, (2000).

129 Dehio, C. Molecular and cellular basis of bartonella pathogenesis. *Annu Rev Microbiol* **58**, 365-390, (2004).

130 Mosepele, M., Mazo, D. & Cohn, J. Bartonella infection in immunocompromised hosts: immunology of vascular infection and vasoproliferation. *Clin Dev Immunol* **2012**, 612809, (2012).

131 Dehio, C. Bartonella interactions with endothelial cells and erythrocytes. *Trends Microbiol* **9**, 279-285, (2001).

132 Dehio, C. Recent progress in understanding Bartonella-induced vascular proliferation. *Curr Opin Microbiol* **6**, 61-65, (2003).

133 Rhomberg, T. A., Truttmann, M. C., Guye, P., Ellner, Y. & Dehio, C. A translocated protein of Bartonella henselae interferes with endocytic uptake of individual bacteria and triggers uptake of large bacterial aggregates via the invasome. *Cell Microbiol* **11**, 927-945, (2009).

134 Truttmann, M. C., Guye, P. & Dehio, C. BID-F1 and BID-F2 domains of Bartonella henselae effector protein BepF trigger together with BepC the formation of invasome structures. *PLoS One* **6**, e25106, (2011).

135 Truttmann, M. C., Rhomberg, T. A. & Dehio, C. Combined action of the type IV secretion effector proteins BepC and BepF promotes invasome formation of Bartonella henselae on endothelial and epithelial cells. *Cell Microbiol* **13**, 284-299, (2011).

136 Schmid, M. C., Scheidegger, F., Dehio, M., Balmelle-Devaux, N., Schulein, R., Guye, P., Chennakesava, C. S., Biedermann, B. & Dehio, C. A translocated bacterial protein protects vascular endothelial cells from apoptosis. *PLoS Pathog* **2**, e115, (2006).

137 Scheidegger, F., Ellner, Y., Guye, P., Rhomberg, T. A., Weber, H., Augustin, H. G. & Dehio, C. Distinct activities of Bartonella henselae type IV secretion effector proteins modulate capillary-like sprout formation. *Cell Microbiol* **11**, 1088-1101, (2009).

138 Selbach, M., Paul, F. E., Brandt, S., Guye, P., Daumke, O., Backert, S., Dehio, C. & Mann, M. Host cell interactome of tyrosine-phosphorylated bacterial proteins. *Cell Host Microbe* **5**, 397-403, (2009).

139 Engel, P. A Genomic Approach to the Role of Type IV Secretion Systems in Bartonella Host Adaptation. *PhD Thesis* 2009).

140 Palanivelu, D. V., Goepfert, A., Meury, M., Guye, P., Dehio, C. & Schirmer, T. Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. *Protein Sci* **20**, 492-499, (2011).

141 Leslie, A. G. The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**, 48-57, (2006).

142 Evans, P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82, (2006).

143 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674, (2007).

144 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, (2010).

145   Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, (2010).

146   Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, (2010).

147   Harms, A. FIC domains of Bartonella effector proteins -- AMPylation and beyond. *Master Thesis* 2010).

148   Engel, P., Goepfert, A., Stanger, F. V., Harms, A., Schmidt, A., Schirmer, T. & Dehio, C. Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature* **482**, 107-110, (2012).

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

150   Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, (2010).

151   Aharoni, A., Gaidukov, L., Khersonsky, O., Mc, Q. G. S., Roodveldt, C. & Tawfik, D. S. The 'evolvability' of promiscuous protein functions. *Nat Genet* **37**, 73-76, (2005).

152   Bone, R., Silen, J. L. & Agard, D. A. Structural plasticity broadens the specificity of an engineered protease. *Nature* **339**, 191-195, (1989).

153   Goody, P. R., Heller, K., Oesterlin, L. K., Muller, M. P., Itzen, A. & Goody, R. S. Reversible phosphocholination of Rab proteins by Legionella pneumophila effector proteins. *Embo J* **31**, 1774-1784, (2012).

154   Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem* **79**, 471-505, (2010).

155   Etienne-Manneville, S. & Hall, A. Rho GTPases in cell biology. *Nature* **420**, 629-635, (2002).

156   Olson, M. F., Ashworth, A. & Hall, A. An essential role for Rho, Rac, and Cdc42 GTPases in cell cycle progression through G1. *Science* **269**, 1270-1272, (1995).

157   Ricos, M. G., Harden, N., Sem, K. P., Lim, L. & Chia, W. Dcdc42 acts in TGF-beta signaling during Drosophila morphogenesis: distinct roles for the Drac1/JNK and Dcdc42/TGF-beta cascades in cytoskeletal regulation. *J Cell Sci* **112 (Pt 8)**, 1225-1235, (1999).

158   Visvikis, O., Maddugoda, M. P. & Lemichez, E. Direct modifications of Rho proteins: deconstructing GTPase regulation. *Biol Cell* **102**, 377-389, (2010).

159   Schmidt, A. & Hall, A. Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. *Genes Dev* **16**, 1587-1609, (2002).

160   Bernards, A. GAPs galore! A survey of putative Ras superfamily GTPase activating proteins in man and Drosophila. *Biochim Biophys Acta* **1603**, 47-82, (2003).

161   DerMardirossian, C. & Bokoch, G. M. GDIs: central regulatory molecules in Rho GTPase activation. *Trends Cell Biol* **15**, 356-363, (2005).

162   Worthylake, D. K., Rossman, K. L. & Sondek, J. Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. *Nature* **408**, 682-688, (2000).

163   Rittinger, K., Walker, P. A., Eccleston, J. F., Smerdon, S. J. & Gamblin, S. J. Structure at 1.65 A of RhoA and its GTPase-activating protein in complex with a transition-state analogue. *Nature* **389**, 758-762, (1997).

164   Oesterlin, L. K., Goody, R. S. & Itzen, A. Posttranslational modifications of Rab proteins cause effective displacement of GDP dissociation inhibitor. *Proc Natl Acad Sci U S A* 2012).

165   Saraste, M., Sibbald, P. R. & Wittinghofer, A. The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**, 430-434, (1990).

166   Wolf, Y. I., Brenner, S. E., Bash, P. A. & Koonin, E. V. Distribution of protein folds in the three superkingdoms of life. *Genome Res* **9**, 17-26, (1999).

167   Pulliainen, A. T., Pieles, K., Brand, C. S., Hauert, B., Böhm, A., Quebatte, M., Wepf, A., Gstaiger, M., Aebersold, R., Dessauer, C. W. & Dehio, C. Bacterial effector binds host cell adenylyl cyclase to potentiate Gαs-dependent cAMP production. *Proceedings of the National Academy of Sciences* **109**, 9581-9586, (2012).

168   Papageorgiou, A. C., Tranter, H. S. & Acharya, K. R. Crystal structure of microbial superantigen staphylococcal enterotoxin B at 1.5 A resolution: implications for superantigen recognition by MHC class II molecules and T-cell receptors. *J Mol Biol* **277**, 61-79, (1998).

169   Samygina, V. R., Popov, A. N., Rodina, E. V., Vorobyeva, N. N., Lamzin, V. S., Polyakov, K. M., Kurilova, S. A., Nazarova, T. I. & Avaeva, S. M. The structures of Escherichia coli inorganic pyrophosphatase complexed with Ca(2+) or CaPP(i) at atomic resolution and their mechanistic implications. *J Mol Biol* **314**, 633-645, (2001).

170   Wigley, D. B., Davies, G. J., Dodson, E. J., Maxwell, A. & Dodson, G. Crystal structure of an N-terminal fragment of the DNA gyrase B protein. *Nature* **351**, 624-629, (1991).

171   Sissi, C. & Palumbo, M. In front of and behind the replication fork: bacterial type IIA topoisomerases. *Cell Mol Life Sci* **67**, 2001-2024, (2010).

172   Fairweather, N. F., Orr, E. & Holland, I. B. Inhibition of deoxyribonucleic acid gyrase: effects on nucleic acid synthesis and cell division in Escherichia coli K-12. *J Bacteriol* **142**, 153-161, (1980).

173   Bahassi, E. M., O'Dea, M. H., Allali, N., Messens, J., Gellert, M. & Couturier, M. Interactions of CcdB with DNA gyrase. Inactivation of Gyra, poisoning of the gyrase-DNA complex, and the antidote action of CcdA. *J Biol Chem* **274**, 10936-10944, (1999).

174   Jiang, Y., Pogliano, J., Helinski, D. R. & Konieczny, I. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of Escherichia coli gyrase. *Mol Microbiol* **44**, 971-979, (2002).

175   Alvarez-Martinez, C. E. & Christie, P. J. Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**, 775-808, (2009).

7 Acknowledgments

I conducted the work for this thesis in a close collaboration between the groups of Tilman Schirmer in the Core Program Structural Biology and Biophysics and Christoph Dehio from the Focal Area Infection Biology at the Biozentrum of the University Basel.

Foremost, I would like to thank my two supervisors Prof. Tilman Schirmer and Prof. Christoph Dehio who have given me the opportunity to work on this interdisciplinary and exciting project and for their guidance and support through my PhD.

A special thank to the FIC team for fruitful scientific interactions and discussions that heavily contributed to the success of this work.

Many thanks to present and past members of the Dehio and the Schirmer labs for the great working atmosphere – Christoph Wirth, Caroline Peneff, Claudia Massa, Franziska Zähringer, Amit Sundriyal, Frédéric Stanger, Cedric Hutter, Stefanie Kauer, Tillmann Heinisch, Chee-Seng Hee, Camille Peitsch, Dietrich Samoray, Maxime Quebatte, Arto Pulliainen, Raquel Conde, Alexander Harms, Alain Casanova, Jérémie Gay-Fraret, Shyan Huey Low, Yun-Yueh Lu, Kathrin Pieles, Mathias Dick, Houchaima Ben Tekaya, Simone Eicher, Rusudan Okujava, Pauli Rämö, Claudia Mistl, Simone Muntwiler, and Mario Emmenlauer.

I am very grateful to Christophe Wirth and Maxime Quebatte for always having an open ear for me.

485 lab, I will never forget the absolutely hilarious time we had at the lab retreat.

I would like to acknowledge the constant support by the many people of the technical and administrative staff of the 3rd and 4th floors. Thanks you very much for all those things running in the back of the every day business of science.

Thank you, Jessica, for your love, encouragements, and for being endlessly patient with my vile temper. I find happiness in loving you.

I want to acknowledge my dog Guizmo for the funny and relaxing moments we share together.

I express my gratitude to my parents for their unconditional support and for having inculcated me the never say die attitude.

8 Curriculum Vitae

Curriculum Vitae

Name: Arnaud Goepfert

Phone: **office:** +41 79 5297665   **private:** +33 6 18 93 30 57

Email: arnaud.goepfert@novartis.com

Current position: Presidential Postdoctoral fellow

Novartis Institutes for Biomedical Research

CPC/ Structural Biophysics

Novartis Campus, CH-4002, Basel, Switzerland

Date of birth: 30.04.1984

Country of residence: France

Country of origin: France

Education

2008 - 2012	**Ph.D. thesis**
	Structural Biology and Infection Biology, Biozentrum, University of Basel, Switzerland
	Supervisor: Prof. Tilman Schirmer and Prof. Christoph Dehio
	Thesis title: "Fic-mediated adenylylation: catalysis and regulation"
2006 - 2007	**Graduate studies**
	Awarded M.Sc. in Structural Biology and Bioinformatics
	Louis Pasteur University, Strasbourg, France
2002 - 2005	**Undergraduate studies**
	Bachelor in Biochemistry
	Louis Pasteur University, Strasbourg, France

Research and Professional Experiences

2013 - present   **Presidential Postdoctoral fellow**
Center for Proteomic Chemistry/ Structural Biophysics,
Novartis Institutes for Biomedical Research (NIBR)
Supervisor: Dr. Jean-Michel Rondeau

2012 - 2013   **Postdoctoral fellow**
Structural Biology and Infection Biology, Biozentrum,
University of Basel, Switzerland
Supervisor: Prof. Tilman Schirmer and Prof. Christoph
Dehio

2008 - 2011   **Assistant and tutor for 3rd year undergraduate students**
in practical courses "Microbiology", Biozentrum,
University of Basel, Switzerland

2008 - 2011   **Assistant and tutor for 3rd year undergraduate students**
in practical courses "Structural Biology", Biozentrum,
University of Basel, Switzerland

Feb 2007 – Dec 2007   **Trainee in structural biology**, Center for Proteomic
Chemistry, Novartis Institutes for Biomedical Research
(NIBR), Basel, Switzerland

Oct 2006 – Jan 2007   **Trainee in structural biology**, Institut de Génétique et de
Biologie Moléculaire et Cellulaire (IGBMC), Strasbourg,
France

Publications

- **Arnaud Goepfert**, Frédéric V. Stanger, Christoph Dehio, Tilman Schirmer, (2013) **"Conserved inhibitory mechanism and competent ATP binding mode for adenylyltransferases with Fic fold."** PLOS ONE, 8(5).

- **Arnaud Goepfert**, Alexander Harms, Tilman Schirmer, Christoph Dehio (2012) **"Type II Toxin-Antitoxin Loci: The *fic* Family."** Prokaryotic Toxin-Antitoxins, Springer, 177-187.

- Philipp Engel*, **Arnaud Goepfert***, Frédéric V. Stanger, Alexander Harms, Alexander Schmidt, Tilman Schirmer, Christoph Dehio (2012) **"Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins."** Nature, 482(7383), 107-110. ***These authors contributed equally to this work**.

- Dinesh V Palanivelu, **Arnaud Goepfert**, Marcel Meury, Patrick Guye, Christoph Dehio, Tilman Schirmer (2011) **"Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA."** Protein Sci. 20(3):492-499.

- Joerg Kallen, **Arnaud Goepfert**, Anke Blechschmidt, Aude Izaac, Martin Geiser, Gisele Tavares, Paul Ramage, Pascal Furet, Keiichi Masuya, Joanna Lisztwan (2009) **"Crystal structures of Human MdmX (HdmX) in complex with p53 peptide analogues reveal surprising conformational changes."** J. Biol. Chem. 284(13):8812-8821.

Oral Presentations

2012	**Presentation at the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC),** Strasbourg, France
	Title: "Keeping the spirit in the bottle: how Fic proteins regulate their potentially lethal adenylylation activity"
2012	**Rhine-Knee Regio Meeting on the structural biology of macromolecules**, Lausanne, Switzerland
	Title: "Keeping the spirit in the bottle: how Fic proteins regulate their potentially lethal adenylylation activity"
2011	**Biozentrum Symposium**, Basel, Switzerland
	Title: "Fic-mediated AMPylation: catalysis and regulation"
2009	**Rhine-Knee Regio Meeting on the structural biology of macromolecules**, Leissigen, Switzerland
	Title: "*Bartonella henselae* effector protein A: putative role in disturbing host cell signalling by AMPylation"
2009	**Presentation at the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC),** Strasbourg, France
	Title: "FIC domain of the *Bartonella henselae* effector BepC"

Poster Presentations

2012	Poster at the **Biozentrum Symposium**, Basel, Switzerland
	Title: "Keeping the spirit in the bottle: how Fic proteins regulate their
	potentially lethal activity"
2011	Poster at the **Biozentrum for the 10 years of Infection Biology**
	celebration, Basel, Switzerland
	Title: "Fic-mediated adenylylation is controlled by a conserved
	inhibition motif"
2011	Poster at the **Federation of European Microbiological societies**
	congress, Geneva, Switzerland
	Title: "Fic-mediated adenylylation is controlled by a conserved
	inhibition motif"
2011	Poster at the **BioValley Life Science week**, Basel, Switzerland
	Title: "Intra- or intermolecular active site obstruction
	controls Fic-mediated adenylylation"
2009	Poster at the **Biozentrum symposium**, Basel, Switzerland
	Title: "*Bartonella henselae* effector protein A: putative role in
	disturbing host cell signaling by AMPylation"