Structure - function studies on FIC - mediated AMPylation and deAMPylation by class I Fic proteins

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Stefanie Tamegger

Basel, 2023

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Erstbetreuer: Prof. Dr. Tilman Schirmer / Prof. Dr. Christoph Dehio

Zweitbetreuer: Prof. Dr. Dirk Bumann

Externer Experte: Prof. Dr. Markus Wiederstein

Basel, den 22. Juni 2021

Prof. Dr. Marcel Mayor

Dekan

Table of contents

Statement of my thesis

This work was performed in the groups of Prof. Tilman Schirmer and Prof. Christoph Dehio in the Focal Area Structural Biology and Biophysics and the Focal Area Infection Biology, respectively, at the Biozentrum of the University of Basel.

My PhD advisory Committee consisted of:

Prof. Tilman Schirmer

Prof. Christoph Dehio

Prof. Dirk Bumann

Prof. Markus Wiederstein

My thesis is written in a cumulative format. It consists of an abstract, an introduction covering several aspects related to my work, a result section containing one published scientific article and one scientific article in preparation with supplementary results followed by concluding remarks

Abstract

Proteins containing a conserved FIC (filamentation induced by cyclic AMP) domain can be found in all domains of life, where they modify the function of target proteins via post-translational modifications such as AMPylation describing the transfer of AMP onto the threonine, tyrosine, or serine side chain of their respective targets.

First studies on AMPylation activity by two bacterial proteins VopS from *Vibrio parahaemolyticus* and IbpA from *Histophilus somni* revealed inhibition of RhoGTPases causing disruption of the actin cytoskeleton leading to cell death. Since then, the AMPylation activity of several Fic proteins containing a highly conserved FIC domain has been described. Fic proteins are controlled by the presence of an inhibition motif and depending on its location the enzymes can be separated into three classes. Proteins such as human FICD or NmFic from *Neisseria meningitidis* have this inhibition motif either on the N-terminus or the C-terminus of the toxin itself and represent class II and class III Fic proteins, respectively. FICD mediates AMPylation of the Hsp70 chaperone BiP in the endoplasmatic reticulum (ER), which leads to BiP's inactivation when the level of unfolded proteins is low. Recently it was shown that FICD deAMPylates BiP and removes the modification, which recruits the target back into chaperone cycle when the load of unfolded proteins is high. Studies on the class III Fic protein EfFic revealed deAMPylation activity of the protein and that Fic proteins can act as bifunctional enzymes. A conserved glutamate in the inhibition motif of both proteins plays a significant role in regulation of the reactions. Class I Fic proteins such as VbhT from *Bartonella schoenbuchensis* are inhibited by a small protein antitoxin consisting of the inhibition motif. Another example for class I Fic proteins are the effector proteins of the α-proteobacterial genus *Bartonella*, which have a growing number of species and are studied as model for evolution of bacterial pathogenesis. *Bartonella* effector proteins (Beps) contain a diverse ensemble of FIC domains, which have evolved in parallel in three *Bartonella* lineages from a single ancestral toxin-antitoxin module.

In *research article I* we use X-ray crystallography, structural modelling, and phylogenetic analysis to gain more insight into the variety of Beps, which includes nine crystal structures and 99 non-redundant sequences. Minor structural changes of the core FIC domain indicate functional and regulatory variability of Beps.

In *research article II* we show that a recent developed nucleotide quantification assay is a sensitive method to obtain real-time enzymatic progress curves. This assay was chosen to characterize the AMPylation and deAMPylation reaction mediated by the FIC domain of the class I toxin VbhT (VbhT(FIC)) from *Bartonella schoenbuchensis,* which is regulated by its cognate antitoxin VbhA. Autoradiography assays previously revealed that VbhT(FIC) AMPylates the DNA gyrase subunit B (GyrB) leading to its inactivation and abolishing cell growth, which is inhibited when the antitoxin was present. We show that the VbhT/VbhA toxin-antitoxin complex acts as a bifunctional enzyme causing AMPylation and deAMPylation of a 43 kilodalton (kDa) subunit of GyrB (GyrB43). Mutation of the glutamate in the inhibition motif, which is known to have strong modifying effects, reveals enhanced AMPylation and deAMPylation activity, indicating a different role of the glutamate in the two reactions compared to class II FICD and class III EfFic.

1   Introduction

1.1   Post-translational modifications

Post-translational modifications (PTMs) describe a variability of modifications utilized by eukaryotic and bacterial cells, and have been topic of research for decades. These alterations can range from the attachment of a few atoms to polypeptide chains, and are especially used by bacteria to manipulate host cell functions for successful invasion and infection of the host (Hedberg and Itzen 2015).

1.2   Adenylylation (AMPylation)

The most prominent PTM is phosphorylation of proteins and refers to the transfer of the γ-phosphate of adenosine triphosphate (ATP) onto a serine, threonine or tyrosine side chain of host cell proteins mediated by enzymes called kinases. The α-phosphate of ATP can be part of a modification as well, which is called adenylylation or AMPylation leading to the release of pyrophosphate (PPi) and the modified AMPylated target (Yarbrough and Orth 2009) (Hedberg and Itzen 2015). This modification has been known for decades and two functions are reported. AMPylation was first described as a stable post-translational modification in the 1960's with the discovery of the covalent attachment of an adenosine monophosphate (AMP) moiety on *Escherichia coli* (*E.coli*) glutamine synthetase I mediated by glutamine synthetase adenylyltransferase regulating the function of the protein (Kingdon et al. 1967) (Stadtman 2001). The second function is the accumulation of an efficient leaving group in mechanisms that indirectly use the energy input from ATP hydrolysis to allow thermodynamically unfavorable overall reactions (Itzen et al. 2011).

1.2.1   Transient adenylylation (AMPylation)

Enzymes performing or mediating a transient adenylylation reaction catalyze the condensation between the weakly nucleophilic carboxylic acid and the weakly electrophilic phosphate. This type of AMPylation appears in two steps. In the first step ATP and the substrate are bound, and reactive substrate-adenylate as intermediate and pyrophosphate

(PPi) are generated. The process proceeds via a negatively pentavalent phosphor that needs to be stabilized. During the second step a nucleophile reacts with the intermediate leading to the release of the final product and AMP (Figure 1) (Schmelz and Naismith 2009).
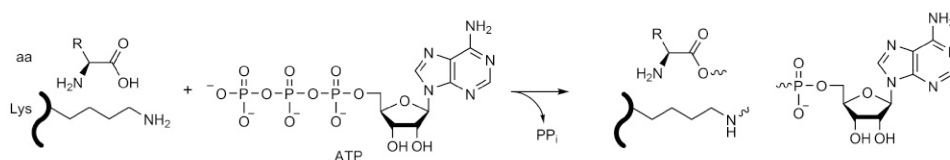

Figure 1: Reaction scheme of the transient AMPylation of proteins (Schmelz and Naismith 2009). This modification of carboxyl groups of amino acids (aa) is utilized to form mixed anhydrides, which act as activated intermediates in the coupling of transfer ribonucleic acids (tRNAs) mediated by the aminoacyl tRNA synthetase.

In several solved crystal structures an arginine, lysine or histidine residue connects the α-phosphate at the reaction center and point mutations of these residues decreased the enzymatic activity. Enzymes catalyzing this reaction such as aminoacyl-tRNA synthetases (aaRSs) and E1 enzymes have diverse roles in the metabolic pathways of prokaryotes and eukaryotes (Schmelz and Naismith 2009) (Pang et al. 2014) .

1.2.1.1 Aminoacyl-tRNA synthetases

Aminoacyl-tRNA synthetases (AaRSs) are a family of enzymes responsible for the first step of protein synthesis, which is an aminoacylation reaction attaching an amino acid to its cognate transfer RNA (tRNA) in a highly specific two-step reaction (Eriani et al. 1990) (Pang et al. 2014). The proteins differ widely in their size and oligomeric state, but two classes of aaRSs have been formed based on their chemical properties, architecture and consensus sequence (Eriani et al. 1990) (Ribas de Pouplana and Schimmel 2001). Class I enzymes tend to appear as monomers, whereas class II aaRSs are usually in dimer or tetrameric form s(Brick et al. 1989) (Cusack et al. 1990) (Ribas de Pouplana and Schimmel 2001). Class I aaRSs contain a Rossmann ATP-binding fold with two consensus sequences. The KMSKS motif stabilizes the pyrophosphate moiety during amino acid activation, the HIGH motif supports the phosphate backbone (Brick et al. 1989). Class II aaRSs have three consensus motifs, in which motif 1 is located at the dimer, and motifs 2 and 3 are part of the aminoacyl activation site (Cusack et al. 1990). The two different cores of the aaRS classes carry out the two-step reaction using different mechanisms. Class I enzymes bind

ATP in an extended conformation, whereas class II aaRSs use a bent conformation to catalyze the reaction (Brick et al. 1989) (Arnez and Moras 1997). In the first step or activation step ATP binds together with the amino acid to the active site of the protein. The α-phosphate of ATP is attacked by the amino acid leading to the formation of 5'-aminoacyl-adenylate, which stays in the active site while pyrophosphate is the other product. In the second step or transfer step the intermediate is transferred to the 2' or 3'-OH of the tRNA amino acid arm, resulting in stable charged tRNA and an AMP moiety is released (Figure 2) (Berg and Offengand 1958) (Giege 2006) (Francklyn and Mullen 2019).
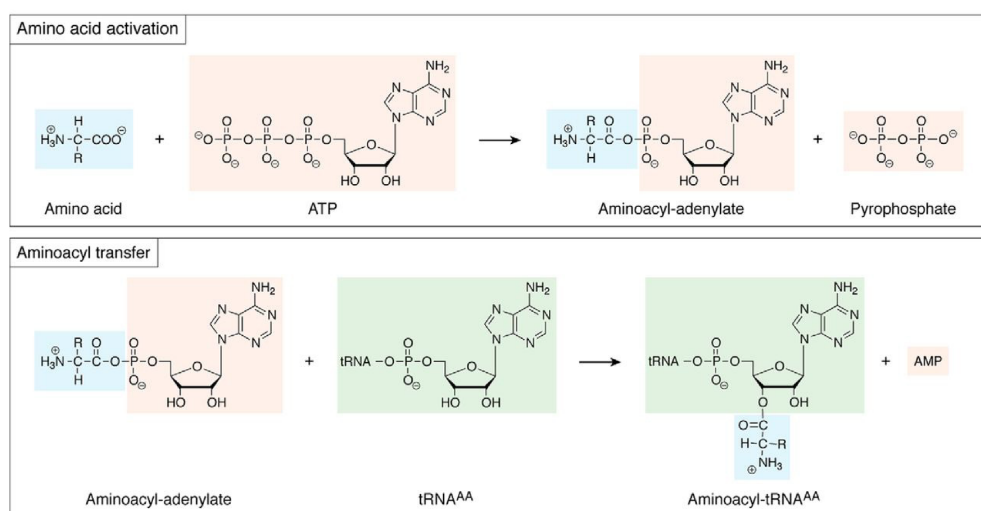
Figure 2: Aminoacylation reaction by aminoacyl-tRNA-synthetases (Francklyn and Mullen 2019). The first step is the activation step, which involves ATP binding leading to an anhydrate adenylate intermediate and PPi. In the second reaction the intermediate is transferred to the tRNA resulting in a stable tRNA molecule and AMP is released.

1.2.1.2 Ubiquitin activation

Transient adenylylation or AMPylation also plays a significant role in the activation of ubiquitin (Ub), which is conjugated to other proteins through a peptide bond between its C-terminal glycine and a primary amine located on the substrate. This process is dependent on the activities of the enzymes E1, E2 and E3. Several protein families were discovered, which are evolutionary related to ubiquitin and are referred to as ubiquitin-like proteins (Ubl's). These proteins share structural and evolutionary similarities to ubiquitin, and are divided into two types depending on the presence of a conjugation (Hochstrasser 2009) (Cappadocia and Lima 2018). E1 activating enzymes are essential for activation of

Ub/Ubl's along the conjugation cascade and catalyze three chemical reactions (Figure 3). In the first step E1 enzymes bind ATP, magnesium and Ub/Ubl's to build up a high-energy level acyladenylate intermediate accompanied by the release of PPi. In the second reaction the catalytic cysteine of E1 attacks the intermediate to induce the building of a high energy thioester bond between the enzyme and Ubl (E1~Ubl) with the release of AMP. In the last step E1~Ubl E2 catalyzes transthioesterification to an E2 resulting in a high energy thioester-linked E2~Ubl product (Figure 3) (Hershko and Ciechanover 1998) (Lu et al. 2010) (Cappadocia and Lima 2018).
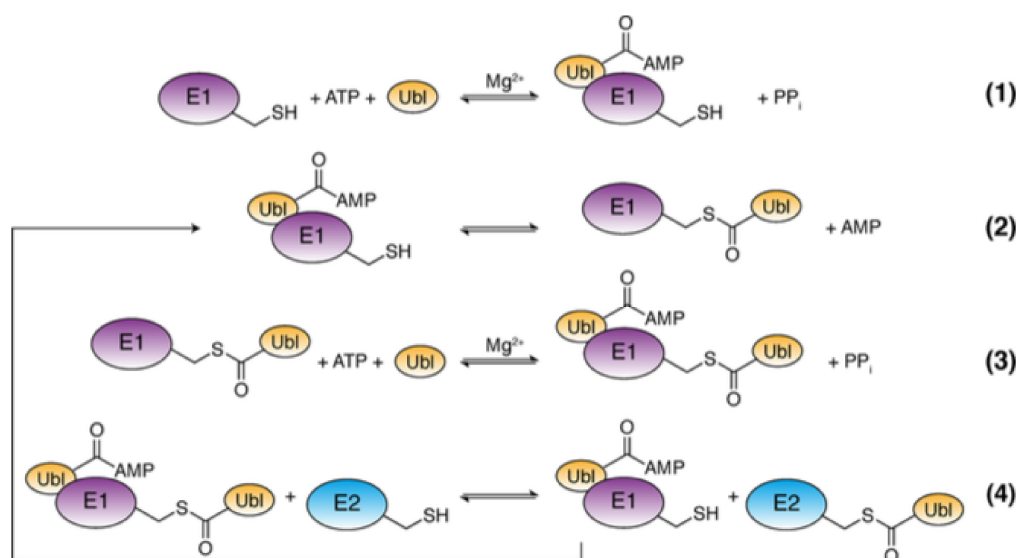
Figure 3: Reaction scheme of E1 mediated AMPylation (Cappadocia and Lima 2018).
In the first step (1) E1 binds Ubl and ATP leading to AMPylation of Ubl, which is then attacked by the catalytic cysteine (2) leading to formation of E1~Ubl thioester. In the third reaction (3) E1 AMPylates another Ubl, which is then transferred to an E2 (4) for further reactions.

E3 is the ubiquitin protein ligase, which catalyzes the linkage of the C-terminal glycine of ubiquitin to the terminal amino group of a lysine chain (Hershko and Ciechanover 1998). These coupled reactions are crucial due to the involvement of ubiquitin in marking proteins for degradation, in promoting or preventing protein interactions and it can affect the activity of proteins (Cappadocia and Lima 2018).

1.2.2   Stable adenylylation (AMPylation)

In contrast to transient adenylylation, stable AMPylation is an irreversible protein PTM, which uses the high energy substrate ATP to affect the function of proteins. It generates a reversible phosphodiester bond and requires a counteracting enzyme to remove the modification (Yarbrough et al. 2009) (Kinch et al. 2009).

1.2.2.1   AMPylation of the Glutamine synthetase

E. coli glutamine synthetase (GS) occupies a very central role in the nitrogen metabolism catalyzing the condensation of ammonia with glutamate to produce glutamine. Glutamine is needed in the biosynthesis of many metabolites, in which some act as a feedback inhibitor for GS binding to different allosteric sites of the protein. This complex mechanism is referred to as cumulative feedback (Woolfolk and Stadtman 1967) (Almassy et al. 1986) (Stadtman 2001).

In the 1960's Earl Stadtman and his team showed that this process is under tight control of a mechanism involving the covalent attachment of AMP onto a specific amino acid residue of the protein (Kingdon et al. 1967) (Stadtman 2001). The modification is mediated by an enzyme called glutamine synthetase adenylyl-transferase (GS-AT), requires the presence of Mg^{2+} and is activated by glutamate (Kingdon et al. 1967). The enzyme contains two domains, which are responsible for two accompanying reactions. The C-terminal domain has adenylyltransferase (AT) activity and is responsible for inactivation of GS by transfer of an adenylyl group from ATP to GS tyrosine 397 (Y397) (Xu et al. 2010). The GS forms a dodecamer (12 subunits) arranged in two rings of six subunits with the modified Y397 obtaining a position at the interface between the subunits and close to the active site (Almassy et al. 1986) (Janson et al. 1986). Depending on the incubation time and concentration of the transferase 1-12 modifications could be obtained. Preparations of GS and inhibition of a certain preparation by various feedback inhibitors is governed by the number of modified residues (Stadtman 2001). Studies showed that the GS-AT catalyzes the opposite reaction using the N-terminal domain, which has adenylylremovase (AR) activity causing removal of the adenylyl group from GS. GS-AT contains a regulatory domain between the two catalytic domains, which binds P_{II}, a signal transduction protein and nitrogen sensor for the cell (Xu et al. 2010). This protein itself exists in two forms. Firstly, in presence of glutamine, a product of GS, the unmodified form stimulates the adenylylation of GS by interactions at the active site of GS-AT. Secondly, in presence of

α-ketoglutarate and absence of glutamine the uridylylated P_{II} supports deadenylylation of GS due to interactions with the AR domain (Stadtman 2001) (Xu et al. 2010). P_{II} uridylylation involves attachment of a uridylyl group from UTP to a tyrosine residue in each of the three subunits mediated by uridylyl transferase (UTase), whereas deuridylylation is a hydrolytic process (Stadtman 2001) (Adler et al. 1975). This complex cascade and therefore tight control of the GS is crucial for its role in the metabolism of nitrogen (Figure 4).
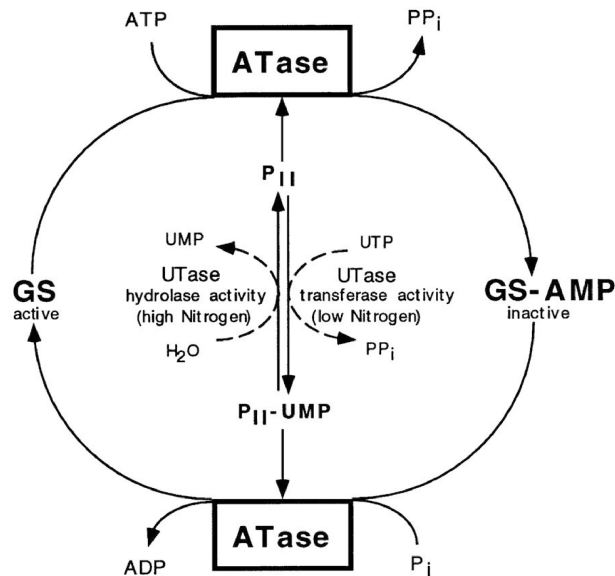
Figure 4: Regulation of glutamine synthetase activity (Jaggi et al. 1997).

The enzyme uridylyl transferase (UTase) acts as a sensor for the nitrogen levels and maintains P_{II} or P_{II}-UMP, which regulates the adenylylation or deadenylylation activity of GS mediated by the ATase or GS-AT.

1.2.2.2   Fic proteins

40 years after the discovery of GS-AT mediated AMPylation the same activity was discovered for several other bacterial effector proteins containing a conserved FIC (filamentation induced by cyclic AMP) domain (Kingdon et al. 1967) (Worby et al. 2009) (Yarbrough et al. 2009). VopS, an effector of *Vibrio parahaemolyticus (V. parahaemolyticus)* causes cell rounding due to the modification of GTPases in HeLa cells and the underlying mechanism was identified as AMPylation (Yarbrough et al. 2009). At the same time another Fic protein called immunoglobulin-binding protein A (IbpA) from *Histophilus somni* (*H. somni*) showed retraction of the plasma membrane and disruption of the actin stress fibers which was also linked to AMPylation of GTPases (Worby et al. 2009).

Fic proteins and their role in affecting host proteins through AMPylation are the main topic of my thesis and will be discussed in more detail in the following sections.

1.3   Classification of FIC domain containing proteins

The original description of Fic was derived from a mutation in the *E.coli fic-1* gene, which led to drastic cell filamentation at elevated cyclic AMP (cAMP) levels and high temperature (Utsumi et al. 1982) (Kawamukai et al. 1988). The precise function of this domain was unknown until the AMPylation activity of VopS and IbpA was discovered 27 years later (Yarbrough et al. 2009) (Worby et al. 2009). *V. parahaemolyticus* is a gram-negative bacterium, which contains two type III secretion systems (T3SS) responsible for the delivery of proteins into the cytosol of host cells during infection inducing autophagy, cell rounding and cell lysis (Daniels et al. 2000) (Ghosh 2004). Infection studies showed that VopS is inhibiting the Ras homolog (Rho) family of guanosine triphosphatases (GTPases) (RhoGTPases) including Ras-related C3 botulinum toxin substrate (Rac), Rho and cell division control protein 42 homolog (Cdc42) causing severe rounding of HeLa cells (Figure 5A, middle panel). The effector contains a C-terminal FIC domain with a conserved motif consisting of the amino acids HPFx(D/E)GNGR. It was revealed that mutation of the histidine abolishes the cell rounding phenotype (Figure 5A, right panel) (Yarbrough et al. 2009). A similar phenotype was observed for the fibrillar surface antigen IbpA from *H. somni*, which has two conserved FIC domains, Fic1 and Fic2, at the N-terminus of the protein. Transfection experiments with each Fic motif expressed separately revealed the collapse of the cytoskeleton due to the transfer of AMP onto the RhoGTPases RhoA, Rac and Cdc42 by Fic2 (Figure 5B, middle panel). IbpAFic2 contains a conserved Fic motif HPFx(D/E)GN(G/K)R and mutation of the histidine shows that the disruption of the cytoskeleton is abolished after transfection of HeLa cells (Figure 5B, right panel) (Worby et al. 2009).
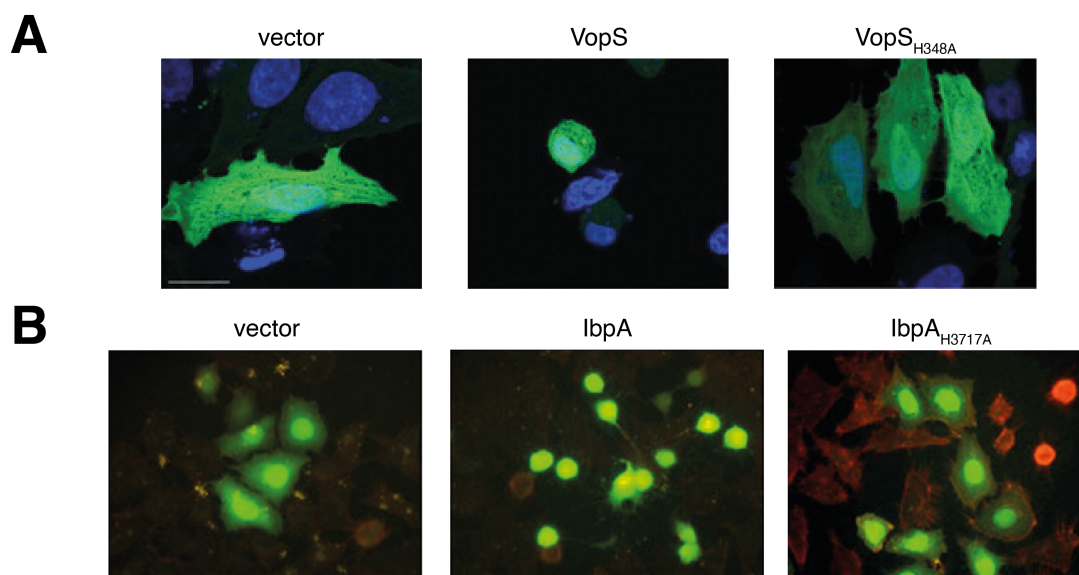
Figure 5: VopS and IbpA cause disruption of the actin cytoskeleton (modified from (Yarbrough et al. 2009) (Worby et al. 2009)).

A: Transfected HeLa cells revealed a severely rounded phenotype due to VopS-induced inhibition of RhoGTPases (middle). Mutation of the conserved histidine of the Fic motif to alanine (H348A) abolished the effect (right). An empty vector was used as control (left) (Yarbrough et al. 2009).

B: Transfected HeLa cells showed disruption of the host actin cytoskeleton in presence of wildtype IbpAFic2 (middle). Mutation of the histidine to alanine (H3717A) in the Fic2 signature motif abolished the effect (right). An empty vector was used as a control (left) (Worby et al. 2009).

Proteins containing this domain belong to the Fic/Doc (Fido) family, which includes death on curing (Doc) domains and the avirulence protein B (AvrB) (Yarbrough et al. 2009) (Kinch et al. 2009). Fic and Doc domains share a conserved central motif composed of the amino acids HxFx(D/E)GN(G/K)R, which is not present in AvrB (Kinch et al. 2009). The Doc toxin can be found in the *E. coli* phage P1 and is part of the toxin-antitoxin module Doc/PhD. In absence of its antitoxin the toxin interacts with the 30S ribosomal subunit leading to inhibition of the translational elongation (Lehnherr et al. 1993).

1.3.1 Three-dimensional structure of Fic proteins

In 2006 crystal structures of Fic proteins were solved as a part of the Structural Genomic Initiative revealing a common architecture (Veyron et al. 2018). The four solved structures HpFic from *Helicobacter pylori* (*H. pylori*) (PDB: 2F6S), NmFic from *Neisseria*

meningitidis (*N. meningitidis*) (PDB: 2G03), BtFic from *Bacteroides thetaiotaomicron* (*B. thetaiotaomicron*) (PDB: 3CUC) and SoFic from *Shewanella oneidensis* (*S. oneidensis*) (PDB: 3EQX) show several conserved α-helices (α1-α6) as depicted in figure 6 using NmFic as representative. The FIC core domain of smaller proteins is composed of helices α2-α5, but can be decorated with additional helices like in NmFic and HpFic (Kinch et al. 2009) (Palanivelu et al. 2011). The loop located between helices α4 and α5 contains the consensus sequence HxFx(D/E)GNGRxxR and together with the N-terminal cap of helix α5 it forms the catalytic center of the protein (Kinch et al. 2009) (Palanivelu et al. 2011). The role of the active center in modification of target proteins will be further discussed in the following chapters.
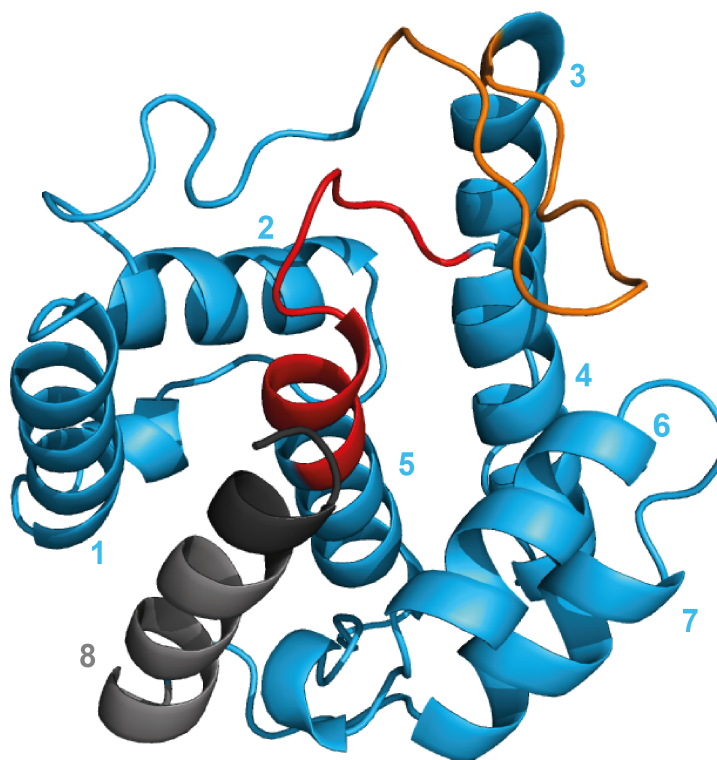
Figure 6: Domain organization of NmFic (modified from (Kinch et al. 2009)).
Cartoon representation of the Fic protein NmFic (PDB: 2G03). The core domain (α2-α5) with additional helices (α1, α6-α7) are depicted in blue. The signature motif is displayed in red and the target recognition site named flap in orange. Helix α8 (grey) plays are role in the regulation of Fic proteins and will be discussed in the following chapters.

A β-sheet called flap is located between helices α2 and α3, which plays an important role in the interaction of the target or surrogate target. It can be found in various orientations but closes down onto the active site of the protein as interactions of IbpAFic2

in complex with its target AMPylated Cdc42 display (Kinch et al. 2009) (Xiao et al. 2010) (Palanivelu et al. 2011) (Goepfert et al. 2013). IbpAFic2 is mainly interacting with the switch1 and switch2 region of Cdc42. The tyrosine residue, which gets modified, is positioned in the switch1 region of the GTPase. The side chains of a lysine and a leucine residue located on IbpAFic2 form a clamp, which locks the tyrosine side chain in a suitable orientation for modification. Mutation of those residues show reduced AMPylation activity of the Fic protein (Xiao et al. 2010). Crystal structures revealed that the flap's conformation ranges from a solid β-sheet like in the structure of IbpAFic2 to a partly disordered loop as seen in *Bartonella* effector protein A (BepA) (Xiao et al. 2010) (Palanivelu et al. 2011). The sequence of the flap is not conserved across Fic proteins, but bioinformatic studies clustered the sequences into seven subfamilies based on a stretch of four residues NLTK of IbpAFic2, which were predicted to be important for protein substrate recognition (Khater and Mohanty 2015).

It was unknown how Fic proteins regulate their AMPylation activity until the discovery of a conserved mechanism of ATP-binding-site obstruction involving an α-helix ($\alpha_{inh}$) inhibiting the toxic effect of a Fic protein found in *Bartonella schoenbuchensis* (*B. schoenbuchensis*). The FIC domain containing toxin VbhT from *B. schoenbuchensis* induces cell filamentation upon ectopic expression in *E. coli*, which is repressed when the toxin is co-expressed with its protein antitoxin VbhA. Comprehensive analysis identified 158 bacterial *vbhA* homologues containing a central (S/T)xx(I/L)EG motif. The crystal structure of the FIC domain of VbhT (VbhT(FIC)) in complex with VbhA revealed that the antitoxin is embracing the toxin tightly with three anti-parallel α-helices (Figure 7A). The α_{inh} is located on the C-terminus of the antitoxin close to the active site of the toxin competing with ATP binding. Comparison with already solved crystal structures revealed that an equivalent of the helix containing the inhibition motif can be part of the FIC domain itself. Fic proteins can be separated into three classes based on the location of the α_{inh} (Figure 7B). Class I Fic proteins interact with an antitoxin consisting of the inhibitory α-helix. Class II and class III Fic proteins have their α_{inh} either at the N-terminus or at the C-terminus of the FIC fold itself, respectively. The overall consensus sequence of the inhibition motif consists of the amino acids (S/T)xxxE(G/N) and revealed that the glutamate is strictly conserved interacting with the second arginine of the Fic motif, which is important for toxin inhibition (Engel et al. 2012). 90% of the Fic proteins containing the

canonical active site motif could be classified based on the assumption that all of these proteins have a mechanism inhibiting their enzymatic activity (Goepfert et al. 2013).
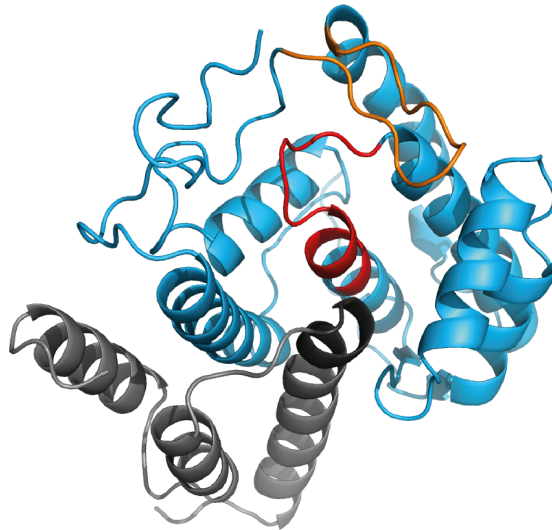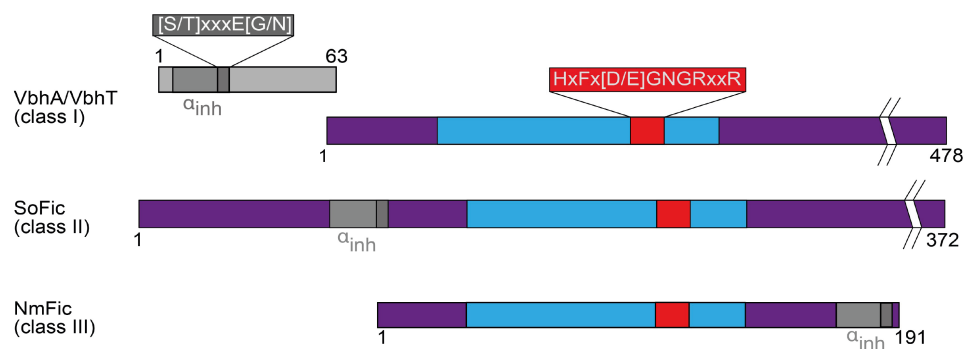
Figure 7: Domain architecture of the VbhT/VbhA toxin-antitoxin complex and classification of Fic proteins (modified from (Engel et al. 2012) (Goepfert et al. 2013)).

A: Overview of the VbhT toxin (PDB: 3ZC7) shown as a blue cartoon with the Fic motif in red and the flap in orange. The antitoxin VbhA is depicted in grey with the inhibition motif in dark grey.

B: Classification of Fic proteins based on the location of the $\alpha_{inh.}$ and representatives.

The following chapters give more insight into the different classes of Fic proteins and their identified targets.

1.3.2   Class I Fic proteins

Class I FIC domain containing toxins (FicT) were identified by Psi-Blast for finding of VbhA-homologous peptides, which are encoded in the upstream region of the *fic* loci. Class I toxins are characterized by the presence of a small, cognate antitoxin (FicA), which contains the inhibitory α-helix $α_{inh}$. The toxin can only be active if the $α_{inh}$ is removed or degraded, whereas class II and III toxins require intramolecular processes to remove their inhibition. Expression of the toxin without the antitoxin is lethal to the bacteria therefore this toxin-antitoxin arrangement is crucial for survival of the bacteria (Engel et al. 2012) (Goepfert et al. 2013) (Harms et al. 2016). The regulatory arrangement of FicTs and their respective antitoxins is suitable for the definition of type II TA systems (Engel et al. 2012) (Goepfert et al. 2013).

1.3.2.1   Bacterial toxin-antitoxin systems

Bacterial toxin-antitoxin (TA) systems are small genetic elements that encode a stable toxin and its cognate, instable antitoxin. The toxin is responsible for inhibition of cell growth caused by interference with vital cell processes while the antitoxin protects the cell from the toxic activity. The molecular activities of these systems have been studied for decades, but the biological function of the majority is still unknown and so far only three biological functions have been discovered. TA modules play a role in post-segregational killing, abortive infection and persister formation or antibiotic tolerance (Unterholzner et al. 2013) (Page and Peti 2016) (Harms et al. 2018).

Four major systems have been described based on the origin and mechanism of the antitoxin. Type I and III TA systems have small noncoding RNAs functioning as antitoxins, whereas type II and IV systems consist of small antitoxin proteins (Figure 8) (Unterholzner et al. 2013) (Harms et al. 2018).
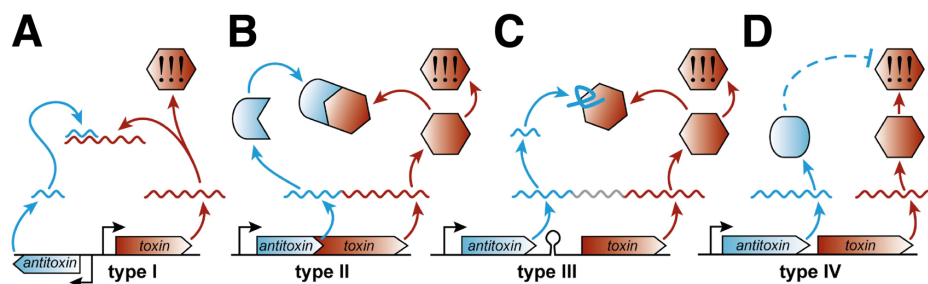
Figure 8: Overview of the major toxin-antitoxin systems (Harms et al. 2018).
The scheme shows the four different interactions of how toxins (brown) are affected by their cognate antitoxins (blue) in type I-IV (A-D) TA modules. Genetic loci and position of the promotors are displayed in the respective colours with black arrows. RNAs are curly lines and active toxins are highlighted by exclamation marks.

Type I TA systems contain a noncoding RNA antitoxin and a protein toxin. The small regulatory antisense RNAs (sRNAs) base-pair to the messenger RNA (mRNA) of the toxin resulting in inhibition of its translation (Thisted et al. 1994). Under normal conditions this mechanism inhibits the binding of the ribosome and the RNA is degraded by RNase III. In the presence of stress, the pool of antitoxin sRNA is decreasing, which leads to translation of the toxin mRNA (Brantl and Jahn 2015) (Page and Peti 2016). An example for such a system is the *symR*/*symE* module of *E. coli*. *SymE* expression is additionally controlled by an SOS-response regulated transcriptional repressor and degradation by the Lon protease (Fernandez De Henestrosa et al. 2000) (Kawano et al. 2007).

Type III TA modules encode a mRNA endonuclease toxin and a RNA antitoxin, which forms pseudoknots binding to the toxin (Page and Peti 2016) (Harms et al. 2018). The best studied member of this class is the *toxIN* model system of *Pectobacterium atrosepticum.* The toxin *toxN* is preceded by a short palindromic repeat followed by a tandem array of nucleotide repeats. ToxN is a RNase that cleaves the *toxIN* transcript into active antitoxin sRNAs, which are post-transcriptionally processed into active subunits. Those subunits inhibit the toxin in a 1:1 stochiometry (Fineran et al. 2009) (Short et al. 2013).

Like type II TA modules type IV toxin-antitoxin modules are proteins but compared to the other systems they do not interact with each other. The antitoxin counteracts the activity of the toxin indirectly by affecting the target. Examples for such types of modules, which have been studied only mechanistically so far are *cbeA*/*cbtA* of *E. coli* K-12 and its two paralogs (Harms et al. 2018). The CbtA toxin and its antitoxin interact independent from each other with the cytoskeletal proteins MreB and FtsZ. The toxin prevents growth due to inhibition of the polymerization of the proteins, while the antitoxin CbeA promotes and stabilizes MreB and FtsZ, leading to cytoskeletal filament bundling (Masuda et al. 2012) (Page and Peti 2016).

Type II TA systems represent the largest and best-studied class of such modules consisting of thousands of loci found in free-living bacteria (Pandey and Gerdes 2005).

Type II antitoxins inhibit their toxins by forming tight complexes via direct interactions which usually cause interference with catalysis at the active site or with the binding of the target. The antitoxin usually contains of a N-terminal deoxyribonucleic acid (DNA)-binding domain critical for transcriptional autoregulation and a C-terminal domain, which directly binds to the toxin resulting in its inhibition (Harms et al. 2018). Antitoxins are highly susceptible to proteolysis in response to cellular signaling such as stress conditions, which as a consequence causes activation of the toxins. This leads to growth arrest caused by inhibition of replication due to the toxin suppressing DNA gyrase activity or translation by for example cleaving mRNA (Bernard et al. 1993) (Christensen-Dalsgaard et al. 2010). The majority of type II antitoxins in *E. coli* are degraded by the Lon protease, but some are target of ClpP and its adapters ClpA and ClpX (Makarova et al. 2009).

Several type II setups contain RNA antitoxins, which use a variety of unrelated protein folds to inhibit their cognate toxins. The majority of type II toxins are endoribonucleases (RNases), which often adopt a microbial RNase fold and bind directly to the ribosome to cleave ribosome-associated mRNA (Yamaguchi et al. 2011).

1.3.2.2 FicTA modules

The regulatory module of class I FIC domain proteins contains a small protein called antitoxin, which can inhibit the activity of its cognate toxin due to tight protein-protein interactions. This interaction abolishes the toxin induced growth inhibition of host cells. The regulatory arrangement fulfills the central definition of type II TA modules, therefore class I toxins are named FIC domain toxins (FicT) and their cognate antitoxins FIC domain antitoxin (FicA) (Goepfert et al. 2013) (Harms et al. 2015). Class I Fic proteins are part of a separate branch in the family of FIC domain containing proteins, which is around five to ten percent of all analyzed proteins. The majority of the proteins in that group comprise a regular HxFx(D/E)GNGRxxR motif or closely related sequence leading to the hypothesis that most of the FicTA modules perform AMPylation as their default activity (Goepfert et al. 2013).

One model for a FicTA module is the VbhT/VbhA complex of *B. schoenbuchensis*, which plays a significant role in *research article II* (Engel et al. 2012). A distantly to VbhT related FicT toxin is YeFicT of *Yersinia enterocolitica* strain 8081, consisting of a canonical signature motif (HPFREGNGRAQR). EcFic, which was the first discovered FIC domain containing protein is not suitable as a model due to its altered active site (Harms et al. 2015) (Stanger et al. 2016). Expression of different FicTs without their cognate

antoxins resulted in strong growth defects in *E. coli* due to their AMPylation activity (Figure 9). Mutation of the catalytic histidine in the active side abolished the effect. Subsequent expression of the cognate antitoxins reversed the growth defects and demonstrates that the activity leads to a bacteriostatic state common for many TA modules (Pedersen et al. 2002).
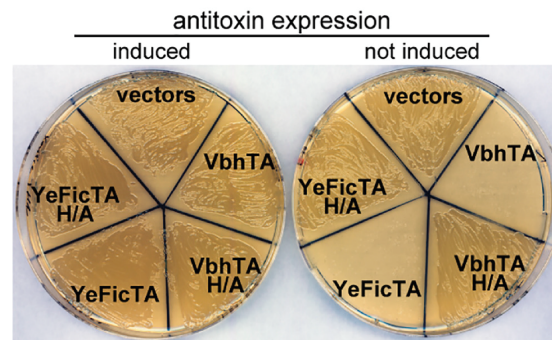
Figure 9: Inhibition of cell growth by FicT toxins is abolished upon antitoxin expression (Harms et al. 2015).

Right: FicT toxins VbhT and YeFicT inhibit *E. coli* growth upon induction in absence of their cognate antitoxins. Expression of catalytically inactive FicT toxins harboring a mutation of the catalytic histidine do not show growth inhibition.

Left: Cell growth defects by VbhT and YeFicT are inhibited due to expression of their cognate antitoxins.

The Fic-1 toxin of *Pseudomonas aeuroginosa* consists of a canonical active side motif mediating AMPyaltion activity, which is regulated via interaction with a specific inhibitor anti-Fic-1 (AntF1) in a dose-dependent manner. A 1:27 molar ratio of Fic-1 to AntF1 suppresses the activity significantly, while a 1:3 ratio completely abolished the toxic effect of Fic-1 on GyrB (Lu et al. 2016).

1.3.2.3 *Bartonella schoenbuchensis* VbhT/VbhA

VbhT is an interbacterial effector protein of the mammalian pathogen *B. schoenbuchensis*, which is comprised of a N-terminal FIC domain and a C-terminal relaxase-derived *Bartonella* effector protein (Bep) intracellular domain (BID) domain (Figure 10C) (Dehio et al. 2001). The BID domain mediates translocation of proteins into target cells via type IV secretion system or conjugation machinery (Engel et al. 2012) (Harms et al. 2017).

Expression of VbhT in *E. coli* leads to growth arrest after induction, which is repressed when the catalytic histidine of the conserved Fic motif HPFREGNGRTLR is mutated to alanine (VbhT$_{H136A}$). Co-expression of the toxin with VbhA resulted in repression of the VbhT effect (Figure 10A). The antitoxin is encoded by a small open reading frame *vbhA* located directly upstream of *vbhT*. Wildtype VbhT causes cell filamentation, whereas the mutant VbhT$_{H136A}$ or the presence of VbhA (VbhT/VbhA) lead to normal cell morphology.

Autoradiography experiments with radioactive labeled α^{32}P-ATP revealed *in vitro* AMPylation of a putative *E. coli* target protein of approximately 80kDa mediated by wildtype VbhT, which is not observed when VbhT$_{H136A}$ or the VbhT/VbhA complex is used (Figure 10B) (Engel et al. 2012).
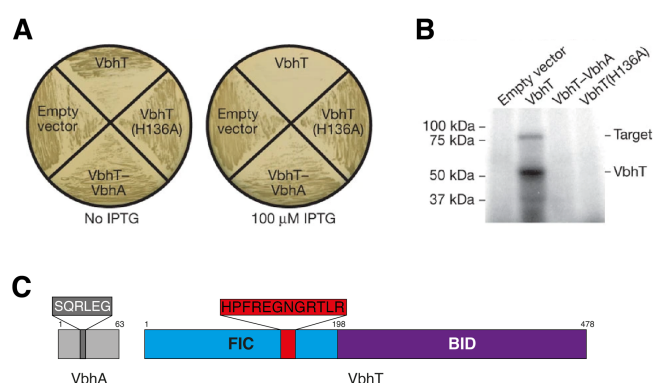
Figure 10: Expression of VbhA suppresses the toxic activity of VbhT (modified from (Engel et al. 2012) (Stanger et al. 2017)).

A: Expression of VbhT shows toxic effects upon IPTG induction (right panel). Mutation of the catalytic histidine to alanine VbhT$_{H136A}$ or co-expression of the toxin with its cognate antitoxin VbhA (VbhT/VbhA) does not have any effect on bacterial growth due to inhibition of the toxic activity.

B: Autoradiography assay with cell lysates of *E. coli* ectopically expressing the same plasmids as shown in A in presence of radioactive α^{32}P-ATP display target AMPylation at 80kDa when only VbhT is used.

C: Domain architecture of VbhT/VbhA with the signature motif in red and the inhibition motif in dark grey.

Several crystal structures of the toxin FIC domain in complex with its antitoxin were solved and reveal that the antitoxin consists of three α-helices embracing the α1 helix of the toxin (Kinch et al. 2009) (Engel et al. 2012). The C-terminus of the antitoxin contains the α_{inh} with the amino acid sequence SQRLEG, which takes a position close to the N-cap of the helix following the active site loop. The residues serine 20 (S20) and glutamate 24 (E24) of the α_{inh} form a hydrogen bond and a salt bridge with the arginine 147 (R147)

causing interference with ATP binding (Engel et al. 2012). Crystal structures of the FIC domain of VbhT (VbhT(FIC)) in complex with VbhA and ATP revealed that the nucleotide is in a pocket between the helices α4 and α6, and the β-hairpin flap. The 3'-hydroxyl of the ribose forms a hydrogen bond with the E24 of the $α_{inh}$ and the triphosphate is interacting with the anionic nest built by the N-terminus of helix α5 of the toxin (Goepfert et al. 2013). These interactions lead to an orientation, which is not suitable for AMP transfer resulting in the inhibition of the toxin. The $α_{inh}$ glutamate was identified to play a major role in the inhibitory effect since mutation of the residue leads to sufficient space for the nucleotide to reorientate into a conformation favorable for AMPylation. The crystal structure of VbhT(FIC) in complex with the antitoxin containing a glycine instead of a glutamate at position 24 ($VbhA_{E24G}$) shows a curved conformation of the triphosphate. The γ-phosphate is close to the ribose moiety and forms a tight salt bridge to the R147 in the FIC motif (Goepfert et al. 2013). The presence of growth defects in *E. coli* cells upon expression of this modified toxin-antitoxin complex revealed that the glutamate is indeed important for the inhibition of the toxic effects of VbhT due to competition with ATP substrate binding (Engel et al. 2012).

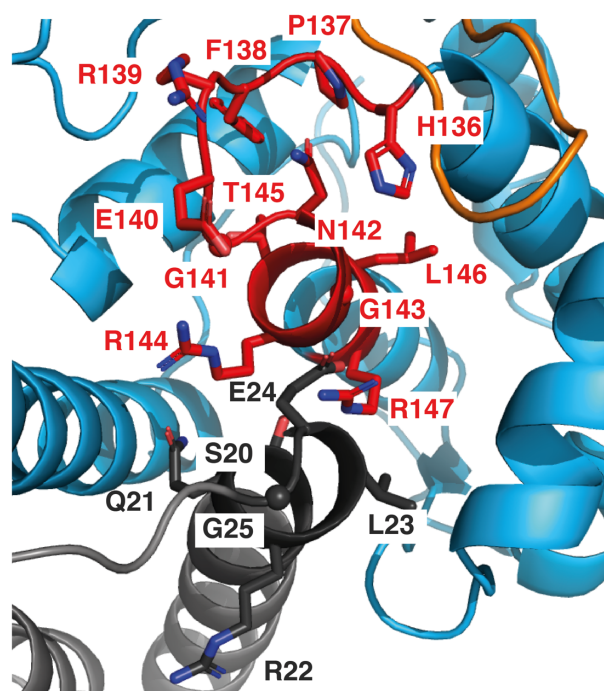**Figure 11: Crystal structure of the wildtype VbhT(FIC)/VbhA (PDB: 3ZC7) (modified from (Goepfert et al. 2013).**
VbhT(FIC) is shown as a blue cartoon with the FIC motif in red and the target binding flap in orange. The antitoxin VbhA is depicted as grey cartoon with the inhibition motif in dark grey. The residues of the FIC motif and the inhibition motif are displayed as sticks or spheres (G141, G143, G25).

The active site of the VbhT toxin is partly covered by a β-hairpin called flap, which recognizes the modifiable side chain to the active site as it was shown in the IbpAFic2:Cdc42 complex (Xiao et al. 2010). The structure of the wildtype toxin-antitoxin complex revealed an additional density located above the active site close to the flap. It was interpreted as a four residues peptide, which has an extended conformation and is associated to the edge of the two-stranded β-hairpin with three main chain-main chain interactions (Goepfert et al. 2013).

1.3.2.4 Targets of *B. schoenbuchensis* VbhT/VbhA

Radioactive *in vitro* AMPylation assays with *E. coli* lysates were performed to unravel the mechanism behind the growth inhibition mediated by VbhT and an AMPylated peptide of around 80kDa, which is similar to the theoretical weight of GyrB. Autoradiography of ectopically expressed GyrB and its paralog ParE of *E. coli* and *B. schoenbuchensis* revealed modification by full-length VbhT in bacterial lysates *in vitro* (Figure 12) (Harms et al. 2015).
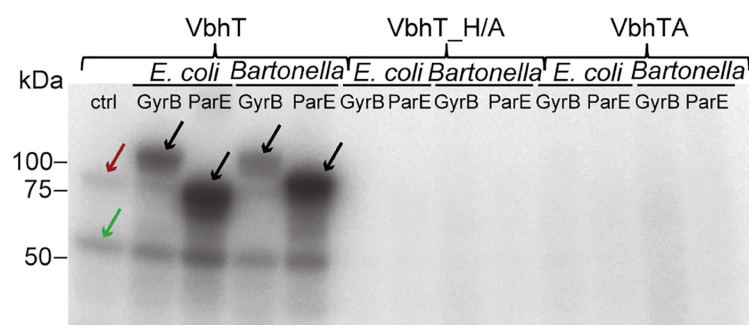
Figure 12: AMPylation of GyrB and ParE by VbhT (Harms et al. 2015).
Autoradiography of cleared *E. coli* lysates with different expressed full-length VbhT constructs and target candidates as indicated. VbhT shows autoAMPylation at 50kDa (green arrow), and causes AMPylation of endogenous GyrB (red arrow) as well as ectopically expressed GyrB and ParE of *E. coli* and *B. schoenbuchensis* (black arrows).

GyrB and ParE belong to the family of topoisomerases, which play a significant role in controlling cellular DNA topology by maintaining negative supercoiling and removing chromosomal entanglements (Sissi and Palumbo 2010). During transcription the structure of double stranded (ds) DNA gets positively supercoiled by polymerases and helicases, while generation of chromosomal knots or catenates can happen during

replication. Depending on the degree of supercoiling and the removal of catenates the cleavage of the phosphodiester bond is needed, which is followed by DNA strand passage and resealing of the linkage. These processes are mediated by topoisomerases, which can be separated into two different types (Liu and Wang 1987) (Postow et al. 2001). Type I topoisomerases cause a single-stranded DNA break, type II topoisomerases generate double-stranded DNA breaks (Liu and Wang 1979) (Brown et al. 1979). GyrB and ParE are part of the type II topoisomerases DNA gyrase and topoIV, which have various functions. Using the free energy generated by ATP hydrolysis DNA gyrase introduces negative supercoils into closed dsDNA leading to a maintained chromosome during transcription, whereas topo IV removes catenants and knots after replication (Schvartzman et al. 2013).

DNA gyrase is responsible for the introduction of negative supercoils into covalently closed dsDNA molecules by using the free energy produced by ATP hydrolysis. It is composed of two subunits, GyrA and GyrB, with the active complex being an A_2B_2 dimer of around 400kDa (Wigley et al. 1991).

The GyrA subunit consists of a N-terminal winged helix domain, followed by a tower domain, a coiled-coil domain and a C-terminal DNA binding domain. The GyrB subunit has 3 domains, a ATPase domain responsible for ATP hydrolysis, a central transducer domain and a TOPRIM domain, which is interacting with the GyrA subunit (Figure 13) (Sissi and Palumbo 2010). A 43kDa N-terminal fragment of GyrB (GyrB43) comprises of the ATPase domain and the transducer domain. The structure of this protein in complex with the substrate analog AMPPNP revealed a tight dimer with the contact interface mainly located in the ATPase domain and a hole through which the DNA is transferred (Wigley et al. 1991). The analog is bound to the canonical site of the GHKL-type ATPase domain, but also interacts with the QTK loop of the transducer domain and an N-terminal "brace" belonging to the other subunit of the dimer. The monomer can bind ATP, but hydrolysis can only be performed in the dimeric form, which leads to product release and dissociation into the monomer. The Monomer-monomer interaction is too weak for stabilization of a dimer in absence of the nucleotide (Ali et al. 1993) (Stanger et al. 2014).

The dimerization interface has three contact areas or gates, labeled N-, DNA- and C-gate, which work in a coordinated manner to allow DNA passage. The process of how type II topoisomerases induce topological changes is a two-gate mechanism. The enzyme binds a segment of dsDNA called the gate segment or G-segment, to the DNA-gate. This

process is followed by trapping of the transfer segment or T-segment due to ATP-induced dimerization of the N-gate. The cleavage of the G-segment and opening of the N-gate allows the T-segment to pass through the gate. In the end steps the resealed G-segment is released as well as the T-segment due to the opening of the C-gate (Figure 13B) (Sissi and Palumbo 2010).
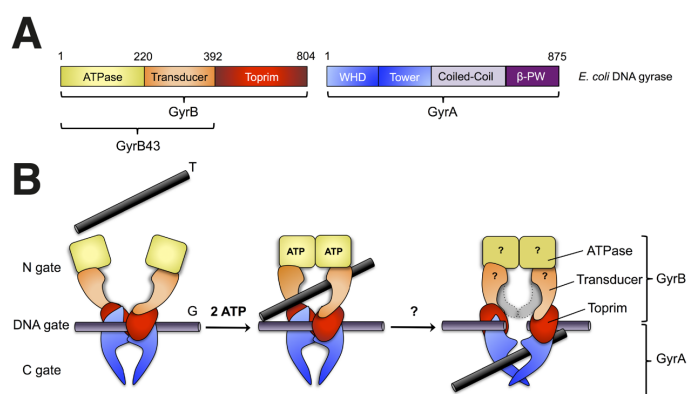
Figure 13: Domain organisation of *E.coli* DNA gyrase and scheme of its activity (Stanger et al. 2014).
A: *E. coli* DNA gyrase is comprised of two domains. GyrB contains of an ATPase (yellow), a transducer (orange) and a TOPRIM (red) domain. GyrA has a winged-helix (WHD, blue) a tower (blue), a coiled-coil (light purple) and a β-pinwheel (β-PW, dark purple) domain.
B: A dsDNA segment (G-segment or G in the scheme) is locked in the DNA-gate, which opens the N-gate and allows binding of the transfer segment (T-segment or T). ATP binding leads to dimerization of the domains and trap the T-segment in the transducer domain. Upon ATP hydrolysis the G-segment is cleaved resulting in opening the DNA-gate through which the T-segment can pass.

Due to its essential functions in the cell cycle DNA gyrase is target of quinolones and coumarins, which are antibacterial drugs inhibiting DNA supercoiling. Quinolones such as ciprofloxacin affect the A subunit of DNA gyrase by interrupting with the DNA breakage and the reunion steps for DNA supercoiling. Coumarin drugs like novobiocin are naturally occurring compounds, which affect the B subunit by suppressing ATP hydrolysis (Reece and Maxwell 1991) (Ali et al. 1993). Crystal structures of the ATPase domain with several coumarin drugs revealed that the sugar ring of the drugs partially overlaps with the adenine ring of ATP at the binding site (Maxwell and Howells 1999). Novobiocin introduces large conformational changes of the dimer subdomains. A loop in the active site is stabilized by strong dimeric contacts and results in a closed conformation of the N-gate

upon ATP binding while the presence of novobiocin leads to an open conformation (Lamour et al. 2002).

Topo IV has a similar distribution of functional domains. The C-terminal domain of ParC has the DNA binding domain, whereas the ATP binding region is located on the N-terminal domain of ParE like in GyrB (Sissi and Palumbo 2010).

The effect of VbhT on time-dependent inactivation of GyrB activity was investigated, which showed that AMPylation abrogates the function of GyrB and blocks all effects *in vitro*. Those effects are repressed when the antitoxin is co-expressed, or the toxin is lacking the catalytic histidine (Harms et al. 2015).

Inhibition of DNA gyrase would lead to a decrease of negative supercoiling of cellular DNA, which can be caused due to treatment of *E. coli* with high novobiocin concentrations resulting in collapse of negative supercoiling. Expression of VbhT or VbhT(fic) that is strongly impaired with bacterial growth only caused a slight stretch of the topoisomerase distribution toward DNA relaxation. Flow cytometry experiments to record DNA gyrase inhibition in single cells revealed that VbhT only has a weak but detectable effect, while VbhT(fic) elicited a stronger response. This effect can happen due to differences in expression or due to the sterically hinderance of the C-terminal BID domain which could be partially compensated by a stronger effect of the FIC domain. This also led to the suggestion that inhibition of DNA gyrase is not the primary driving force of VbhT-induced growth defects (Harms et al. 2015). Expression of VbhT and VbhT(fic) greatly inhibited DNA segregation and bacterial cell division, which was first evidence for inactivation of topo IV *in vivo*, but only cells expressing the FIC domain displayed a highly condensed nucleoid morphology. An independent read-out for topo IV inactivation showed that both constructs induce detectable DNA knots and therefore robust topo IV inhibition. VbhT shows clear signs of catenates with various node numbers while VbhT(fic) reveals a weaker effect. This difference compared to full-length VbhT confirms stronger inhibition of DNA gyrase because the accompanying inactivation of gyrase and topo IV causes rapid arrest of DNA replication and consequently prevents formation of catenates (Witz and Stasiak 2010).

The induced growth arrest relies on modification of a tyrosine residue in the B subunits of DNA gyrase and topo IV causing inactivation of the proteins. The modified residues were identified as tyrosine 109 (Y109) in GyrB and tyrosine 105 (Y105) in ParE using MS. Y109 is a highly conserved residue in the ATP lid loop of GyrB and borders the

binding site (Brino et al. 2000) (Stanger et al. 2014). Experiments with recombinant proteins revealed that AMPylation mediated by an inhibition-relieved mutant of VbhT(FIC)/VbhA$_{E24G}$ prevents all ATP-dependent activities of both targets such as supercoiling of a relaxed reporter plasmid by DNA gyrase and relaxation of supercoiled reporter plasmid and decatenation by topo IV. Due to the inhibition of ATP hydrolysis by VbhT-mediated AMPylation the enzyme acts similar to drugs like aminocoumarins or gyramides, which have a similar mechanism on suppressing topoisomerase activity (Harms et al. 2015).

1.3.2.5 *Bartonella* effector proteins (Beps)

VbhT is formed by the fusion of a FicT-like FIC domain and a relaxase-derived BID domain, which is also the most common and probably ancestral setup of effectors belonging to the α-proteobacterial genus *Bartonella* (Pieles et al. 2014) (Harms et al. 2017). Those so-called Beps are secreted by VirB/D4 type IV secretion system (T4SS) and have a bipartite domainarchitecture consisting of a N-terminal FIC domain and a C-terminal BID domain. An oligo-saccharide binding fold or OB-fold is connecting the two domains (Palanivelu et al. 2011). *Bartonella* consists of a group of ubiquitous mammalian pathogens and is characterized by a stealth infection strategy due to successful host-specific adaptation unlike frontal attack strategies like *V. parahaemolyticus* or *H. somni* causing maximal host cell damage facilitated by VopS or IbpA (Merrell and Falkow 2004) (Harms and Dehio 2012) (Harms et al. 2017). Three phylogenetic lineages have been linked to *Bartonella's* host adaptability which is enabled by lineage-specific acquisition of a VirB/D4 T4SS and parallel evolution of effector proteins (Harms et al. 2017).

Phylogenetic analysis show that *Bartonella* comprises two deeply rooting lineages (*B.apis* and *B. tamiae*) and the lineage of eubartonellae, which split into previously described four lineages (L1, L2, L3, L4) and *Bartonella australis* (Guy et al. 2012). The sequenced strains of *B. baciliformis* of L1, and L2 do not contain a VirB/D4 T4SS, while *B. ancashensis* has this specific secretion system encoded on a distinct position on the chromosomal loci such as L3 and L4. Comparison of the genetic organization of *virB/D4/bep* genes in L3, L4 and *B. ancashensis* revealed that the genes of *B. ancashensis* and L4 are encoded on a single genomic island, while L3 shows a more dispersed organization (Harms et al. 2017).

Bep genes have evolved via duplication, diversification and reshuffling of domains leading to a modular gene architecture (Schulein et al. 2005) (Engel et al. 2011). Most of the Beps encoded by L3, L4, and *B. ancashensis* consist of a characteristic and ancestral FIC-BID domain organization (Figure 14) (Engel et al. 2011) (Harms et al. 2017). The C-terminal BID domain is present in all sequenced Beps as it functions as secretion signal for transport via VirB/D4 T4SS (Engel et al. 2011). Some Beps of L4 and *B. anchashensis* contain more than one BID domain suggesting that the domains have secondarily evolved functions in addition to acting as secretion signal (Engel et al. 2011). Some effectors of L3 and L4 contain an array of tyrosine phosphorylation motifs, which are modified by host kinases after translocation. An effector of *B. ancashensis* also displays tyrosine phosphorylation motifs as an extension after the regular FIC-BID domain architecture. This leads to the suggestion that these motifs have evolved de novo three times in lineages L1, L3 and L4 (Harms et al. 2017).

Beps are part of class I Fic proteins due to presence of an antitoxin called BiaA. An inhibition relieved BiaA in complex with Bep2 displayed successful AMPylation of vimentin in autoradiography (Pieles et al. 2014). *Bartonella* carrying a VirB/D4 T4SS all encode one or two copies of BiaA, with a conserved glutamate in the inhibition motif suggesting a regulation mechanism for Bep FIC domain proteins. Bep3 and Bep4 do not have a separate antitoxin, but carry an N-terminal extension homologous to BiaA (Figure 14) (Harms et al. 2017).
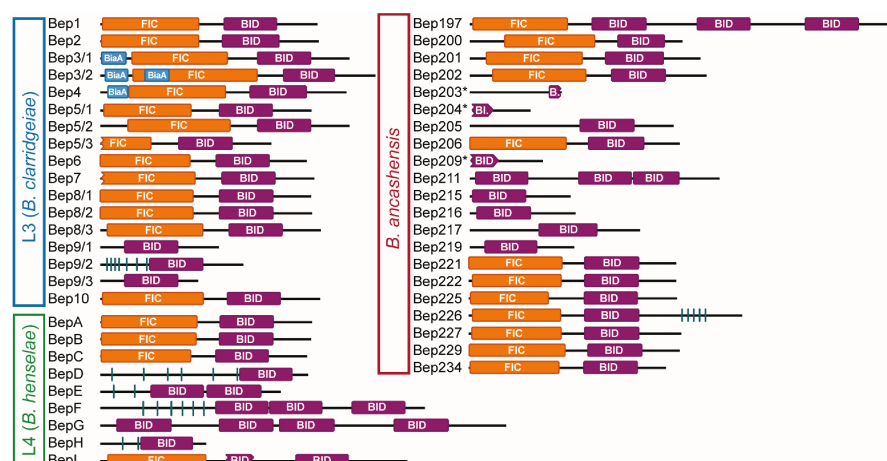
Figure 14: Domain architecture of three Bep repertoires (Harms et al. 2017).
Domain organization of ortholous groups of Beps found in L3, L4 and *B. ancashensis*. Two representative organisms of L3 (*B. clarridgeiae)* and L4 (*B. henselae*) are shown. The FIC domain is displayed in orange,

the BID domain in purple and BiaA-like modules are shown in blue. The presence of tyrosine phosphorylation sites is depicted as vertical, cyan lines.

The first solved crystal structure of an effector from *Bartonella* was BepA from *B. henselae*, which consists of a canonical active site motif and displays the same helical arrangement (α1-α7) as other Fic protein structures such as HpFic or NmFic, but also reveals some differences compared to known crystal structures. Helix α8 is packed peripherally connecting an OB-fold and a specific element between α1 and α2 only present in Beps (Palanivelu et al. 2011).
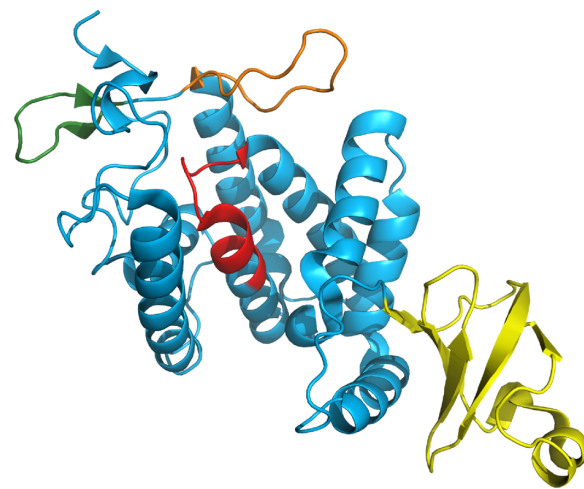
Figure 15: Crystal structure of a truncated BepA (BepA$_{tr}$) (residues 12-302) (PDB: 2VY3) (Palanivelu et al. 2011).
The FIC domain of BepA$_{tr}$ is shown in red, the target binding flap is depicted in orange, the OB fold in yellow. Another element specific for Bep (Bep element) is shown in green, which will be further discussed in *research article I*.

Detailed investigations of the FIC domains of L3 and L4 revealed that only orthologs of Bep1, Bep2 and Bep3 consistently express a canonical active site suggesting AMPylation activity (Harms et al. 2017). Bep2 mediates AMPylation of the protein vimentin, which is a component of the cytoskeleton and has no homology to small GTPases therefore representing a new class of targets (Palanivelu et al. 2011). Bep1 selectively AMPylates the four members of the Rac subfamily (Rac1, Rac2, Rac3 and RhoG), while IbpA and VopS modify a wide range of RhoGTPases on a conserved tyrosine or threonine residue in the switch region of the target. The crystal structure of the Bep1 FIC domain reveals an extended target recognition flap in comparison to other FIC structures. The tip

of this elongated β-sheet forms two salt bridges with the GTPase, which is crucial for selectivity (Dietz et al. 2021). Bep1 and Bep2 orthologs display similar motifs at this position in the flap region, while paralogous Beps show higher variability (Harms et al. 2017).

Other orthologs of Beps have conserved, but non-canonical signature motifs like most of L3 Bep4s, which have a cysteine at the position of the catalytic histidine suggesting other catalytic activities than AMPylation (Harms et al. 2017). The L4 effector BepC has a lysine located at the position of the metal- coordinating aspartate or glutamate. The translocated effector recruits GEF-H1 to the plasma membrane causing activation of the RhoA/ROCK signalling pathway and leading to formation of actin stress fibers (Marlaire and Dehio 2021). A third group of Beps such as BepA/B or Bep5 containing no consistent sequence motif have either different activities or play a role in target binding without having enzymatic activity (Harms et al. 2017).

A detailed analysis of the structural aspects of Bep FIC domains and their possible enzymatic functions is displayed in *research article I.*

1.3.3   Class II Fic proteins

Structural comparison of several Fic proteins showed that the inhibitory α_{inh} can be part of the core domain itself either at the N-terminus or at the C-terminus, which lead to classification of those proteins into class II and class III Fic proteins, respectively. Class II FIC domain containing proteins have the α-helix with the inhibition motif at the N-terminus of the protein as it was revealed in the crystal structure of SoFic from *S. oneidensis* and BtFic from *B. thetaiotaomicron*. After classification of two-thirds of all Fic proteins it was shown that class II contains the majority of them, but little is known about their molecular activities and biological functions (Engel et al. 2012) (Harms et al. 2016).

The human Fic protein Huntingtin yeast interacting protein E (HYPE) or FICD is one of the best studied class II proteins. It is ubiquitously expressed in mammalian cells with highest expression levels in the pituitary gland, pancreatic islets, liver and prefrontal cortex, where it interacts with huntingtin, a protein that plays a significant role in development and long-term memory storage (Faber et al. 1998) (Sanyal et al. 2015). FICD consists almost entirely of α-helices, which build the two tetratricopeptide repeat (TPR) domains and the FIC domain, with a single α-helix as a linker between them. It is a stable,

asymmetric dimer with interactions between the two FIC domains, while the TPR domain does not contribute to dimerization. The crystal structures of FICD with non-hydrolysable variants of ATP and ADP suggested that the protein is involved in protein AMPylation. FICD is located in the endoplasmic reticulum (ER) as a type II, single-pass transmembrane protein with a short cytoplasmic section and a large catalytic domain facing the ER lumen (Bunney et al. 2014). At this location FICD AMPylates the heat shock protein 70 (Hsp70) chaperone BiP, which plays a role in the unfolded protein response (UPR). Modification of BiP by FICD enhances the ATPase activity of the target necessary for refolding of unfolded proteins upon ER stress (Sanyal et al. 2015). Threonine 518 (T518) was identified as AMPylation site of BiP using MS, whereas no modification appeared on serine 365 (S365) and threonine 366 (T366). Mutation of T518 to alanine (T518A) abolished all modification by FICD and altering the ability of BiP to form oligomers. Mutations distant from the modification site had impact on the AMPylation efficiency by affection the access to T518 or changing the overall conformation. Certain mutations lead to stronger modifications by enhancing the access of FICD to BiP due to the reduction of oligomerization interactions or interfere with ATP binding or hydrolysis of BiP causing changes in conformation of the target. Mutants affecting the ATP binding of BiP and undergoing allosteric transitions, but are not able to mediated nucleotide hydrolysis, also showed enhanced AMPylation rates compared to wildtype BiP (Preissler et al. 2015).

AMPylation of BiP is connected to its conformation and the compact ATP-bound state is the preferred substrate of FICD. Substrate-free, ATP-associated BiP can enter the substrate cycle, which is influenced by the concentration of unfolded proteins in the ER or AMPylation by FICD causes inactivation. The ATPase activity of the modified target is resistant to the stimulation of J-domain protein co-factors and reduces BiP's ability to form complexes with its substrates (Preissler et al. 2017). Therefore, FICD inactivates Bip, but the modification is temporarily since the levels of AMPylated BiP are depending on the changes in the ER protein folding load (Preissler et al. 2015). FICD's activity is inhibited by the conserved glutamate in the inhibition motif and relief of this suppression is likely possible via allosteric transitions (Preissler et al. 2015) (Preissler et al. 2017).

Recently it was shown that if the levels of ER stress increases, BiP can be activated via deAMPylation and recruited into the chaperone cycle. FICD is necessary for BiP inactivation via AMPylation and is also able to remove the modification through deAMPylation of the target protein (Preissler et al. 2017). A more detailed view on the deAMPylation of target proteins is depicted in chapter 1.4.4.

1.3.4   <u>Class III Fic proteins</u>

Class III Fic proteins are a small group of proteins, which have their regulatory α_{inh} located on the C-terminus of the Fic core domain such as NmFic from *N. meningitidis* and HpFic from *H. pylori*. Initial studies of those two proteins revealed strong autoAMPylation activity in autoradiography, which is significantly reduced after mutation of the catalytic histidine or residues involved in ATP binding (Xiao et al. 2010). Class III Fic proteins are highly conserved single-domain proteins and can be found across all classes of Proteobacteria suggesting that they are stand-alone autoregulated functional proteins (Stanger et al. 2016). Detailed analysis of the functional and structural role of class III Fic proteins was executed with NmFic due to its size and known crystallization conditions (Xiao et al. 2010) (Engel et al. 2012).

1.3.4.1   *Neisseria meningitidis* NmFic

First experiments with the class III Fic protein showed significant autoAMPylation in its wildtype form, which was reduced when the catalytic histidine or other residues corresponding to ATP binding residues in IbpAFic2 were mutated (Xiao et al. 2010). The auto-modification site was traced to a tyrosine residue (Y183) located on the α_{inh} of the protein using MS. Based on the proximity of the residue to the active side auto-modification was suggested to happen after partial unfolding or detachment of that helix. NmFic did not show additional bands after *E. coli* lysate was added indicating absence of targets or inhibition of the protein. Mutation of the serine 182 (S182) and glutamate 186 (E186) to alanines revealed transfer of radioactivity onto a protein of approximately 80kDa and displayed enhanced autoAMPylation with an additional modification site at tyrosine 188 (Y188). Strong target AMPylation was also detected when the whole α_{inh} was deleted (NmFic$_{\Delta 8}$) confirming the importance of these residues while only weak auto-modification appeared due to the lack of tyrosine acceptor residues (Engel et al. 2012). Independent of the presence of a target, which is discussed in detail in the following chapter, wildtype NmFic and inhibition-relieved NmFic$_{E186G}$ display strong autoAMPylation (Figure 17) (Stanger et al. 2016).

Crystallization of NmFic revealed a tetrameric arrangement in different forms and ligation states formed by two independent interfaces. Interface 1 is built by apolar

interaction of phenylalanine 70 (F70) and tyrosine 77 (Y77) with their symmetry mates, and two salt bridges formed between arginine 71 (R71) and glutamate 102 (E102) (Figure 16). This interface is strongly conserved among the analyzed class III proteins. Interface 2 consists of the apolar residues L155 and F159, and the two salt bridges of R149 and E156. Compared to interface 1 the residues of interface 2 are not as strongly conserved as the other residues. Instead, the interacting residues reveal covariation with striking charge changes in a subset of class III proteins. To analyze the stability of the NmFic wildtype tetramer in solution the dissociation constants were measured and both values were determined in the low micromolar range. The presence of ATP additionally leads to a better stabilization of the tetramer. Charge reversal point mutations of either E156 or E102 to arginine (E156R or E102R) lead to disruption of the tetramer but revealed a concentration-dependent monomer/dimer equilibrium. The crystal structures displayed virtually identical interfaces compared to the interfaces of wildtype NmFic. A double mutant NmFic$_{E102R,E156R}$ resulted in a monomeric mutant (NmFic$_{mono}$), which is not able to form oligomers (Stanger et al. 2016).
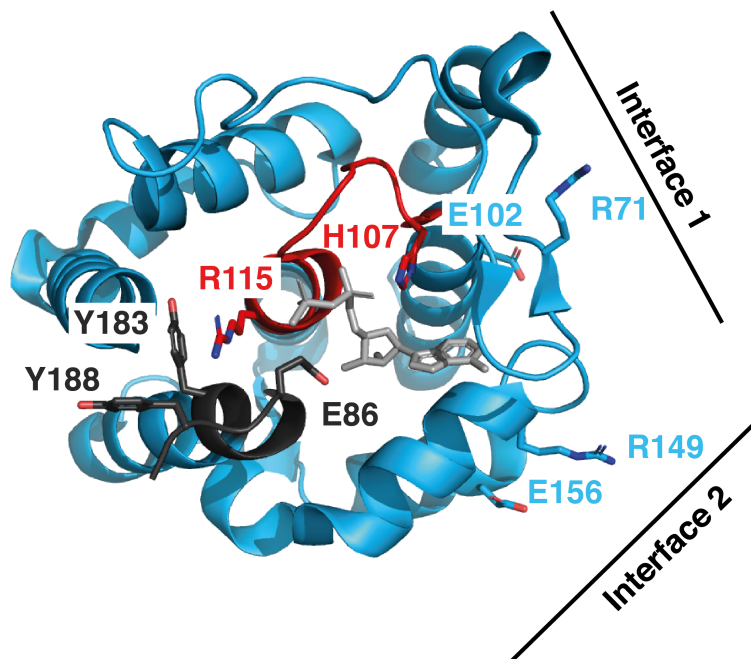
Figure 16: Interface of NmFic (modified from (Stanger et al. 2016)).
Cartoon representation of NmFic (PDB: 3S6A) in blue with the active side in red and the inhibition motif in dark grey. E102 and R71 are part of interface 1 and E156 and R149 are part of interface 2. Both interfaces are important for tetramerization. The modifiable tyrosines Y188 and Y183 are shown in dark grey as well as the inhibitory glutamate E86. AMPPNP shown in grey sticks.

1.3.4.2   Targets of NmFic

Similar to VbhT expression of $NmFic_{E186G}$ induced growth inhibition of ectopically expressing *E. coli*, while no effect was observed with the wildtype (Engel et al. 2012). Wildtype NmFic ($NmFic_{wt}$) does not show any AMPylation impact on GyrB whereas $NmFic_{E186G}$ successfully AMPylates *N. meningitides* GyrB and its orthologs in *E. coli* and *Mycobacterium tuberculosis*. While GyrB is modified no AMPylation was detected for ParE of *N. meningitidis* and *E. coli* (Figure 17). Mutation of the previously identified tyrosine 109 (Y109) of GyrB completely abolishes the effect of $NmFic_{E186G}$ (Harms et al. 2015) (Stanger et al. 2016).
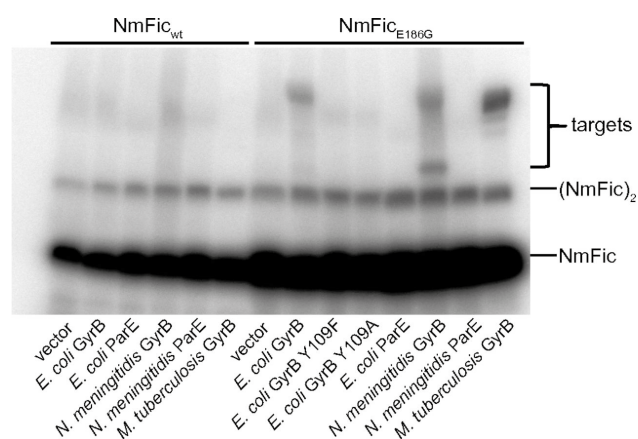
Figure 17: Identification of GyrB as target of NmFic (Stanger et al. 2016).
Autoradiography revealed AMPylation of GyrB (lanes 8, 12, 14) after incubation with $NmFic_{E180G}$, which is not detected in presence of wildtype NmFic.

The interface mutants $NmFic_{E102R}$, $NmFic_{E156R}$ and $NmFic_{E102R,E156R}$ ($NmFic_{mono}$) reveal growth defects, which are not present when the catalytic histidine $NmFic_{H107A}$ is mutated. This effect correlates with GyrB43 AMPylation as shown in autoradiography assays. The interface mutants in combination with a H107A mutant completely abolish modification of GyrB. All interface mutants show AMPylation activity whereas $NmFic_{E156R}$ has a stronger effect compared to the other mutants probably due to the involvement of residue 102 in recognition of the target protein. This showed that

tetramerization of NmFic causes incompetent autoAMPylation and target AMPylation (Stanger et al. 2016).
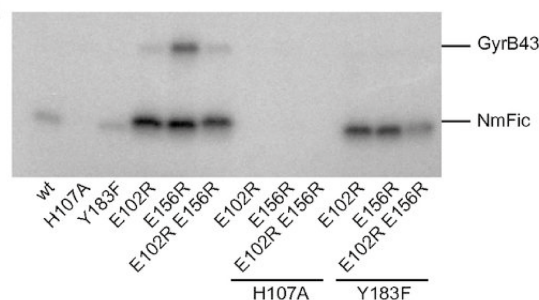
Figure 18: Disruption of the tetramerization interface causes GyrB43 AMPylation (Stanger et al. 2016). Mutation of the residues E102 and E156, which are important for tetramerization, revealed AMPylation of GyrB43 in autoradiography. Addition of a histidine mutant abolished both autoAMPylation and target AMPylation, while mutation of the Y183 only inhibited GyrB43 modification.

1.4 AMPylation and deAMPylation mediated by Fic proteins

1.4.1 Post-translational modifications catalyzed by Fic proteins

FIC domain containing proteins are versatile on molecular and functional levels and can catalyze different PTMs. Most Fic proteins mediate nucleotidyl transfer involving residues of the active side and the target binding region flap. Variations in nucleotide biding can shift the balance from AMPylating enzymes to enzymes having a different substrate specificity. The activities of Doc or AnkX are catalyzed by the same enzymatic machinery, but the proteins have a different substrate binding conformation compared to the typical nucleotide-binding mode. Doc transfers a γ-phosphate from ATP acting as a kinase, while AnkX transfers phosphocholine from CDP-choline after nucleophilic attack. The majority of Fic proteins containing the canonical active site perform AMPylation of target proteins (Harms et al. 2016).

1.4.2 AutoAMPylation of Fic proteins

Prior to the identification of targets AMPylated by Fic proteins automodification by many of these enzymes was already observed. VopS, IbpAFic2, NmFic and HpFic execute AMPylation activity in absence of a target and this modification requires the catalytic histidine of the Fic motif (Kinch et al. 2009) (Xiao et al. 2010). AutoAMPylation was also

shown for *Bartonella henselae* BepA and *Bartonella rochalimae* Bep2 (Palanivelu et al. 2011) (Pieles et al. 2014). After identification of the inhibitory α-helix α_{inh} autoradiography showed autoAMPylation by Fic proteins of class I (VbhT(FIC)/VbhA), class II (SoFic) and class III (NmFic), which is boosted after mutation of the glutamate in the α_{inh} (Figure 19) (Engel et al. 2012) (Goepfert et al. 2013).
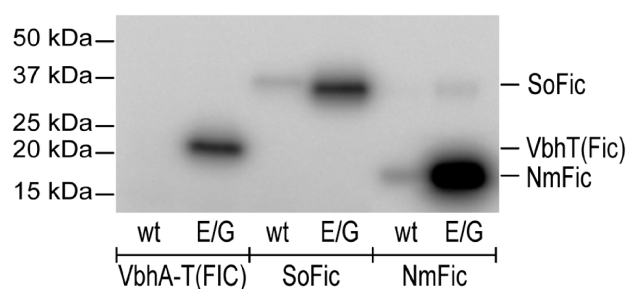
Figure 19: AutoAMPylation of representative FIC domain containing proteins of the three classes (Goepfert et al. 2013).

Incubation of Fic proteins in their wildtype form with radioactive ATP displayed no or weak autoAMPylation in autoradiography, while mutation of the inhibitory glutamate resulted in successful automodification.

The acceptor sites of NmFic automodification were traced to tyrosine 183 (Y183) and tyrosine 188 (Y188) in the α_{inh} using MS. While Y188 is in a favorable position, Y183 is strictly conserved and buried in the hydrophobic core of the protein. Mutations in the α_{inh} such as $NmFic_{S182A,E186A}$ lead to target AMPylation and enhanced autoAMPylation of the protein. Weak autoAMPylation is still observed even after deletion of the whole α_{inh} ($NmFic_{\Delta8}$) (Engel et al. 2012) (Stanger et al. 2016).

Additionally, NmFic tetramerization renders the enzyme incompetent for successful autoAMPylation and target AMPylation, which is abolished when the residues of interface 1 are mutated ($NmFic_{E102R}$, $NmFic_{E156R}$, $NmFic_{mono}$). Combination of the interface mutants and mutation of the catalytic histidine also suppresses automodification and target modification. Circular dichroism (CD) spectroscopy and differential scanning fluorimetry (DSF) using the autoAMPylation competent $NmFic_{mono}$ revealed that partially unfolding of the α_{inh} upon modification leads to the disordered C-terminus to be long enough for Y183 to reach the active site. AutoAMPylation occurs in *cis* with the reaction velocity being independent from the total concentration of the enzyme. $NmFic_{mono}$ contains eight helices in aqueous solution. During automodification the first seven helices are in place, while the α_{inh} is unfolded from position 178 (glycine 178) onward confirmed with resonance (NMR)

spectroscopy. After modification of Y183 the unfolded helix cannot adopt its original position and conformation. The functional role of this crucial residue was revealed after mutation of the tyrosine to phenylalanine. GyrB43 AMPylation is successfully accomplished by the three interface mutants, but not by the respective Y183F variants. Remaining automodification appears due to the presence of three other tyrosines (Y184, Y185 and Y188). Time courses of target AMPylation and autoAMPylation show that NmFic$_{E156R}$ AMPylates GyrB43 with a lag phase, which is not present when the enzyme is pre-activated. Therefore, only the autoAMPylated form of the enzyme can perform GyrB43 AMPylation. NmFic$_{wt}$ shows GyrB43 AMPylation at lower enzyme concentrations with a sharp transition from 250nm to 500nm, which is similar range as the monomer-to-tetramer transition. NmFic is controlled by two processes, tetramerization and *cis*-autoAMPylation resulting in a complex regulatory mechanism. Due to the strong conservation of the interfaces among class III proteins, it is suggested that the tetrameric arrangement is common in this class of proteins (Stanger et al. 2016).

A class I Fic protein revealed hints for a similar regulatory mechanism. Class I Fic proteins are enzymatically controlled by the presence of its cognate antitoxin. Studies with the *Pseudomonas fluorescens* (*P. fluorescens)* Fic-1 showed that the tyrosine residue 5 (Y5) might play a similar role. A Fic-1 Y5A mutant lacking autoAMPylation almost completely lost its activity in GyrB AMPylation. It is therefore possible that this tyrosine on the toxin is crucial in absence of the antitoxin (Lu et al. 2016).

1.4.3 Structural and functional aspects of the AMPylation reaction mediated by Fic proteins

AMPylation mediated by proteins with a conserved FIC domain was first presented for VopS and IbpA, which both target GTPases of the Rho family leading to collapse of the actin cytoskeleton and cell death. The importance of the strictly conserved active site HxFx(/D/E)GNGRxxR in AMPylation was revealed in mutational and bioinformatic studies (Yarbrough et al. 2009) (Worby et al. 2009) (Roy and Mukherjee 2009) (Mattoo et al. 2011). As described in chapter 1.3.1 the core of the FIC domain consists of four helices (α2-α5) with the active site located between α4 and the N-terminal cap of α5 (Kinch et al. 2009) (Palanivelu et al. 2011). The crystal structure of IbpAFic2 in complex with its AMPylated target Cdc42 and BepA in complex with Mg^{2+} and PPi gave first insights into

the catalytic mechanism of the AMPylation reaction (Xiao et al. 2010) (Palanivelu et al. 2011). The IbpAFic2$_{H3717A}$-Cdc42 complex structure displays the essential role of the histidine in transferring a proton from the tyrosine leading to a nucleophilic attack of the α-phosphate of ATP, while the structure of BepA reveals the possible location of the ATP substrate using PPi and the coordination of the substrate via Mg^{2+} and the glutamate 163. Solving the structure of this complex led to a first proposal for the catalytic mechanism of Fic proteins. This model suggested the role of the conserved histidine as a general base for deprotonation of the target hydroxyl. The hydroxyl side chain is in-line with the scissile Pα -O3α bond and a nucleophilic attack generates a penta-coordinated transition state stabilized by the cation and followed by phosphoester bond cleavage to obtain AMPylated target and PPi as leaving group (Xiao et al. 2010) (Palanivelu et al. 2011).

The regulation of Fic protein mediated AMPylation remained elusive until the discovery of an inhibitory α-helix α_{inh} containing the amino acids (S/T)xx(I/L)EG, which obstructs productive ATP binding. The crystal structure of VbhT in complex with its antitoxin VbhA shows the inhibition motif at the C-terminal part of the α_{inh} and is located close to the N-cap of the helix $\alpha5$ of the ATP binding site. Investigations of the inhibitory mechanism using NmFic with the α_{inh} on the C-terminus of the FIC domain itself, in complex with the non-hydrolysable ATP analogue adenylyl imidodiphosphate (AMPPNP) displays the non-productive orientation of the α-phosphate as crucial position. The α-phosphate is needed to be in-line with the scissile Pα -O3α bond for successful AMPylation and in presence of the inhibitory helix this position is occluded by histidine H107 and asparagine N113. Deletion of the whole α_{inh} shows that the ATP γ-phosphate is bound to the second arginine of the Fic motif (R118) using the same position as the inhibitory glutamate of the NmFic wildtype and revealing its crucial function in the inhibition of toxin mediated AMPylation (Engel et al. 2012).
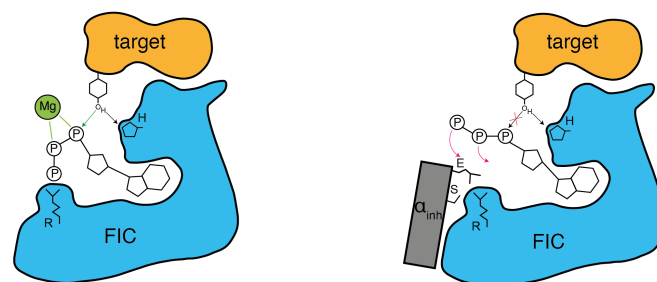
Figure 20: Schematic representation of the AMPylation reaction (left) and its inhibition (right) (modified from (Engel et al. 2012)).

Left: Absence of the α_{inh} leads to interaction of the γ-phosphate of ATP with the second arginine of the FIC motif. This is accompanied by reorientation of the α-phosphate, which is coordinated by the Mg^{2+}. The phosphate is in-line with the target hydroxyl side chain leading to a nucleophilic attack after proton transfer mediated by the catalytic histidine.

Right: The presence of the α_{inh} results in interaction with the inhibitory glutamate, which leads to reorientation of the α-phosphate into a position not suitable for the in-line attack.

Further investigations revealed that mutation of the inhibitory glutamate to glycine in several Fic proteins of different classes such as VbhT(FIC)/VbhA, SoFic and NmFic enhances the autoAMPylation activity and suggest a common inhibitory mechanism as shown in figure 20. Crystal structures with bound ATP locate the base in a pocket built by $\alpha4$, $\alpha6$ and the flap while the ribose 3' hydroxyl forms an H-bond with the glutamate of the α_{inh}. The triphosphate interacts with the anionic nest formed by the $\alpha5$ (Engel et al. 2012) (Goepfert et al. 2013).

The conserved glutamate of α_{inh} is important for the inhibitory effect because truncation of its side chain turns overexpressed Fic proteins of all three classes into toxic proteins in *E.coli* (Engel et al. 2012). Crystal structures of three mutant proteins with bound ATP or AMPPNP show that the base and the ribose moieties have the same location as the wildtype, while the triphosphate reveals a different conformation. The γ-phosphate forms a tight salt bridge with the second arginine (R147 in VbhT) (Figure 21A). The α- and β-phosphate moieties form H-bonds with the backbone amide groups at the $\alpha5$, the β-phosphate forms a salt bridge with the first arginine (R144 in VbhT) and the α-phosphate interacts with the asparagine of the Fic motif (N142) Due to this repositioning of the α-phosphate the new orientation allows in-line attack of a target side chain onto the phosphate leading to AMP transfer onto the target (Engel et al. 2012) (Goepfert et al. 2013).
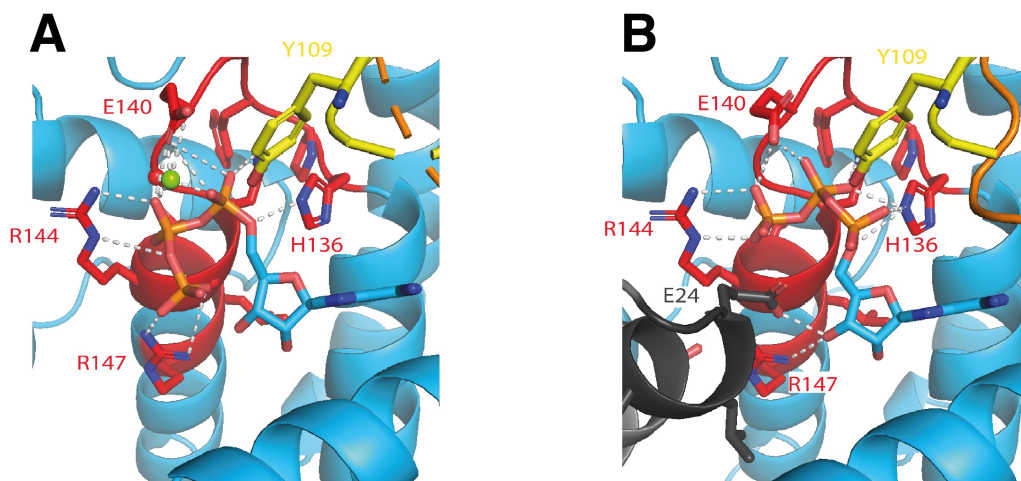
Figure 21: Active site of wildtype VbhT(FIC) (left) or mutated VbhT(FIC)/VbhA$_{E24G}$ (right) with bound ATP (modified from (Goepfert et al. 2013)).

VbhT(FIC) is shown as cartoon in blue with the FIC motif in red and the flap in orange. VbhA is displayed as cartoon in grey with the α_{inh} in dark grey. Selected residues are shown in full. Important interactions of the ATP with the Fic motif or inhibition motif are shown as light grey dashes.

A: The position of the α-phosphate is not suitable for a nucleophilic attack of a target hydroxyl side chain onto the ATP. The metal ion is shown in green.

B: Mutation of the inhibitory glutamate leads to reorientation of the α-phosphate of the substrate and therefore allows in-line attack of the target side chain after proton transfer to the active histidine.

The aspartate or glutamate in the Fic motif display importance in the coordination of the metal ions. In all structures containing an α_{inh} E > G mutation a magnesium ion is found, which interacts with the α- and β-phosphate and is coordinated by the conserved D/E residue in the active site (E140 in VbhT) (Figure 21). The cation is observed only in inhibition-relieved complexes and is crucial for Fic mediated AMPylation due to its finetuning of the ATP phosphates within the anion binding nest (Goepfert et al. 2013).

Studies with the class III Fic protein NmFic revealed an additional mechanism, a combination of autoAMPylation and oligomerization, to regulate AMPylation of target proteins. NmFic forms a tetramer in its wildtype state, which is catalytically not competent due to its buried active site. Monomers of NmFic are autoAMPylated on a tyrosine residue in the α_{inh} in *cis* leading to partial unfolding of the helix resulting in binding of ATP into the active site and AMPylation of target proteins. NmFic reveals a sharp monomer-tetramer transition in the presence of ATP and at low enzyme concentrations. NmFic activity *in vivo* would respond to changes in protein concentrations with a certain delay and there might be self-sufficient molecular timers (Stanger et al. 2016) (Harms et al. 2016).
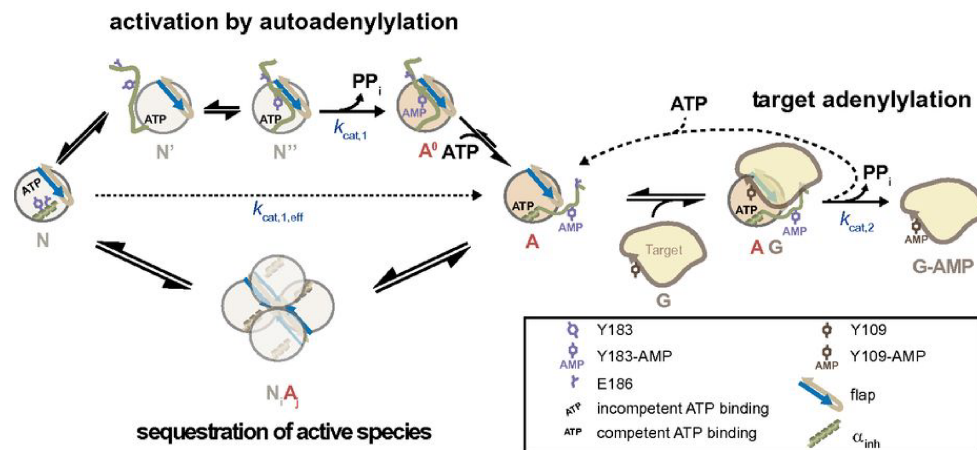
Figure 22: Model of the regulatory mechanism of class III Fic protein NmFic (Stanger et al. 2016).
Monomeric NmFic is activated by autoAMPylation to obtain species A, which involves unfolding of the α_{inh}, binding of the modifiable tyrosine to the unfolded part of the flap and automodification of the tyrosine. Both species are in a dynamic monomer/tetramer equilibrium forming homotetramers or heterotetramers. The remaining monomers are competent for target AMPylation.

1.4.4   Fic proteins as bifunctional enzymes? DeAMPylation mediated by Fic proteins

Several studies have revealed that Fic proteins can act as bifunctional enzymes and remove the transferred modification from their target proteins. One of the first studies about deAMPylation activity by Fic proteins was done on the *Legionella pneumophila* effector SidM or DrrA, which is a multi-domain protein and consists of a nucleotidyl transferase domain responsible for Rab1 AMPylation (Muller et al. 2010). This domain is also found in GS-ATase, which is a bifunctional enzyme catalyzing the addition and removal of AMP from the glutamine synthetase (Stadtman 2001) (Xu et al. 2010). Therefore, deAMPylation was speculated as a function of SidM until SidD was identified as a deAMPylase (Neunuebel et al. 2011) (Tan and Luo 2011). The expression of this protein is temporarily regulated and a catalytic domain, which resembles a metal-dependent phosphatase domain catalyzes deAMPylation causing release of AMP (Neunuebel et al. 2011).

Recent studies discovered the deAMPylating activity of the human Fic protein FICD, which leads to reactivation of the Hsp70 chaperone BiP and its entry into the chaperone cycle (Preissler et al. 2017). The inhibitory glutamate 234 (E234) suppresses BiP AMPylation due to side chain engagement in the active site and turns the enzyme into a deAMPylase (Engel et al. 2012) (Bunney et al. 2014) (Preissler et al. 2017). Mutations of

the residue abolishes the deAMPylation activity and displays it AMPylation activity on BiP (Preissler et al. 2017). The side chain of E234 is flexible and has different orientations as it is revealed in the crystal structures of wild-type FICD or FICD bound to ADP or ATP (Bunney et al. 2014). Due to BiP's role it is speculated that this could represent a mechanism about how the protein deals with the fluctuating levels of unfolded proteins in the ER. In presence of high levels, the E234 side chain is engaged in the active site leading to deAMPylation of BiP and recruiting it back into the chaperone cycle. When the burden of unfolded proteins is low AMPylation takes place due to disengagement of the side chain and causing inactivation of BiP. In FICD-mediated AMPylation deprotonation and attack of the threonine 518 (T518) of BiP on the α-phosphate of ATP leads to modification of the target. In the deAMPylation reaction the hydroxyl of a water molecule, which is supposed to be activated by E234, causes an attack onto the phosphodiester bond of the bound AMP (Figure 23) (Preissler et al. 2017).
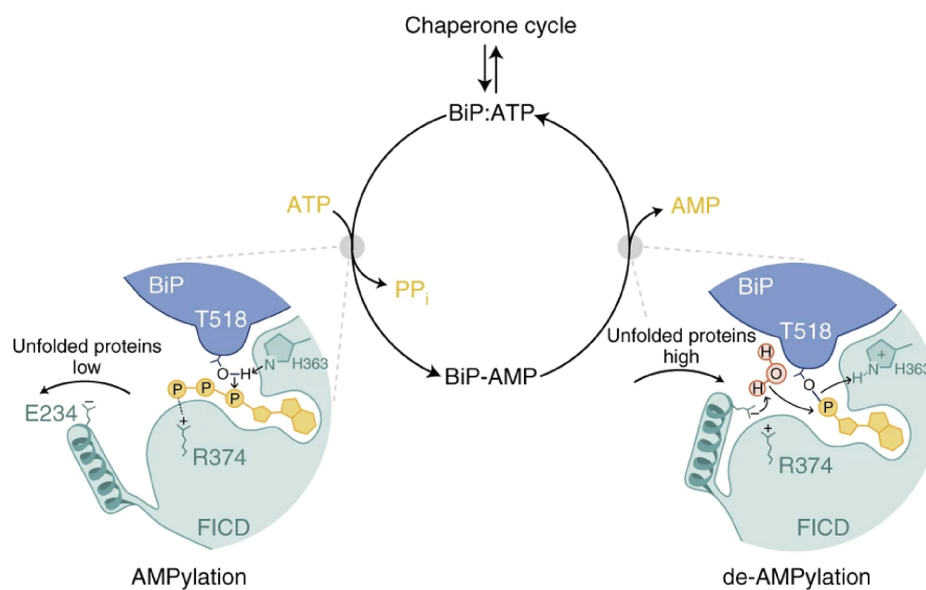
Figure 23: First hypothetical model of FICD-mediated AMPylation and deAMPylation of BiP (Preissler et al. 2017).

When the level of unfolded proteins is low, the α_{inh} with E234 moves out of the active site resulting in an AMPylation-competent state and leading to modification of BiP on threonine 518. If the level of unfolded proteins is high, E234 engages into the active site and coordinates a water molecule which is responsible to attack the phosphodiester bond between the bound AMP and the side chain. The histidine is assumed to protonate the leaving group. DeAMPylation leads to BiP being recruited back to the chaperone cycle.

Both AMPylation and deAMPylation require the conserved histidine, which mediates deprotonation leading to attack of the threonine hydroxyl of BiP on the α-phosphate of ATP. In deAMPylation the hydroxyl of the water molecule performs a hydrophilic attack onto the bound AMP while the histidine is needed for protonation of the BiP T518 leaving group (Preissler et al. 2017).

FICD is a dimer in its wildtype form and restricts the enzyme to deAMPylation activity, while a pool of monomeric FICD show AMPylation. Structural studies suggested that monomerization of the enzyme weakens a network of intramolecular contacts, while in the deAMPylation state the contacts stabilise the E234 leading to unsuccessful binding of MgATP. Dimeric FICD binds ATP in an AMPylation in-competent mode, which is also shown for other Fic proteins (Engel et al. 2012) (Goepfert et al. 2013) (Perera et al. 2019). It was revealed that monomerization of FICD or mutations in residues linking the dimer interface to the active site allows competent binding of ATP in presence of Mg^{2+} despite the inhibitory E234. Induced monomerization of FICD is also connected to increased flexibility of the E234 side chain due to weakening of the dimer relay in solution. Enhanced dimer dissociation is induced by ADP and the increase of the ADP/ATP ratio shifts the monomer-dimer equilibrium in direction of dimerization of FICD (Perera et al. 2019).

In 2019 another Fic protein was revealed to contain bifunctional activity. The class III protein EfFic from *Enterococcus faecalis* (*E. faecalis*) has AMPylating and deAMPylating activities, which are both mediated by the same active site. In absence of a known target purified, AMPylated inhibition-relieved AMP-EfFic$_{E190G}$ was used to study deAMPylation mediated by wildtype EfFic in presence of Ca^{2+}, while mutants of the catalytic histidine (H111), the metal-coordinating glutamate (E115) or the inhibitory glutamate (E190) do not show deAMPylation activity. Mg^{2+} and Ca^{2+} both support deAMPylation of AMP-EfFic$_{E190G}$ by wildtype EfFic but EfFic$_{E190G}$ abolished deAMPylation in presence of Ca^{2+} and partially with Mg^{2+}. Mg^{2+} supports AMPylation which is eliminated in presence of Ca^{2+}. Addition of Ca^{2+} to unchanging Mg^{2+} concentrations switches EfFic from AMPylation to deAMPylation (Veyron et al. 2019) . Ca^{2+} might conflict with Mg^{2+} and cannot interact with the metal-coordinating glutamate, which is known to be critical for modification, and therefore inhibits AMPylation (Engel et al. 2012) (Veyron et al. 2019).

The same study revealed that Ca^{2+} can tune the deAMPylation activity of human FICD. In competition experiments with wildtype FICD and AMPylated BiP in presence

of a fixed Mg^{2+} concentration and increasing Ca^{2+} concentration deAMPylation is decreasing with increasing Ca^{2+}/Mg^{2+} ratio (Veyron et al. 2019). The ER is involved in calcium homeostasis and contains high concentration of free Ca^{2+} under resting conditions (Montero et al. 1995). FICD stimulates the activity of BiP in response to increase of unfolded proteins and therefore it is speculated that inhibition of FICD mediates deAMPylation due to high Ca^{2+} concentrations observed *in vitro* might be also present in ER homeostasis where Ca^{2+} concentration is high and BiP is inactive (Sanyal et al. 2015) (Preissler et al. 2015).

In a recent published study the crystal structure of FICD in complex with AMPylated BiP displays more insight into the coordination of the E234 and its interaction with R1 and R2 of the active side, the cation and the water molecules necessary to perform deAMPylation of BiP. AMP is covalently attached to the threonine 518 of BiP and is coordinated by Mg^{2+}, which is itself held in place by the metal-coordinating aspartate 367 (D367) of the FIC motif. E234 tightly engages a water molecule within the active site, which sits almost directly in-line with the Pα-Oγ(Thr518) phosphodiester bond and is assumed to mediate deAMPylation of the target protein. The positioned catalytic histidine 363 (H363) facilitates protonation of the alkoxide leaving group leading to unmodified BiP and AMP as products (Perera et al. 2021).

2   Aim of my Thesis

FIC proteins mediating post-translational modifications have been of interest over the last decade. Their conserved enzymatic machinery is able to manipulate the activity of various target proteins some of them by the transfer of AMP called AMPylation. Class I FicT toxins, containing a conserved active site, are regulated by their cognate antitoxins FicA, while activity of other Fic proteins is controlled by intramolecular processes.

The aim of my PhD thesis was to gain further insight into the regulatory mechanism of post-translational modifications mediated by class I Fic proteins including the effector proteins of *Bartonella*.

Bartonella's evolutionary success of its stealth infection strategy in some phylogenetic lineages correlates with the acquisition of the VirB/D4 T4SS, which is responsible for secretion of host-targeted effector proteins called Beps into host cells. Beps belong to class I Fic proteins and bind to the FicA homolog named BiaA. In collaboration with the Seattle Structural Genomics Center of Infectious Disease several Beps were crystallized yielding nine crystal structures, two structures in complex with their respective BiaA. One aim of my PhD thesis was further analysis of the solved structures including the comparison of the of the Bep toxin-antitoxin complexes with the already published structures of the VbhT/VbhA and EcFicT/EcFicA.

Recent studies on the human Fic protein FICD revealed deAMPylation of the Hsp70 chaperone BiP, which has significant impact on its role in the chaperone cycle in the endoplasmic reticulum. The second aim of my PhD thesis was to study the AMPylation and deAMPylation of GyrB mediated by VbhT/VbhA using a newly developed nucleotide quantification assay and kinetically characterize both reactions, while analyzing the structural aspects of the deAMPylation reaction.

3 Results

3.1 Research article I

Evolutionary diversification of host-targeted *Bartonella* effectors proteins derived from a conserved FicTA toxin-antitoxin module

Tilman Schirmer[1]*, Tjaart A. P. de Beer[1,4], Stefanie Tamegger[1], Alexander Harms[1], Nikolaus Dietz[1], David Dranow[2], Thomas E. Edwards[2], Peter E. Myler[2], Isabelle Phan[2], and Christoph Dehio[1,*]

[1]Biozentrum, University of Basel, Basel, Switzerland; [2]Seattle Structural Genomics Center for Infectious Disease, Seattle, Washington, USA; [3]Present address: Roche, Basel, Switzerland; [4]Present address: Inchnos Sciences Biotherapeutics SA, Epalinges, Switzerland

3.1.1 <u>Statement of my own contributions</u>

My contribution to research article I was analyzing the interactions between toxin and antitoxin of VbhT/VbhA, EcFicT/EcFicA, *Bro*_Bep1/BiaA and *Bhe*_BepA/BiaA, and visualization of crystal structures displayed in the manuscripts.

3.1.2 <u>"Evolutionary diversification of host-targeted *Bartonella* effectors proteins derived from a conserved FicTA toxin-antitoxin module"</u>

microorganisms

Article

Evolutionary Diversification of Host-Targeted *Bartonella* Effectors Proteins Derived from a Conserved FicTA Toxin-Antitoxin Module

Tilman Schirmer [1,*], **Tjaart A. P. de Beer** [1], **Stefanie Tamegger** [1], **Alexander Harms** [1], **Nikolaus Dietz** [1], **David M. Dranow** [2,3], **Thomas E. Edwards** [2,3], **Peter J. Myler** [2,4,5], **Isabelle Phan** [2,4] and **Christoph Dehio** [1,*]

1   Biozentrum, University of Basel, 4056 Basel, Switzerland; tdebeer@gmail.com (T.A.P.d.B.);
    stefanie.tamegger@unibas.ch (S.T.); alexander.harms@unibas.ch (A.H.); nikolaus.dietz@unibas.ch (N.D.)
2   Seattle Structural Genomics Center for Infectious Disease, Seattle, WA 98109, USA;
    DDranow@be4.com (D.M.D.); tom.edwards@ucb.com (T.E.E.); peter.myler@seattlechildrens.org (P.J.M.);
    Isabelle.Phan@seattlechildrens.org (I.P.)
3   Beryllium Discovery, Bainbridge Island, WA 98110, USA
4   Seattle Children's Research, Seattle, WA 98109, USA
5   Departments of Pediatrics, Global Health and Biomedical Informatics & Health Education,
    University of Washington, Seattle, WA 98109, USA
*   Correspondence: tilman.schirmer@unibas.ch (T.S.); christoph.dehio@unibas.ch (C.D.)

Abstract: Proteins containing a FIC domain catalyze AMPylation and other post-translational modi-
fications (PTMs). In bacteria, they are typically part of FicTA toxin-antitoxin modules that control
conserved biochemical processes such as topoisomerase activity, but they have also repeatedly diver-
sified into host-targeted virulence factors. Among these, *Bartonella* effector proteins (Beps) comprise a
particularly diverse ensemble of FIC domains that subvert various host cellular functions. However,
no comprehensive comparative analysis has been performed to infer molecular mechanisms underly-
ing the biochemical and functional diversification of FIC domains in the vast Bep family. Here, we
used X-ray crystallography, structural modelling, and phylogenetic analyses to unravel the expansion
and diversification of Bep repertoires that evolved in parallel in three *Bartonella* lineages from a single
ancestral FicTA toxin-antitoxin module. Our analysis is based on 99 non-redundant Bep sequences
and nine crystal structures. Inferred from the conservation of the FIC signature motif that comprises
the catalytic histidine and residues involved in substrate binding, about half of them represent AMP
transferases. A quarter of Beps show a glutamate in a strategic position in the putative substrate
binding pocket that would interfere with triphosphate-nucleotide binding but may allow binding of
an AMPylated target for deAMPylation or another substrate to catalyze a distinct PTM. The β-hairpin
flap that registers the modifiable target segment to the active site exhibits remarkable structural
variability. The corresponding sequences form few well-defined groups that may recognize distinct
target proteins. The binding of Beps to promiscuous FicA antitoxins is well conserved, indicating a
role of the antitoxin to inhibit enzymatic activity or to serve as a chaperone for the FIC domain before
translocation of the Bep into host cells. Taken together, our analysis indicates a remarkable functional
plasticity of Beps that is mostly brought about by structural changes in the substrate pocket and the
target dock. These findings may guide future structure–function analyses of the highly versatile FIC
domains.

1. Introduction

Fic proteins, which are characterized by containing a FIC (filamentation induced by
cAMP) domain, form a diverse protein family and are found in all domains of life [1,2].

They are enzymes that mediate AMPylation (also known as adenylylation) and other post-translational modifications (PTMs) of proteins [1,2]. The FIC fold comprises eight conserved α-helices and encompasses a characteristic active site loop between helices α4 and α5 that forms most of the catalytic center and the substrate binding pocket (reviewed by [3–5]). Despite a remarkable diversity of target proteins and substrates, Fic proteins typically catalyze the nucleophilic attack of a target hydroxyl residue onto the diphosphate moiety of a nucleotide substrate, which causes the transfer of a phosphoryl-linked moiety onto the target protein. Most of them display a canonical (HPFx(D/E)GNGRxxR) signature motif that locates the active site loop and is critical for AMPylation activity [2,6–8]. However, several families of Fic proteins carry non-canonical FIC signature motifs, with some of them shown to mediate other PTMs such as phosphocholination or phosphorylation [9,10]. Recent reports showed for some Fic proteins that they are also able to deAMPylate AMPylated targets [11–13].

The enzymatic activity of AMPylating FIC domains is typically tightly controlled by active site obstruction via an inhibitory α-helix (α_{inh}) that prevents productive ATP binding [1] and may be, in fact, crucial for de-AMPylation [11,12]. α_{inh} is found N-terminally or C-terminally at the FIC core in class II and III Fic proteins, respectively, or is part of a small interacting protein known as antitoxin in the case of class I Fic proteins [1]. Additionally, FIC domains generally contain a β-hairpin between helix α2 and helix α3 that is referred to as the "flap" and mediates docking of a target segment in extended conformation via β-strand augmentation [5,14–16]. This sequence-independent interaction ensures the productive insertion of the modifiable hydroxyl residue of the target into the FIC active site. On top of the catalytic core machinery, FIC domains typically contain one or more accessory extensions within or around this core that contribute to target and/or substrate specificity [3].

So far, most research on Fic proteins have focused on host-targeted virulence factors of various bacterial pathogens such as VopS of *Vibrio parahaemolyticus* and IbpA of *Histophilus somni* that inactivate Rho family GTPases by AMPylation to cause collapse of the actin cytoskeleton and host cell death [7,8]. Similarly, the type IV secretion (T4S) system effector AnkX of *Legionella pneumophila* manipulates membrane trafficking in host cells by the phosphocholination of Rab1 and Rab35 [17] and AvrAC of the plant pathogen *Xanthomonas campestris* inactivates two immune kinases by UMPylation [18]. Although the example of these proteins highlights the diversity of target proteins and PTMs of FIC domains, they only have secondarily evolved out of a much more abundant pool of Fic proteins that act in a genuine bacterial context [1]. Among these, we have previously shown that class I Fic proteins and their small inhibitory partner encompassing α_{inh} constitute FicTA toxin-antitoxin modules exemplified by VbhTA of *Bartonella schoenbuchensis* and Ye_FicTA of *Y. enterocolitica* [19]. When released from their FicA antitoxin, these FicT toxins cause bacterial growth inhibition by AMPylation and concomitant inactivation of type IIA topoisomerases (i.e., DNA gyrase and topoisomerase IV), causing a disruption of cellular DNA topology [19].

The α-proteobacterial genus *Bartonella* comprises numerous host-restricted species that share a conserved stealth infection strategy to cause long-lasting hemotropic infections in their respective reservoir hosts [20]. Within the genus, the amazing host adaptability and concomitant virtual ubiquity of two particular phylogenetic lineages (lineage 3 and lineage 4) have been linked to the acquisition of the VirB/D4 T4S system and the vast potential of its secreted effectors called *Bartonella* effector proteins (Beps) in order to manipulate host cell functions [21,22]. Interestingly, it appears that the VirB/D4 T4S system of lineage 3 (L3) and lineage 4 (L4) has been acquired independently with a single ancestral effector from a common source, followed by parallel series of duplication and functional diversification of the effector genes in both lineages [21]. Since adaptive radiations of L3 and L4 driven by biodiversification of these host-restricted bacteria to infect a wide range of mammals had only been triggered after the evolution of complex effector sets, it seems clear that the functional diversity of the effectors was a key innovation promoting host

adaptability in the frame of the conserved *Bartonella* stealth infection strategy [21,23]. Of note, we have recently described a third acquisition of the VirB/D4 T4S system in the L1 species *B. ancashensis* that also resulted in the evolution of a complex Bep repertoire. However, this event has not yet resulted in radiation, which indicates a more primordial stage of evolution compared to L3 and L4 [24]. Although novel types of effectors without FIC domains arose in all three linages, a prominent fraction of prototypic Beps in L1 (12 out of 21), L3 (9 out of 12), and L4 (5 out of 10) contain FIC domains that are generally assumed to engage with target proteins in the host, although the molecular activity and biological role of these BepFIC domains have only begun to be revealed [21]. So far, it has been shown that Bep1 AMPylates Rac-subfamily GTPases [16], Bep2 AMPylates the host intermediate filament protein vimentin [25], and BepA causes the AMPylation of two unknown host proteins of ca. 40–50 kDa size [6]. Furthermore, all of these proteins display auto-modification, which is a common feature of FIC domains [2,6,25]. For the FIC domain of BepC with its non-canonical FIC signature motif, no enzymatic activity has been found although it binds to the RhoGEF GEF-H1 and activates this signaling protein by relocalization to the plasma membrane [26,27]. Apart from the FIC domain, all effectors contain at least one BID (*Bartonella* Intracellular Delivery) domain and a positively-charged C-terminus that forms a conserved secretion signal for translocation by the VirB/D4 T4S system [28]. Furthermore, the effectors harbor a conserved OB-fold between FIC and BID domains that was shown for other proteins to mediate interaction with oligonucleotides or oligosaccharides; however, its function in the context of the Beps remains unknown [2,6].

Taken together, in recent years we have learned a substantial amount about structure–function aspects of Fic proteins, including selected Beps, but no comprehensive comparative analysis had been performed so far to study possible molecular mechanisms underlying the biochemical and functional diversification of FIC domains in the vast Bep family. Here, we show that Beps represent genuine class I Fic proteins that can form a tight complex with the promiscuous BiaA antitoxin in *Bartonella*. Our comprehensive analysis of the structural evolution of BepFIC domains using X-ray crystallography, structural modeling, and sequence comparison shed light on the remarkable functional and regulatory plasticity of Beps brought about by minor structural changes of the FIC fold. Our results underline the diversification and specialization of *Bartonella* effectors regarding catalytic activities and target proteins and provide a paradigm for the diversification of a highly conserved enzymatic scaffold in the course of adaptive evolution.

2. Material and Methods

2.1. Bioinformatics

NCBI BLAST [29] was used to search against the nr database by only filtering out *Bartonella* sequences. The *B. rochalimae* Bep1 protein sequence was used as a query and only sequences containing both the FIC and BID domain were retained. An E-value cutoff of $1e^{-3}$ was used, which gave rise to 146 sequences from various *Bartonella* species as listed in Table S1. Redundancy among these sequences was reduced by elimination of closely related homologues and resulted in 99 sequences listed in Table S2. Full-length FIC-BID Beps sequences were annotated and aligned (CLUSTALW [30] routine with default parameters) and sequence logos were generated within Geneious Prime 2020.2.3 (www.geneious.com accessed on 10 March 2021). Protein structures were visualized with PyMOL (pymol.org accessed on 10 March 2021) and Dino (dino3d.org accessed on 10 March 2021).

The subgrouping of sequences according to active site sequence features and the computation of correlations were performed with an in-house Python (python.org accessed on 10 March 2021) routine. The correlation of the occurrence of specific residue types at two specific positions (co-conservation) was computed with another Python script. In short, for a given position in the multiple sequence alignment, the presence/absence of the respective residue type was coded with a binary of 1 or 0, respectively. This allowed the computation of the Pearson correlation between vectors representing the two specified positions. A cladogram was generated by running the Simple Phylogeny routine (European Bioin-

formatics Institute server (https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/ (accessed on 10 March 2021) on the flap segments extracted from the BepFIC multisequence alignment.

2.2. Cloning

B. clarridgeiae strain CIP 104772/73 full-length Bep1 (*Bcl*_Bep1, UniProtKB: E6YFW2, aa 1-558) and the FIC domains of *B. clarridgeiae* strain CIP 104772/73 Bep5 (*Bcl*_Bep5, E6YGF5, aa 14-226), *B. sp.* strain AR 15-3 Bep8 (*B15*_Bep8, E6YQQ1, aa 9-240), *B. sp.* strain AR 1-1C Bep8 (*B11C*_Bep8, E6YV77, aa 10-241), *B. tribocorum* strain CIP 105476/IBS 506 BepC (*Btr*_BepC, A9IWP7, aa 3-220), and *B. quintana* strain Toulouse BepC (*Bqu*_BepC, Q6FYV8, aa 3-220) were each cloned into the uncleavable pBG1861 vector [31] by LIC cloning [32] in order to produce constructs with an N-terminal hexahistidine tag.

The full-length biaA gene that codes for the small ORF directly upstream of bepA gene and part of the bepA gene from B. henselae (Bhe_bepA) encoding the FIC and OB domains (amino acid residues 1-296) were PCR-amplified from genomic DNA. The PCR products for biaA from B. henselae (Bhe_biaA) and the fragment of Bhe_bepA were cloned into the pRSF-Duet1 vector using NcoI/BamHI and NdeI/XhoI restriction sites, respectively. The pRSF-Duet1 vector containing Bhe_biaA and the Bhe_bepA constructs was transformed into *E. coli* BL21-AI (Invitrogen, Waltham, MA, USA). The constructs were expressed and purified as described in for VbhA/VbhT(FIC) [1] and concentrated in a crystallization buffer (20 mM HEPES, 150 mM NaCl, 2 mM MgCl$_2$, 1 mM TCEP) to 20 mg mL^{-1} for crystallization.

2.3. Expression and Protein Purification

Each protein was expressed in BL21 (DE3) *E. coli* cells in 2l of auto-induction media in a LEX bioreactor (Harbinger, Markham, ON, Canada) at 20 °C for 72 h, after which the harvested cells were flash frozen in liquid nitrogen.

*Bcl*_Bep1 and *B15*_Bep8 FIC were purified as described previously [33]. The following protocol was used for all other proteins: The frozen cell pellet was thawed and re-suspended in 20 mM HEPES, pH 7.4, 300 mM NaCl, 5% glycerol, 30 mM Imidazole, 0.5% CHAPS, 10 mM MgCl$_2$, 3 mM β-mercaptoethanol, 1.3 mg/mL protease inhibitor cocktail (Roche, Basel, Switzerland) and 0.05 mg/mL lysozyme. The collected cells were sonicated for 15 min and incubated with 20 μL Benzonase® nuclease (EMD Chemicals, Gibbstown, NJ, USA) for 40 min at room temperature. The lysate was centrifuged and filtered through a 0.45 μm cellulose acetate filter (Corning Life Sciences, Lowell, MA, USA). Recombinant protein was purified by affinity chromatography by using a HisTrap FF 5 mL column (GE Biosciences, Piscataway, NJ, USA) equilibrated in 25 mM HEPES, pH 7.0, 300 mM NaCl, 5% Glycerol, 30 mM Imidazole, 1 mM DTT buffer and eluted with 500 mM imidazole in the same buffer. The concentrated sample was further purified by size exclusion chromatography in 20 mM HEPES, pH 7.0, 300 mM NaCl, 5% glycerol and 1 mM TCEP.

Fractions with pure protein (greater than 90% pure according to Coomassie-stained SDS-polyacrylamide gels) were pooled and concentrated, yielding 4.9 mg/mL for *Bqu*_BepC (expected MW = 26.47 kDa, observed MW = 26 kDa), 22.75 mg/mL of *Btr*_BepC (expected MW = 26.18 kDa, observed MW = 26 kDa), 26.4 mg/mL of *Bcl*_Bep1 (expected MW = 63.81 kDa, observed MW = 65 kDa), 45.3 mg/mL of *Bcl*_Bep5 (expected MW = 26.03 kDa, observed MW = 25 kDa), 42.3 mg/mL of B15_Bep8 (expected MW = 28.17 kDa, observed MW = 26 kDa), and 22.00 mg/mL B11_Bep8 (expected MW = 28.17 kDa, observed MW = 28 kDa. Aliquots of 100–200 μL pure protein samples were flash frozen in liquid nitrogen and stored at −80 °C.

2.4. Crystallography

Crystals for *Bqu*_BepC and *Btr*_BepC, *Bcl*_Bep1, *Bcl*_Bep5, *B15*_Bep8, and *B11C*_Bep8 were grown at 289 K by sitting drop vapor diffusion by mixing 0.4 μL of protein solution

with 0.4 μL reservoir solution. *Bqu*_BepC FIC crystals were grown from a solution of 10% *w/v* PEG 4000, 20% *v/v* glycerol, 0.1 M bicine/Trizma base, pH 8.5, and 0.03 M each of NaF, NaBr, and NaI. *Bqu*_BepC FIC with ADP crystals were grown from a solution of 15%'*w/v* PEG 3350 and 0.1 M sussinic acid. *Btr*_BepC FIC with AMPPNP crystals were grown from a solution of 10% *w/v* PEG 8000, 20% *v/v* ethylene glycol, 0.1 M MES/imdazole, pH 6.5, and 0.02 M each of sodium L-glutamate, DL-alanine, glycine, DL-lysine, and DL-serine. *Bcl*_Bep1 FIC were grown from a solution of 10% *w/v* PEG 20,000, 20% *v/v* PEG MME 5500, 0.1 M MOPS/HEPES-Na, pH 7.5, and 0.02 M each of sodium formate, ammonium acetate, trisodium citrate, sodium potassium L-tartrate, and sodium oxamate. *Bcl*_Bep5 FIC crystals were grown from a solution of 20% *w/v* PEG 3350, 0.1 M sodium citrate/citric acid, pH 4.0, and 0.2 M sodium citrate tribasic. *B11C*_Bep8 FIC crystals were grown from a solution of 30% *v/v* Jeffamine M-600 and 100 mM HEPES free acid/NaOH, pH 7.0. *B15*_Bep8 Fic crystals were grown from a solution of 20% *w/v* PEG 8000, 0.1 M HEPES/NaOH, pH 7.5.

Data for *Bqu*_BepC FIC, *B15*_Bep8 FIC, and *Bcl*_Bep5 FIC were collected at 100 K on a Rayonix MX-300 detector at a wavelength of 0.9786 Å on beamline 21-ID-G at the Advanced Photon Source (APS, Argonne, IL, USA). Data for *Bqu*_BepC FIC with ADP, *Bcl*_Bep1 FIC, and *B11C*_Bep8 FIC were collected at 100 K on a Rayonix MX-225 detector at a wavelength of 0.9786 Å on beamline 21-ID-F at the Advanced Photon Source (APS, Argonne, IL, USA). Data for *Btr*_BepC FIC with AMPPNP were collected on a Rigaku Saturn 944 + detector at a wavelength of 1.5418 Å with our in-house source (Rigaku FR-E + Superbright rotating anode). For all datasets, indexing and integration were carried out by using XDS and the scaling of the intensity data was accomplished with XSCALE [34]. For all structures except *Bqu*_BepC with ADP, the structure was solved by using a molecular replacement with Phaser [35]. For *Bqu*_BepC FIC and *Bcl*_Bep1 FIC, the starting model was the *Bhe*_BepA FIC (PDB: 2JK8). For *BAR*_Bep8 FIC, *Bqu*_BepC FIC with ADP, and *Btr*_BepC FIC, the starting model was *Bqu*_BepC FIC (PDB: 4LU4). For *B11C*_Bep8 FIC, the starting model was *Bcl*_Bep1 FIC (PDB: 4NPS). For *Bcl*_Bep5 FIC, the starting model was *Bhe*_BepA FIC (PDB: 2VZA). For all structures, refinement model building was carried out by using either Refmac5 [36] or Phenix [37], TLS [38], and Coot [39]. All structures were quality checked by Molprobity [40].

The crystals for Bhe_BepA/BiaA were obtained at 22 °C by using the hanging-drop vapor diffusion method upon mixing 0.2 μL protein solution with 0.2 μL reservoir solution. The reservoir solution was composed of 0.2 M di-Sodium malonate 20% *w/v* Polyethylene glycol 3350. For data collection, crystal was frozen in liquid nitrogen with 20% *v/v* glycerol as cryoprotectant. Diffraction data were collected on beam-line X06SA (PXIII) of the Swiss Light Source (λ = 1.0 Å) at 100 K on a Pilatus 3M detector. Data were processed with XDS and the structure was solved by molecular replacement with Phaser by using the Bhe_BepA structure (PDB: 2JK8) as search model. Several rounds of iterative model building and refinement were performed by using Coot, Phenix, and Buster [41], respectively.

3. Results

3.1. Comparative Sequence Analysis of BepFIC Domains

Class I Fic toxins and their cognate antitoxins form the FicTA toxin-antitoxin module that is present in a large number of bacterial species belonging to diverse phyla [19]. A phylogenetic tree based on an alignment of the FIC domains of FicT toxins reveals multiple deep-branching clades (Figure 1). The FIC domains of *Bartonella sp.* (BepFICs) form a monophyletic cluster emerging from a deep-branching clade of rhizobial FicT toxins for which they display considerable sequence similarity. This finding indicates that they are all derived from a common ancestor. Consistent with the phylogenetic tree derived from genome-wide analysis [21,24], the BepFIC domains of *Bartonella* lineages three (Bep1–Bep8, and Bep10) and lineages four (BepA–BepC, BepI, and BepJ) and *B. ancashensis* of lineage one form separate sub-clades.
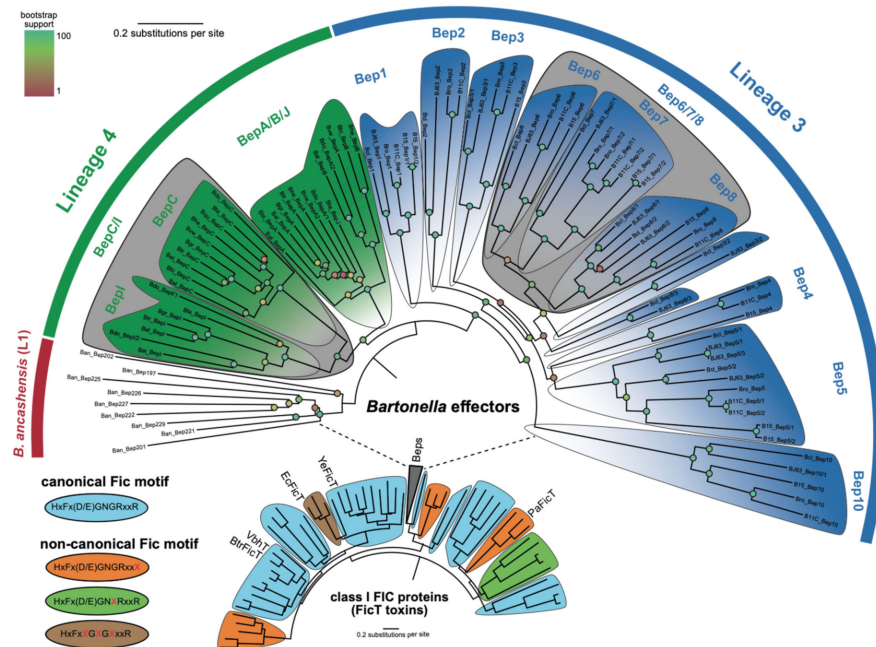
Figure 1. Phylogeny of class I Fic proteins. An extension of the class I Fic phylogeny published in [19], which shows the extended Bep phylogeny. The *Bartonella* Beps form part of a rhizobial clade and have ca. 35% sequence identity with them. Each clade is colored by the representative FIC motif as some Bep clades show a lot of diversity. The BepA/B clade groups together but shows a progressive change between the canonical and non-canonical FIC motif.

The general organization of FIC domain containing Bartonella effector proteins (called FIC-BID Beps) is shown in Figure 2A. The canonical FIC domain is found N-terminally extended by a well-conserved segment of irregular structure (N-ext). C-terminally, the BepFIC domain is followed by a small oligonucleotide/oligosaccharide-binding (OB)-like fold of unknown function and by a BID domain, which, together with the positively charged C-terminus, are responsible for T4S system-mediated translocation [24,28,42,43].

The BepFIC topology and a representative crystal structure (Bhe_BepA) are shown in Figures 2B and 2C. As in other catalytically active Fic proteins (VbhTA, NmFic, etc. [1,44]) and shown for Bhe_BepC further down, the ATP substrate binds to a crevice formed by helices α1, α2, α4, and α6 with the α-phosphate moiety interacting with the N-terminus of α5. Suspended above the bound ATP substrate, there is an irregular β-hairpin (flap, orange) that has been shown to interact with target proteins (e.g., the switch-1 loop of small GTPases [14,16]) through β-sheet augmentation. Another β-hairpin (Bep element, shown in green) found exclusively in BepFICs precedes the flap. Finally, the FIC signature motif HxFx(D/E)GNGRxxR (shown in red) locates to the α4-α5 loop and the N-terminal end of α5.

The large set of non-redundant BepFIC sequences constituting the monophyletic Bep cluster shown in Figure 1 (99 sequences as listed in Table S2) constitutes a valuable basis for the identification of functionally important residues and the classification of potential sub-groups. In the following section, BepFIC sequence conservation as represented by the sequence logo in Figure 3, is discussed in light of the known structural and functional roles of specific residues. The mean pairwise sequence identity of the BepFIC part is 46%, which is somewhat larger than the overall sequence identity of 37% calculated for the full-length FIC-BID sequences (for the full-length sequence logo, see Figure S1).
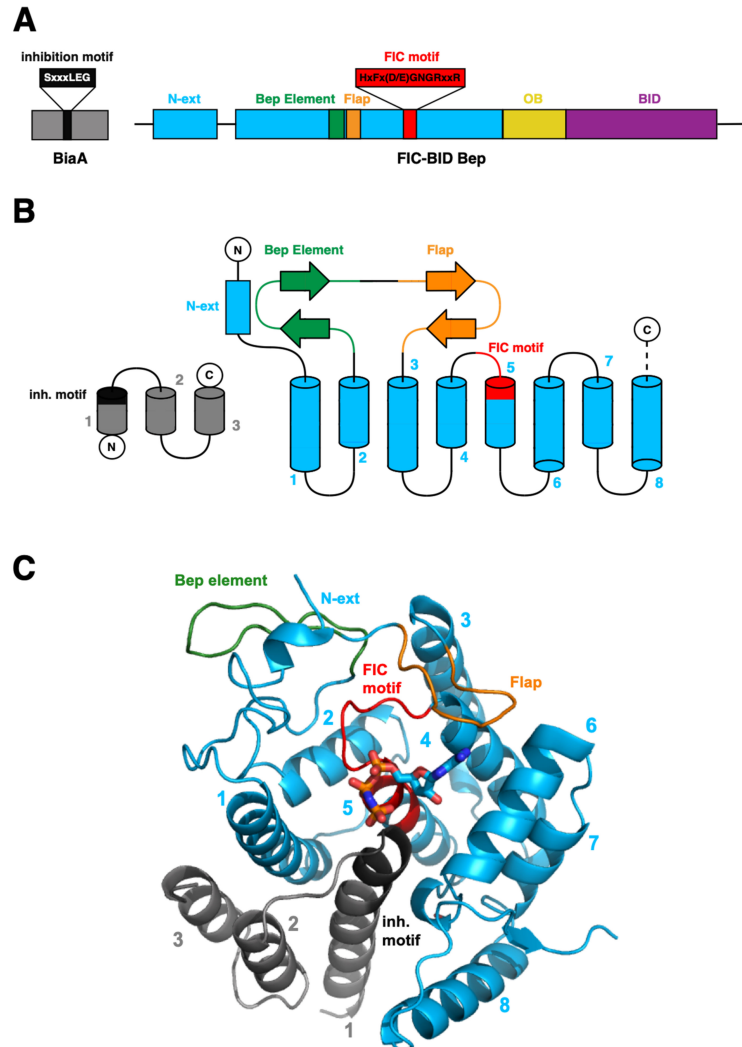
Figure 2. Structural organization of the BepFIC-domain and the interacting BiaA antitoxin. (**A**) Domain organization of FIC-BID proteins characterized by the FIC-fold with an N-terminal extension (N-ext) shown in blue, followed by an OB-fold domain (yellow) of unknown function and the BID-domain (magenta), which, together with the positively charged C-terminus, are responsible for T4S system-mediated secretion. (**B**) Topology of the BepFIC fold consisting of N-ext (blue), helices α1 to α8 (blue), and two β-hairpins (Bep element in green; flap in orange) between α2 and α3. The signature motif locates to the α4-α5 loop and the N-terminal end of α5 (both shown in red) that form the major part of the active site. The topology of the separate BiaA antitoxin is shown in grey. (**C**) Cartoon representation of the BepFIC domain of *Bhe*_BepA in complex with antitoxin *Bhe*_BiaA (PDB ID: 5NH2). The ATP substrate analog AMPPNP shown in sticks has been modeled based on the *Btr*_BepC/AMPPNP complex structure (4WGJ).
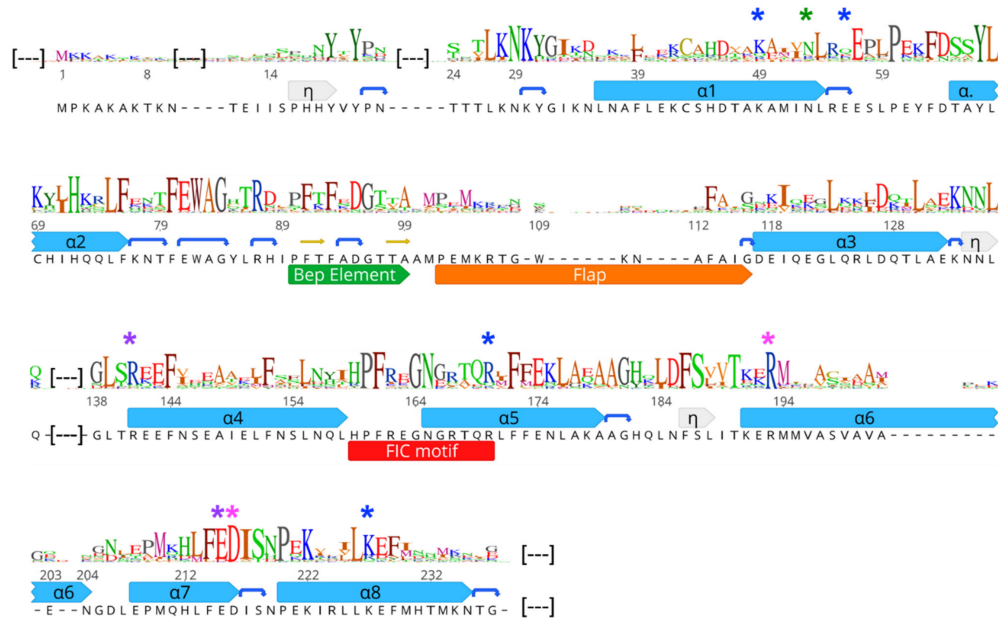
Figure 3. Sequence logo of BepFIC domains derived from the 99 FIC-BID Beps included in the phylogeny given in Figure 1. Sequence, numbering, and secondary structure (helices α1 to α8; 3₁₀-helices; labeled η; in grey; strands in yellow; β-turns shown as blue arcs) correspond to Bhe_BepA (PDB: 5NH2). Partners of two salt-bridges are indicated by magenta and pink asterisks, respectively. Residues forming salt-bridges and the asparagine forming the Asn–Asn interaction with the antitoxin Bhe_BiaA are shown as blue and green asterisks, respectively. Functionally important segments are annotated below the sequence. Large gaps in the alignment of Bhe_BepA, due to non-representative insertions in other sequences, have been cropped for convenience and are marked by (—). The logo covering the entire length of FIC-BID Beps proteins is provided in Figure S1.

　　　　Conserved residues and segments are distributed, though not evenly, across the entire BepFIC sequence (Figure 3). In the following, the roles of these residues are discussed in relation to their position in the Bhe_BepA structure. The canonical FIC helices show only few well conserved residues and most of them are apolar and contributes to the hydrophobic core (e.g., residues L54, L75, F76, I119, L130, F144, F172, L174, L213, F214, and I217; Bhe_BepA residue numbering). The strictly conserved H72 of α2 is completely buried and the two imidazole nitrogens form H-bonds with the FIC signature loop and the loop following α2, thereby effectively tethering these loops together. There are two largely conserved inter-helix salt-bridges, R141-E215 and R193-D216, that join α4 with α7 and α6 with α7, respectively. The conservation of the aforementioned residues strongly suggests the conservation of the FIC-fold for all investigated Beps, which was confirmed by crystal structure analyses (see below).

　　　　The N-terminal extension preceding α1 shows a strongly conserved YxYPxxxxLKNKxGI motif with well-defined structure [6] (Figure 4). Noteworthy, the extension is also present in *Ec*FicT and VbhT (Figure S2). Residues Y19 and L27 of the N-terminal extension together with F80 and W82 of the FEWAG motif of the loop following α2 form a tightly packed hydrophobic core.
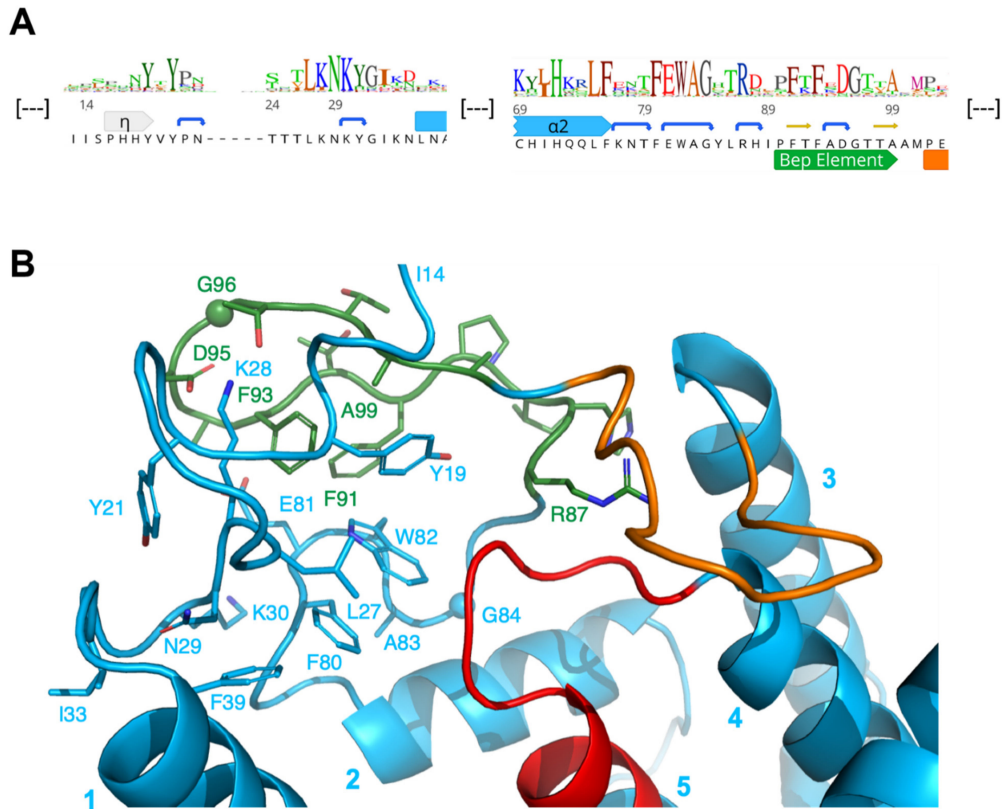
Figure 4. Sequence and structure of the N-terminal that is part of *Bhe*_BepA. (**A**) Excerpt of the N-terminal part of the sequence logo shown in Figure 3 with Bhe_BepA sequence below. (**B**) Detailed view of the N-terminal that is part of the Bhe_BepA structure (5NH2) showing the close interactions between the segments preceding α1 (residues 2–33), the FEWAG segment following α2 (80–84), and the Bep element (90–99, green). Various conserved residues are labeled.

Following the FEWAG motif, there is a conserved insertion of 11 residues relative to *Ec*FicT and VbhT (Figure S6). This Bep specific segment (Bep element) is folded to a β-hairpin (Figures 2C and 5B). Inward facing residues (F91, F93, and A99) and the tip (D95 and G96) are conserved; this suggests that it adopts the same relative position with respect to the domain core in all Beps. The Bep element is also present in the small group of rhizobial FicT toxins that is most closely related to the BepFIC domains, e.g., the FicT homologs of *Ochrobactrum anthropi* (UniProt: A6X7M7) and *Agrobacterium vitis* (UniProt: B9K658).

The subsequent "flap" involved in target registration is folded to an irregular and wide β-hairpin and hovers over the substrate binding site above the signature loop (Figure 2). The sequence of the flap shows surprising diversity both in residue composition and length. This will be discussed in further detail.
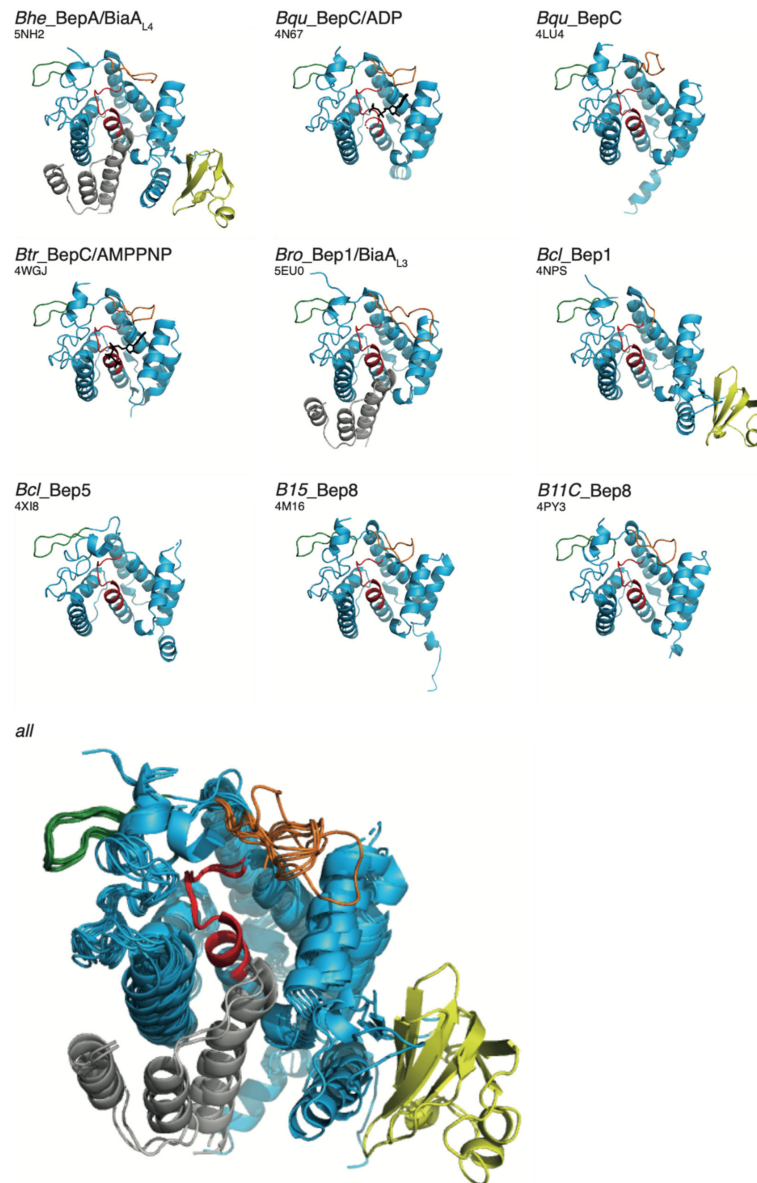
Figure 5. Crystal structures of BepFIC domains from various Bep sub-clades. The proteins are named according to the code provided in Table S2 and PDB codes are provided in parentheses. The structures are represented as in Figure 2. The BepFic structures of *Bhe*_BepA and *Bro*_Bep1 have been determined in complex with their cognate anti-toxins (grey) BiaA$_{L4}$ and BiaA$_{L3}$, respectively. The structures of *Bhe*_BepA and *Bcl*_Bep1 encompass the BepFIC and the OB-fold in yellow (see also [6]). A superposition of all structures onto *Bhe*_BepA (using α4, Fic loop, and α5) is shown at the bottom. All structures with the exception of 5EU0 [16] have been determined in this study and crystallographic details are provided in Table S3.

The sequence of the α3-α4 loop (Figure 3), which is located at the surface of the domain opposite to the flap, is well conserved (with the exception of Bep10, the sequence of which differs drastically from the other Beps for this loop and for α4). Most of the conserved residues (N135, L136, G138, and L139) play a structural role. The Fic signature motif shows a considerable degree of variation, which will be discussed further below. The α5-α6 loop is rather large and its sequence is mostly conserved for structural reasons. The loop allows the perpendicular arrangement of the antiparallel α6/α7 helix pair with α4 (Figure 2).

3.2. Comparative Structure Analysis of BepFIC Domains

As the basis for a detailed structure–function relationship study, Bep constructs were subjected to crystallization and resulted in eight new crystal structures of representative FICBeps with resolutions ranging between 3 and 1.6 Å (Table S3). Figure 5 shows a side-by-side view and a superposition of these crystal structures and the recently determined Bro_Bep1 structure [16]. The comparison shows that the conformation and relative position/orientation of N-ext, the Bep element (green), and the FIC signature loop (red) are invariant. This contrasts with the large variability of the flap (orange), which may relate to the adaptation to various targets (see further down).

Most of the Bep crystallization constructs encompassed only the BepFIC domain, with the exception of Bro_Bep1 and Bhe_BepA, which included the adjoining C-terminal OB-fold. In these cases, the relative position/orientation of the OB-fold with respect to the FIC domain is similar but not identical. Noteworthily, in the constructs that do not have the OB-fold included, the C-terminal α8 is not fully folded (in 4WGJ, it is fully disordered) and show various orientations. Most likely, the proper folding of α8 requires the presence of the OB-fold.

3.3. Variability of FIC Signature Motif and Substrate Binding Site

Most FIC proteins hitherto studied catalyze target AMPylation by using ATP as substrates. These proteins show a typical HPFx(D/E)GNGRxxR sequence (FIC signature motif) for the α4-α5 linker and the adjoining start of α5. This allows the α-phosphate of the substrate to be accommodated in an anion-nest at the N-terminal end of α5 and allows the β-phosphate and γ-phosphate to be coordinated with the two arginyl residues of the motif (Figure 6A, see also references [1,14]). Latter interaction explains the second G (G2) in the motif since a side-chain in this position would interfere with binding. The histidyl side-chain is thought to act as a catalytic base to deprotonate the incoming hydroxyl side-chain of the target protein. Finally, the proline probably rigidifies the particular FIC-loop conformation, the phenylalanine anchors the motif to the hydrophobic core, the D/E side-chain takes part in Mg^{2+} coordination, the first G (G1) is required for steric reasons, and the asparagine side-chain stabilizes the loop by H-bonding back to the main-chain of the phenylalanine (see Figure 3A in [6]).

There are two more residues that are non-contiguous with the FIC signature motif but contribute to nucleotide binding in *Bqu*_BepC (this study; 4N67, 4WGJ) and VbhT ([1]; 3ZCB). These are (1) a phenylalanine from the end of the flap (*Bhe*_BepA residue number 113, also observed in IbpA and NmFic) that interacts perpendicularly with the adenine base of the substrate (and is probably also important for the fixation of the catalytic histidine) and (2) a serine/cysteine side-chain (*Bhe*_BepA number 198) that forms an H-bond with the ribose moiety of the nucleotide (Figure 6A). The overall sequence logo (Figure 6B (top)) for the non-contiguous sequence alignment (Figure S3) shows a strictly conserved PFxxGN motif, which suggests an invariant loop conformation for all of the BepFICs. However, for about half of the sequences, deviations from the strict FIC signature motif are found in the other positions suggesting that these proteins have acquired other function than catalyzing AMP transfer. This prompted us for a comprehensive and comparative sequence analysis.
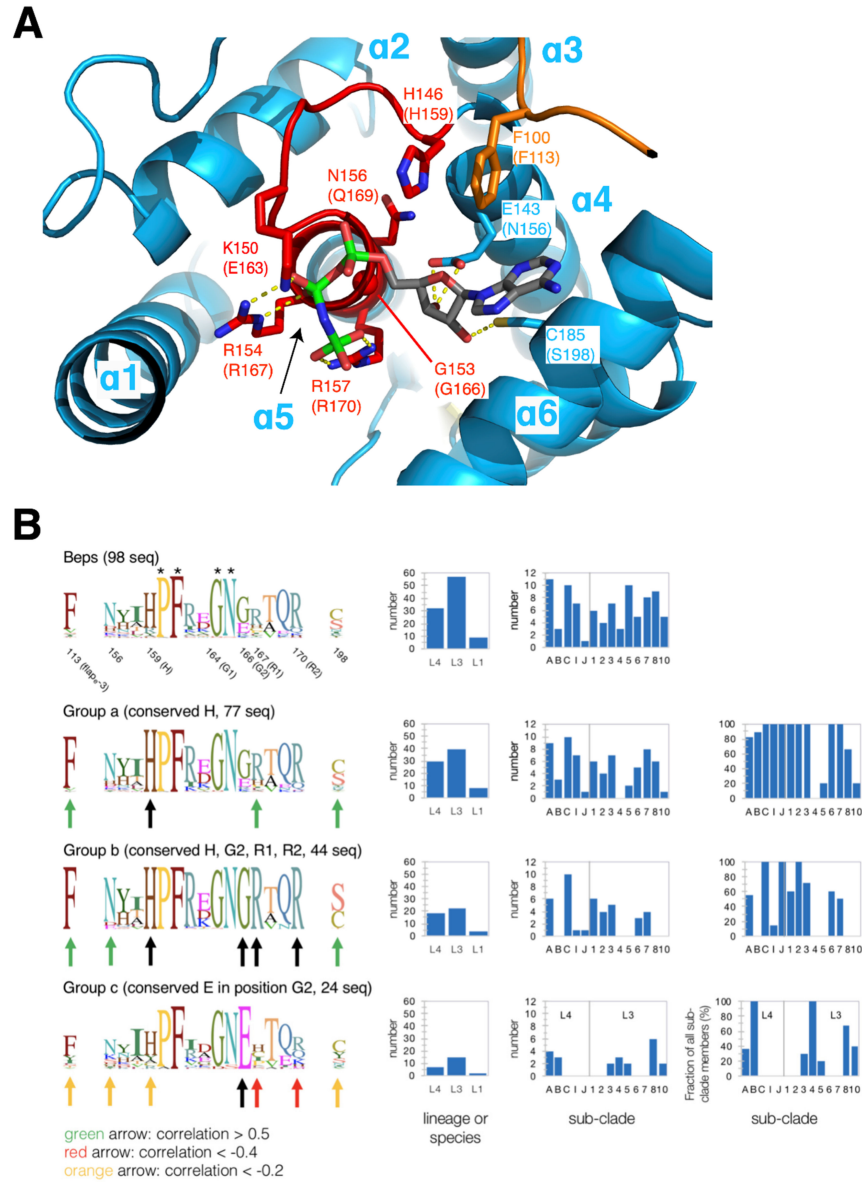
Figure 6. FIC substrate binding site: conservation and correlation. (**A**) Bhe_BepC structure with bound substrate analog AMPPNP (PDB ID: 4WGJ) corresponding to Figure 2C, but zoomed into the ligand binding pocket. The catalytic histidine H146 and residues interacting with the ligand are labeled (Bhe_BepA residue types and numbers in brackets). (**B**) Sequence logos encompassing the residues are shown in panel (**A**). The overall logo derived from all BepFic sequences as presented in Figure 3 is shown at the top followed by the logos of selected groups that comply to the criteria provided above the corresponding graph. Residues that show correlation or anti-correlation with the selected residue(s) (black arrows) are indicated by green (corr. > 0.5), orange (corr. < −0.2), or red (corr. < −0.4) arrows. A break-down of the group members into lineages and sub-clades (only for L3 and L4 members) is provided on the right side.

The histidine required for the catalysis of AMP transfer is present in about 75% of the BepFIC sequences and the logo representing this sub-group (group a) is shown in the second row of Figure 6B. It reveals enrichment (green arrows) of F, R, and S/T in positions 113, 167, and 198, respectively, as expected for Fic proteins that are competent for nucleotide binding.

By further filtering for the presence of the second glycine (G2) and the two arginines (R1 and R2) of the canonical motif, 44 proteins are yielded (group b). Interestingly, this group almost invariably shows canonical residue types in positions 113 and 198 and N/D in position 156 (poised for ribose binding; see, e.g., Figure 6A). Therefore, we can classify group b as bona fide AMP transferases.

The second most common residue type in G2 position is glutamate (see overall logo in Figure 6B). The presence of a bulky residue in this position appears incompatible with triphosphate-nucleotide binding as will be discussed further below (Bep8 structures). 20 sequences show a glutamate in position G2 (group c) and the corresponding sequence logo in Figure 6B (bottom) clearly demonstrates that residue types required for ATP binding or AMP transfer are strongly under-represented. This suggests that group c comprises proteins with unknown functions, but see below for a further discussion.

In order to substantiate the above analysis, the co-conservation of ATP binding residues was quantified by simple pair-wise correlation calculations on all aligned sequences (see Materials and Methods). Indeed, there is a significant (>0.4) correlation between the residues that are involved in ATP binding in bona fide AMP transferases (Figure S4). Noteworthily, this correlation has no structural reason since the residues are exposed and do not interact with each other, but rather points to their common functional role. The strong anti-correlation between an E in the G2 position with canonical residue types in the other positions again points to a distinct function of the respective BepFICs.

For each group shown in Figure 6B, enumerations of Beps aggregated by lineages (L4, L3, or L1) and by orthologous sub-clades (A–C, I, J, and 1–10) are shown at the right side. Most sub-clades, but not B, 4, 5, 8, and 10, contain bona-fide AMP transferases (group b), with many orthologous groups possessing members that belong to group c (E in position G2). The remaining 20 sequences (not classified as b or c) belong to sub-clades A, I, 5–8, and 10.

3.4. Variability of Target Binding Flap

Fic proteins catalyze covalent modification of target hydroxyl side-chains. For this, the backbone of the target segment following the modifiable side-chain has to register to the N-terminal strand of the FIC flap (β-sheet augmentation, see Figure 7A). In our definition, the flap extends from the end of the Bep element (flap_start, P102 in *Bhe*_BepA) to the start of helix α3 (flap_end, G116 in *Bhe*_BepA, Figure 7B). Part of the flap residues point towards the FIC core and would, thus, not be in direct contact with a bound target protein. The phenylalanine in position flap_end-3 (F113 in *Bhe*_BepA) has been shown to interact with the catalytic histidine and the adenine moiety in AMP transferases (Figure 6A). The residues in positions flap_start+2 and flap_start+4 are also pointing inwards and they form two opposing walls of a slot into which the modifiable target side-chain is inserted as is observed in the IbpA/Cdc42 complex (residues L3668 and K3670 in Figure 7A). The latter residue is often a lysine or arginine which hovers over the α-phosphate of the AMPylated target tyrosine in IbpA/Cdc42 (Figure 7A) or of the ATP substrate in NmFic [1] and most likely has the additional role of stabilizing the transition state.

As noted before, Bep flaps show considerable sequence (Figure 3) and structure (Figure 7C) variation. Since the flap appears to be in a strategic position to contribute to target recognition (in addition to the sequence independent β-sheet augmentation), its variability may reflect distinct targets for individual Beps. Moreover, loss of function has to be considered.
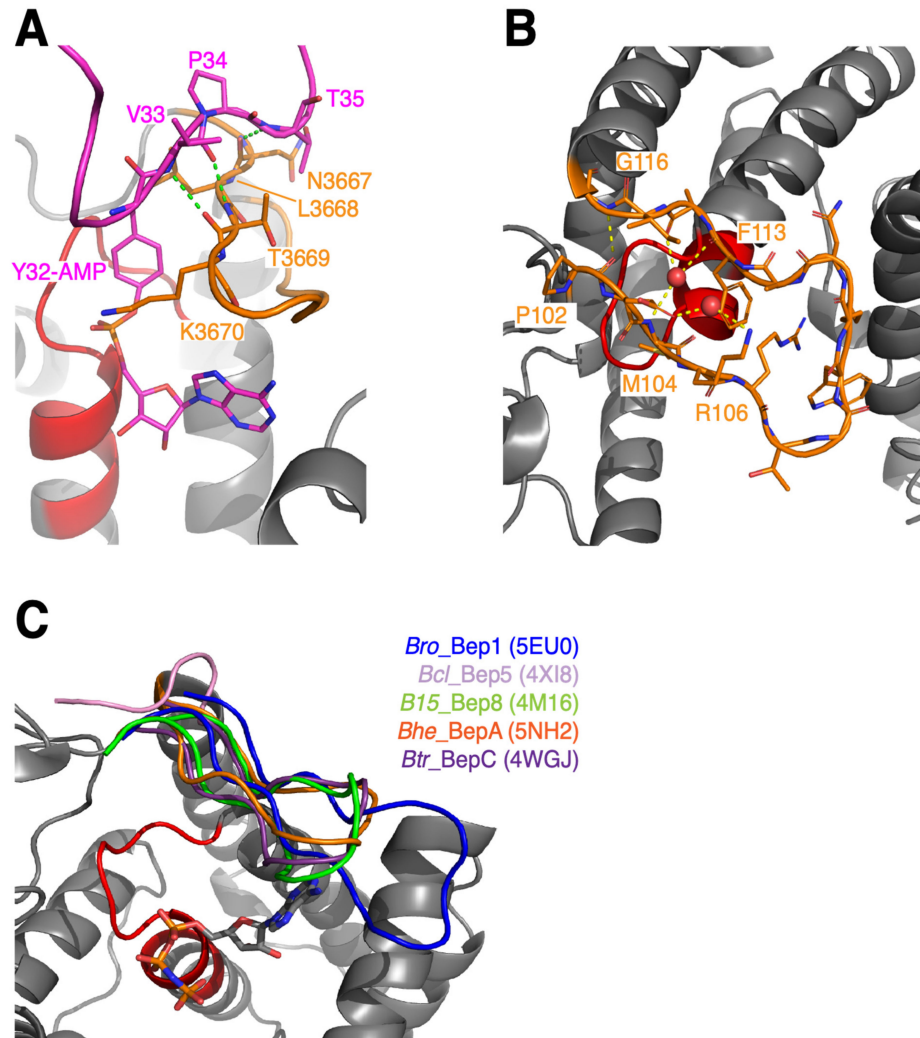
Figure 7. FIC flap structures as observed in the IbpA/Cdc42 complex and in BepFICs. The FIC signature motif locates to the elements shown in red and flaps are shown in orange or as indicated. (**A**) IbpA(fic2)/Cdc42-AMP (PDB: 4ITR, [14]) product complex showing registration of the target switch-1 loop with AMPylated Y32 (pink) to the N-terminal strand of the IbpA(fic2) flap via anti-parallel β-strand augmentation. (**B**) Structure of *Bhe*_BepA (PDB: 5NH2) showing the flap in detail. Note that the flap shows an open β-hairpin conformation, with some main-chain interactions (yellow dashes) mediated by water molecules (red spheres). (**C**) Comparison of BepFic flap structures (with color code as indicated) after the superposition of FIC core structures. The grey cartoon shows *Bhe*_BepA with bound AMPPNP ligand as in Figure 2C.

In order to obtain insight into the variability and sub-grouping of the flap sequences, we extracted the corresponding segments from the global BepFic alignment and computed a cladogram. For this, the flap sequences were not re-aligned to ensure their unchanged relative position with the (conserved) adjoining elements. The cladogram (Figure S5) allows the definition of 12 branches composed of five to eight members each with pair-wise sequence identities between 86 and 25% (Figure S6). Although the sub-grouping may

be ambiguous in some instances due to the short sequence length, the cladogram clearly indicates some well-defined branches (see the excerpt of the cladogram shown in Figure 8). Such segregation seems indicative of the functional diversification of the flaps, which is possibly related to their interaction with distinct targets.
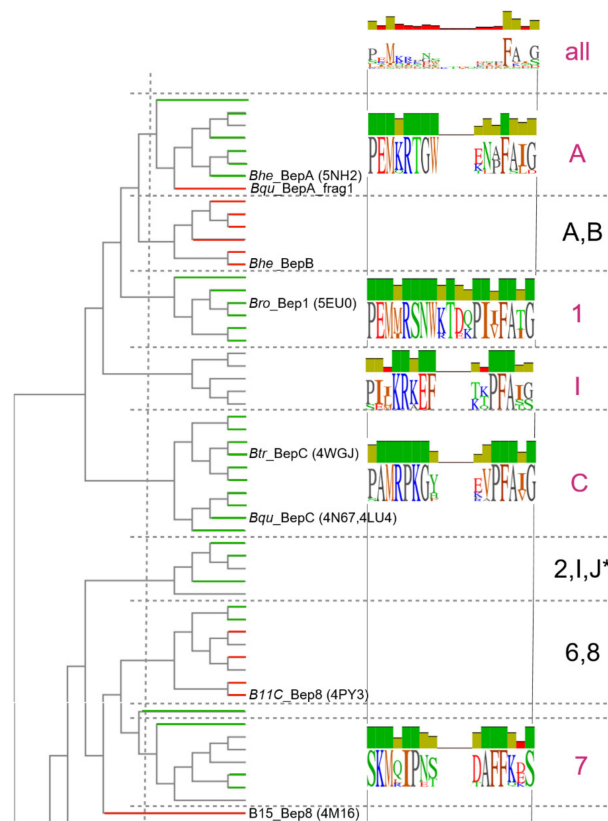
Figure 8. Excerpt of cladogram derived from BepFIC flap sequences with annotated sequence logos for selected branches. Horizontal lines colored in green or red indicate Beps with canonical FIC motif or with a glutamate in position G2 (groups b and d, respectively, in Figure 6B). The full cladogram with all logos and sequence names is shown in Figure S5. Branches were selected for being comprised of at least 5 members that have a pairwise identity score that is larger than the conservation score of the respective full-length sequence, see Figure S6. Branches are labeled according to the sub-clades of their members and the asterisk indicates the additional presence of Ban members (lineage 1).

Branches A, C, I, 1, and 7 (all being composed of members of only one sub-clade) show top-scoring pairwise flap sequence identities, which are larger than the overall sequence identity of the respective BepFICs (Figure S6). This suggests that these flap sequences are conserved for functional reasons. Figure 8 shows their sequence logos together with the pertinent part of the cladogram.

All members of branch 1 are predicted to catalyze AMP transfer (indicated in green in the cladogram). This has been verified for *Bro*_Bep1, which AMPylates the switch 1 loop of small GTPases [16]. Most members of branch A are putative AMP transferases as well.

Although their target(s) have not yet been identified, it is demonstrated that *Bhe*_BepA catalysed AMPylation of proteins from the Hela cell lysate [6].

The A,B branch is interesting, since it shows a similar flap logo as branch A (Figure S5) suggesting the same target, but all of its members have a very distinct FIC motif with an E in G2 position (indicated in red in the cladogram). *Bro*_Bep2, which belongs to branch 2, was shown to AMPylate vimentin [25]. Its flap sequence resembles that of branch A, but shows a conserved D in the flap_start+6 position. However, this residue is predicted to point inwards and so it may not directly be responsible for target specificity.

3.5. Distinctive Features of BepFIC Domains

In general, the protein function is encoded in its three-dimensional structure. Thus, it is valuable to analyze the defining features of BepFIC sub-clades (as revealed by the sequence analyses) in the light of the representative crystal structures. The role of the individual residues of the canonical FIC motif in AMP transfer has been discussed in detail in the first chapter and elsewhere [15].

The structures of Bhe_BepA, Bcl_Bep1, Bqu_BepC, and Btr_BepC exhibit all features defining a FIC AMP transferase as described in the first chapter and elsewhere [15]. Indeed, the complex of Btr_BepC with AMPPNP (Figure 9A) demonstrates competent tri-phosphate nucleotide binding as has been observed for, e.g., the inhibition-relieved VbhT/VbhA(E24G) complex [15]. However, BepC proteins are special in that they have a lysine instead of a Mg^{2+}-coordinating aspartate/glutamate in the FIC motif. The structure shows that this lysine (K150) is interacting directly with the α-phosphate. Thus, BepC proteins appear to require no divalent cation for substrate binding, in contrast to all FIC AMP transferases characterized so far. The structure also verifies that a cysteine (C185) can substitute the ribose binding serine (last position of the active site motif used in Figure 6B).

BepFIC domains with a glutamate (E159) in position G2 are represented by the *B15*_Bep8 and *B11C*_Bep8 structures, which are virtually identical. Although the motif is highly degenerated to PFxxGNE, the fold of the domain and, in particular, of the active loop are canonical (Figure 9B). Nevertheless, the binding of a triphosphate-nucleotide appears to be impossible due to the presence of the side-chain in position G2 (compare Figures 9A,B). However, simple outward rotation of the glutamate side-chain (Figure 9C) allows the accommodation of an AMPylated side-chain in a constellation very similar to that of the Tyr-AMP target side-chain in the IbpA/Cdc42 complex [14]. Intriguingly, the model would allow binding of water molecule to the N-terminal end of α5 ready for a nucleophilic attack onto the phosphorous, which is in line with its phosphoester bond with the tyrosine, i.e., it is ready for deAMPylation of the modified side-chain. The strict conservation of a glutamate in G2 may then point to a role in water binding and proton abstraction, i.e., the glutamate would function as general base.

Members of the Bep5 sub-clade are characterized by the absence of the flap. Indeed, the structure of Bcl_Bep5 (Figures 5 and 9D) shows that the Bep element is directly followed by α3. The structurally required residues of the FIC signature motif (PFxx(N/S)G) are conserved, explaining the canonical fold of the FIC loop, but most other residues of the FIC signature motif are distinct (Figure 9D).
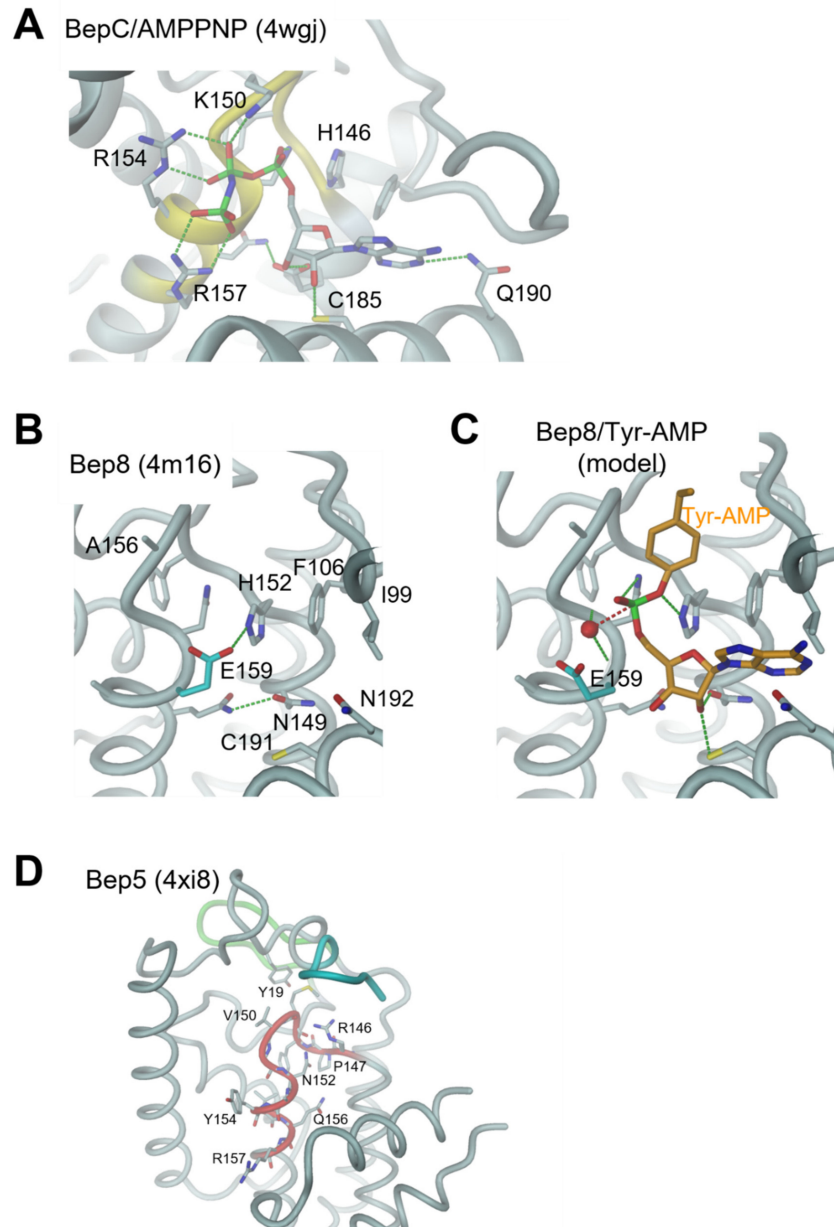
Figure 9. Structural details of the BepFIC domains. (**A**) *Btr*_BepC in complex with AMPPNP. (**B**) *B15*_Bep8 with E159 in position G2 highlighted in cyan. (**C**) Modeled complex of Bep8 with Tyr-AMP (orange) in a position as derived from the IbpA/Cdc42 complex 4ITR. The side-chain conformation of E159 (cyan) has been altered to allow Tyr-AMP binding. The putative hydrolytic water (red sphere) is bound to the N-terminus of α5 and possibly to E159. (**D**) *Bcl*_Bep5; note the missing flap that in other BepFIC domains is inserted between α2 and the Bep element (green). Due to disorder, K153 is shown only up to Cβ. The N-termal His-tag is indicated in cyan.

3.6. Interaction of Beps with FicA Antitoxins

Class I FicT toxins are typically controlled by small FicA antitoxins. These proteins are comprised of three antiparallel helices which bind (mainly via hydrophobic interactions) to α1 of the FicT toxin, as has been shown for VbhT/A [1] and Ec_FicT/A [45]. The C-terminal end of the first anti-toxin helix (α_{inh}) typically carries an (S/T)xxxE(G/N) motif with the glutamate forming a salt-bridge with the second arginine (R2) of the FIC signature motif. As a result, this arginine can no longer engage in ATP γ-phosphate binding, leaving the bound ATP substrate in a non-reactive orientation [1].

We have shown previously that all investigated Bartonellae encode one (or some even two) FicA homolog, which are called BiaA [24] and are homologous to VbhA. To obtain further insight into their mode of binding, several of the investigated Beps were co-expressed with their cognate BiaA and subjected to crystallization. Crystals were obtained and the structures solved for *Bhe*_BepA/*Bhe*_BiaA and *Bro*_Bep1/*Bro*_BiaA (Figures 4 and 10). The two BiaA structures and their association modes with the BepFICs are very similar. The 3-helix up-and-down fold of the antitoxin is complemented by α1 of the toxin to form a canonical antiparallel 4-helix bundle, where α1 of the toxin is arranged anti-parallel to the antitoxin α1-helices and α3-helices. A large part of the antitoxin surface (about 1100 Å^2, amounting to about 25% of the solvent accessible surface area) is buried and numerous interactions of all kinds are formed (see annotations below the sequences in Figure 10C). Although VbhA and Ec_FicA are only distantly related to BiaA antitoxins (22% and 12% sequence identity, respectively, w/r to *Bhe*_BiaA), a similar arrangement is observed in the VbhTA and Ec_FicTA complexes (Figure S7). In particular, the conserved serine/threonine and glutamate of the antitoxin motif are, again, forming interactions with the R2 side-chain of the toxin.
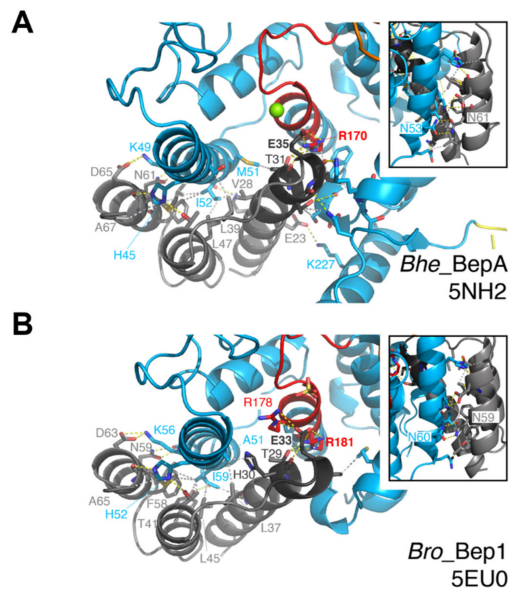

Figure 10. *Cont.*
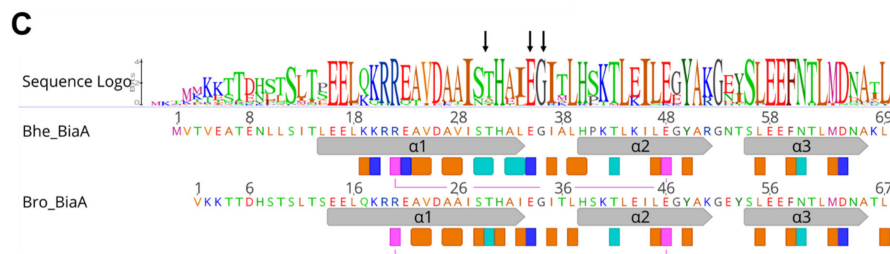
Figure 10. Structures of BepFIC/antitoxin complexes and antitoxin logo. (**A**,**B**) Crystal structures of *Bhe*_BepA/BiaA$_{L4}$ (5NH2) and *Bro*_Bep1/BiaA$_{L3}$ (5EU0); representation as in Figure 2, with interface residues shown in full. The insets show the view from the back to visualize the conserved intermolecular Asn–Asn interaction. The H-bonds between residues (cut-off 3.7 Å) are shown as yellow dashes and apolar interactions (cut-off 3.7 Å) are shown as grey dashes. (**C**) BiaA sequence logo derived from 23 *Bartonella* sequences (as defined in Figure 6 of [24], but without the divergent BiaA/2 from *B. taylori*) with annotated *Bhe*_BiaA and *Bro*_BiaA sequences underneath. Arrows above the logo indicate the residues of the inhibitory (S/T)xxxE(G/N) motif [1]. Secondary structure (grey arrows) and residues interacting with the toxin (as determined by PISA) are indicated by rectangles below the sequences (orange: apolar contact with at least 50% of the residue buried in the complex, light blue: H-bond, dark blue: salt-bridge). The magenta rectangles connected by a line indicate a conserved intra-molecular salt-bridge between α1 and α2.

4. Discussion

The host-targeted Bep effector repertoires that evolved in parallel in three distinct *Bartonella* lineages represent a paradigm for diversifying evolution with the FIC domain mediating PTMs of target proteins representing one of their central functional units. In this work, we have explored the evolutionary trajectories of BepFIC domains along their structural and functional diversification from a single common ancestor and, ultimately, their exaptation for host interaction from a bacterial toxin-antitoxin module. As a heritage of this deep ancestry, all Bartonellae encoding FIC-BID Beps also encode at least one BiaA protein. As BiaAs forms stable complexes with BepFICs, which is exemplified by *Bro*_Bep1/*Bro*_BiaA [16] or *Bhe*_BepA/*Bhe*_BiaA (this work), they may—similarly to the homologous FicA antitoxins—inhibit AMPylation activity and possibly mediate de-AMPylation, but could also serve as chaperones for BepFICs before translocation into host cells via the T4S system. Since T4S is believed to require partial or full unfolding of its substrates, the FicA homolog will be stripped off from BepFIC domains upon translocation and thereby unleash the Beps' activity in the eukaryotic target cell. The exceptions are the few Beps that harbor an N-terminally fused copy of an antitoxin sequence (Bep3/1, Bep3/2, and Bep4; see [24]) and they are, thus, representative of the class III Fic proteins according to the definition by Engel et al. [1].

Most class I Fic proteins were proposed to act as FicT toxins by AMPylating endogenous topoisomerases, an activity that is kept in check by the cognate FicA antitoxin [19]. For this reason, they exhibit a conserved ATP binding site and a conserved juxtaposed β-hairpin, which has the dual function to register the modifiable target segment to the active site and to contribute to target affinity/specificity.

This study showed that both the BepFIC overall structure and the backbone structure of the active site are very well conserved. However, the residue types of the active site and the sequence and structure of the target docking site show large variability, which is believed to reflect functional diversification. About half of the BepFIC domains display a canonical FIC signature motif indicating that they are phosphotransferases operating according to the well-studied catalytic mechanism with the histidine acting as a general base and a triphosphate-nucleotide as a substrate. Indeed, AMP transfer activity has been confirmed experimentally for some representatives (i.e., Bep1, Bep2, and BepA). For BepC, despite its canonical signature motif, no catalytic activity has yet been found, but its BepFIC

domain mediates protein–protein interaction with the RhoGEF GEF-H1, thereby activating RhoA signaling [26,27]. It should be noted that the presence of the FIC signature motif does not permit inference about the kind of substrate nucleobase that is used by the enzyme. This is exemplified by the UMP transferase of AvrAC with its canonical motif [18].

About one quarter of BepFIC domains display a glutamate in the G2 position which would clearly interfere with the triphosphate-nucleotide binding mode observed in AMP transferases, i.e., with the γ-phosphate forming a salt-bridge with R2. However, binding of an AMPylated target to members of this group of BepFICs appears possible, which suggests that they might act as deAMPylases with the glutamate in G2 involved in binding and deprotonation of the hydrolytic water. DeAMPylation activity has been reported for the class II Fic protein FICD in humans [12] and the class III Fic protein *Ef*Fic from *Enterococcus faecalis* [11] and appears to require the "inhibitory" glutamate in a similar position as the glutamate in G2. It should be noted that only half of the members of this group have the histidine conserved and, therefore, would qualify for potential deAMPylases. It may be noted that the distantly related Fic protein Doc also has a side-chain (lysine) in G2 position, which would prevent canonical ATP binding. Indeed, it has been reported that ATP binds in an apparently inverted orientation, which explains the kinase as opposed to AMP transfer activity of Doc [10].

The remaining quarter of Beps are highly variable in their substrate binding pockets and, thus, may catalyze other PTMs or may lose enzymatic activity, but may still mediate specific effector functions such as those found for BepC [26,27].

The flap predicted to interact with the target displays considerable variability in length and amino acid composition among the Bep paralogs, but is conserved among orthologous Bep groups (Figure 8). None of the identified logos matches the well-conserved V(D/E)IxKxxxxF(A/C)xx motif of the flap of FicT toxins (see also [19] and subfamilies 1 and 2 as identified in reference [46]) from which the BepFIC domains emerged as a mono-phyletic group.

Very recently, it has been shown that small GTPases of the Rac-subfamily are the target for Bep1 and that two residues in the extended flap are critical for the exquisite target specificity [16]. With the use of modeling, this finding was rationalized by suggesting that they form two salt-bridges with Rac-subfamily specific residues. It will be most relevant for a deep understanding of BepFIC-target interactions to identify more of the targets and eventually determine the structures of the respective complexes. Whether BepA and BepB recognize the same target—due to their distinct active sites—but confer different actions (e.g., AMPylation and deAMPylation) remain an interesting question.

BepFIC domains generally carry a β-hairpin preceding the flap and a highly conserved N-terminal segment that is absent from other FIC domains. The function of this unique feature of the BepFIC domains is presently unknown, but may be related to the fact that these domains must partially unfold in the process of T4S system-mediated translocation.

The crucial role of Bep diversification for the host adaptability of the Bartonella lineages L3 and L4 is obvious from the fact that, in both cases, the adaptive radiation of these now ubiquitous groups of pathogens was triggered only after the full extant effector repertoires had evolved in their respective common ancestors [21,24]. Although much remains unknown about the biological functions of this diversity, including the more recent expansion of a complex Bep repertoire in *B. ancashensis* of L1 [24], it is clear that the variability of the BepFIC domains stick out of the huge pool of class I Fic proteins that act as FicT toxins by AMPylating bacterial topoisomerases and exhibiting a canonical or near-canonical FIC signature motif [19].

Since the lifestyle and infection cycle of *Bartonella* have largely cut them off from horizontal gene flow as an alternative source of pathogenicity factors, it seems plausible to assume that strong selective pressures have forced them to exploit the intrinsic functional versatility and plasticity of the FIC domain to the maximum. This is the opposite concept of what is represented by, e.g., *Salmonella* bacteria that profit from a continuous flow of innovations and fully functional upgrades for their virulence machineries in the spirit of

"open source evolution" [47,48]. With a reference to popular culture, we favor the view that a lack of readily available and fully evolved virulence factors forced *Bartonella* to become the "MacGyver of bacterial pathogens". In this model, strong selective pressures promoted the use of a couple of suitable raw materials to evolve highly effective and innovative, yet not necessarily elegant or elaborate pathogenicity factors de novo.

Our analysis revealed an unexpected functional plasticity of Beps that is brought about by minor structural changes in specific elements of the substrate pocket and the target dock. Future structure–function studies will be key for a better understanding of the remarkable functional plasticity of FIC domain proteins and how they contribute to host adaptability as a crucial feature of *Bartonella* evolution. This studies may also allow us to address the more fundamental question of how a conserved enzymatic scaffold, such as that of the Bep-ancestral FicT toxin that AMPylates bacterial topoisomerases, can functionally diversify over short evolutionary timescales as exemplified by the Beps in the process of host adaptation.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/microorganisms9081645/s1, Figure S1: Full-length sequence logo of FIC-BID Beps, Figure S2: Alignment of Bhe_BepA with outgroup members VbhT and Ec_FicT, Figure S3: Excerpt of the overall BepFic sequence alignment covering Bhe_BepA active site, Figure S4: Correlation (see Methods) between selected residues of the non-contiguous active site motif shown in Figure 6, Figure S5: Cladogram derived from all BepFIC flap sequences used in this study (with the exception of the flap-less Bep5 sequences), Figure S6: Overall (blue) and flap (green) sequence identity of the flap branches as shown in the cladogram of Figure S5, Figure S7: Toxin/Antitoxin interaction, Table S1: List of species abbreviations and genome references, Table S2: List of FIC-BID Beps used in this study, Table S3: Data collection and refinement statistics.

Author Contributions: Conceptualization, T.S., T.A.P.d.B., A.H., P.J.M., I.P. and C.D.; methodology, T.S., T.A.P.d.B., A.H., S.T., A.H., D.M.D. and T.E.E.; validation, T.S. and I.P.; formal analysis, T.S., A.H., N.D. and T.A.P.d.B.; investigation, T.S., T.A.P.d.B., A.H., S.T., A.H., D.M.D. and T.E.E.; resources, P.J.M. and I.P.; data curation, T.S., S.T., N.D. and A.H.; writing—original draft preparation, T.S., T.A.P.d.B., A.H. and C.D.; writing—review and editing, T.S. and C.D.; visualization, T.S., A.H. and S.T.; supervision, T.S. and C.D.; funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Crystal structures have been deposited in the PDB database (www.pdb.org accessed on 1 June 2021) under accession codes 5NH2, 4N67, 4LU4, 4WGJ, 4NPS, 4XI8, 4M16, and 4PY3.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Engel, P.; Goepfert, A.; Stanger, F.; Harms, A.; Schmidt, A.; Schirmer, T.; Dehio, C. Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature* **2012**, *482*, 107–110. [CrossRef]
2. Kinch, L.N.; Yarbrough, M.L.; Orth, K.; Grishin, N.V. Fido, a Novel AMPylation Domain Common to Fic, Doc, and AvrB. *PLoS ONE* **2009**, *4*, e5818. [CrossRef]
3. Garcia-Pino, A.; Zenkin, N.; Loris, R. The many faces of Fic: Structural and functional aspects of Fic enzymes. *Trends Biochem. Sci.* **2014**, *39*, 121–129. [CrossRef]

4. Harms, A.; Stanger, F.V.; Dehio, C. Biological Diversity and Molecular Plasticity of FIC Domain Proteins. *Annu. Rev. Microbiol.* **2016**, *70*, 341–360. [CrossRef]
5. Roy, C.R.; Cherfils, J. Structure and function of Fic proteins. *Nat. Rev. Genet.* **2015**, *13*, 631–640. [CrossRef] [PubMed]
6. 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.* **2011**, *20*, 492–499. [CrossRef]
7. 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* **2009**, *34*, 93–103. [CrossRef]
8. 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* **2009**, *323*, 269–272. [CrossRef] [PubMed]
9. Campanacci, V.; Mukherjee, S.; Roy, C.; Cherfils, J. Structure of the Legionella effector AnkX reveals the mechanism of phosphocholine transfer by the FIC domain. *EMBO J.* **2013**, *32*, 1469–1477. [CrossRef] [PubMed]
10. Castro-Roa, D.; Garcia-Pino, A.; De Gieter, S.; Van Nuland, N.A.; Loris, R.; Zenkin, N. The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nat. Chem. Biol.* **2013**, *9*, 811–817. [CrossRef] [PubMed]
11. Veyron, S.; Oliva, G.; Rolando, M.; Buchrieser, C.; Peyroche, G.; Cherfils, J. A Ca(2+)-regulated deAMPylation switch in human and bacterial FIC proteins. *Nat. Commun.* **2019**, *10*, 1–10. [CrossRef] [PubMed]
12. Perera, L.A.; Rato, C.; Yan, Y.; Neidhardt, L.; McLaughlin, S.H.; Read, R.J.; Preissler, S.; Ron, D. An oligomeric state-dependent switch in the ER enzyme FICD regulates AMPylation and deAMPylation of BiP. *EMBO J.* **2019**, *38*, e102177. [CrossRef]
13. Veyron, S.; Peyroche, G.; Cherfils, J. FIC proteins: From bacteria to humans and back again. *Pathog. Dis.* **2018**, *76*. [CrossRef]
14. Xiao, J.; Worby, C.A.; Mattoo, S.; Sankaran, B.; Dixon, J.E. Structural basis of Fic-mediated adenylylation. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1004–1010. [CrossRef] [PubMed]
15. Goepfert, A.; Stanger, F.V.; Dehio, C.; Schirmer, T. Conserved Inhibitory Mechanism and Competent ATP Binding Mode for Adenylyltransferases with Fic Fold. *PLoS ONE* **2013**, *8*, e64901. [CrossRef]
16. Dietz, N.; Huber, M.; Sorg, I.; Goepfert, A.; Harms, A.; Schirmer, T.; Dehio, C. Structural basis for selective AMPylation of Rac-subfamily GTPases by Bartonella effector protein 1 (Bep1). *Proc. Natl. Acad. Sci. USA* **2021**, *118*. [CrossRef]
17. Mukherjee, S.; Liu, X.; Arasaki, K.; McDonough, J.A.; Galán, J.E.; Roy, C.R. Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **2011**, *477*, 103–106. [CrossRef]
18. Feng, F.; Yang, F.; Rong, W.; Wu, X.; Zhang, J.; Chen, S.; He, C.; Zhou, J.-M. A Xanthomonas uridine 5′-monophosphate transferase inhibits plant immune kinases. *Nature* **2012**, *485*, 114–118. [CrossRef] [PubMed]
19. Harms, A.; Stanger, F.; Scheu, P.D.; de Jong, I.G.; Goepfert, A.; Glatter, T.; Gerdes, K.; Schirmer, T.; Dehio, C. Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Rep.* **2015**, *12*, 1497–1507. [CrossRef] [PubMed]
20. Harms, A.; Dehio, C. Intruders below the Radar: Molecular Pathogenesis of *Bartonella* spp. *Clin. Microbiol. Rev.* **2012**, *25*, 42–78. [CrossRef]
21. Engel, P.; Salzburger, W.; Liesch, M.; Chang, C.-C.; Maruyama, S.; Lanz, C.; Calteau, A.; Lajus, A.; Médigue, C.; Schuster, S.C.; et al. Parallel Evolution of a Type IV Secretion System in Radiating Lineages of the Host-Restricted Bacterial Pathogen Bartonella. *PLoS Genet.* **2011**, *7*, e1001296. [CrossRef]
22. 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.* **2007**, *39*, 1469–1476. [CrossRef]
23. Guy, L.; Nystedt, B.; Toft, C.; Zaremba-Niedzwiedzka, K.; Berglund, E.C.; Granberg, F.; Näslund, K.; Eriksson, A.-S.; Andersson, S.G.E. A Gene Transfer Agent and a Dynamic Repertoire of Secretion Systems Hold the Keys to the Explosive Radiation of the Emerging Pathogen Bartonella. *PLoS Genet.* **2013**, *9*, e1003393. [CrossRef]
24. Harms, A.; Segers, F.H.; Quebatte, M.; Mistl, C.; Manfredi, P.; Körner, J.; Chomel, B.B.; Kosoy, M.; Maruyama, S.; Engel, P.; et al. Evolutionary Dynamics of Pathoadaptation Revealed by Three Independent Acquisitions of the VirB/D4 Type IV Secretion System in Bartonella. *Genome Biol. Evol.* **2017**, *9*, 761–776. [CrossRef] [PubMed]
25. Pieles, K.; Glatter, T.; Harms, A.; Schmidt, A.; Dehio, C. An experimental strategy for the identification of AMPylation targets from complex protein samples. *Proteomics* **2014**, *14*, 1048–1052. [CrossRef]
26. Marlaire, S.; Dehio, C. Bartonella effector protein C mediates actin stress fiber formation via recruitment of GEF-H1 to the plasma membrane. *PLoS Pathog.* **2021**, *17*, e1008548. [CrossRef]
27. Wang, C.; Zhang, H.; Fu, J.; Wang, M.; Cai, Y.; Ding, T.; Jiang, J.; Koehler, J.E.; Liu, X.; Yuan, C. Bartonella type IV secretion effector BepC induces stress fiber formation through activation of GEF-H1. *PLoS Pathog.* **2021**, *17*, e1009065. [CrossRef]
28. Schulein, R.; Guye, P.; Rhomberg, T.A.; Schmid, M.C.; Schroder, G.; Vergunst, A.; 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. USA* **2005**, *102*, 856–861. [CrossRef] [PubMed]
29. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [CrossRef] [PubMed]
30. Larkin, M.; Blackshields, G.; Brown, N.; Chenna, R.; McGettigan, P.; McWilliam, H.; Valentin, F.; Wallace, I.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2. *0. Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef] [PubMed]
31. Alexandrov, A.; Vignali, M.; LaCount, D.J.; Quartley, E.; de Vries, C.; De Rosa, D.; Babulski, J.; Mitchell, S.F.; Schoenfeld, L.W.; Fields, S.; et al. A Facile Method for High-throughput Co-expression of Protein Pairs. *Mol. Cell. Proteom.* **2004**, *3*, 934–938. [CrossRef] [PubMed]

32. Aslanidis, C.; de Jong, P.J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* **1990**, *18*, 6069–6074. [CrossRef] [PubMed]

33. Bryan, C.M.; Bhandari, J.; Napuli, A.J.; Leibly, D.J.; Choi, R.; Kelley, A.; Van Voorhis, W.C.; Edwards, T.E.; Stewart, L.J. High-throughput protein production and purification at the Seattle Structural Genomics Center for Infectious Disease. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2011**, *67 Pt 9*, 1010–1014. [CrossRef]

34. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66 Pt 2*, 133–144. [CrossRef]

35. 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.* **2007**, *40*, 658–674. [CrossRef] [PubMed]

36. Murshudov, G.N.; Vagin, A.A.; Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53 Pt 3*, 240–255. [CrossRef]

37. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66 Pt 2*, 213–221. [CrossRef]

38. Painter, J.; Merritt, E.A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62 Pt 4*, 439–450. [CrossRef]

39. Emsley, P.; Cowtan, K.D. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60 Pt 12*, 2126–2132. [CrossRef]

40. Chen, V.; Arendall, W.B.; Headd, J.J.; Keedy, D.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *66 Pt 1*, 12–21. [CrossRef]

41. Blanc, E.; Roversi, P.; Vonrhein, C.; Flensburg, C.; Lea, S.; Bricogne, G. Refinement of severely incomplete structures with maximum likelihood in BUSTER–TNT. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60 Pt 12*, 2210–2221. [CrossRef]

42. Stanger, F.; de Beer, T.A.; Dranow, D.M.; Schirmer, T.; Phan, I.; Dehio, C. The BID Domain of Type IV Secretion Substrates Forms a Conserved Four-Helix Bundle Topped with a Hook. *Structure* **2017**, *25*, 203–211. [CrossRef]

43. Wagner, A.; Tittes, C.; Dehio, C. Versatility of the BID Domain: Conserved Function as Type-IV-Secretion-Signal and Secondarily Evolved Effector Functions Within Bartonella-Infected Host Cells. *Front. Microbiol.* **2019**, *10*, 921. [CrossRef]

44. Stanger, F.; Burmann, B.M.; Harms, A.; Aragão, H.; Mazur, A.; Sharpe, T.; Dehio, C.; Hiller, S.; Schirmer, T. Intrinsic regulation of FIC-domain AMP-transferases by oligomerization and automodification. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E529–E537. [CrossRef]

45. Stanger, F.; Harms, A.; Dehio, C.; Schirmer, T. Crystal Structure of the Escherichia coli Fic Toxin-Like Protein in Complex with Its Cognate Antitoxin. *PLoS ONE* **2016**, *11*, e0163654. [CrossRef] [PubMed]

46. Khater, S.; Mohanty, D. Deciphering the Molecular Basis of Functional Divergence in AMPylating Enzymes by Molecular Dynamics Simulations and Structure Guided Phylogeny. *Biochemistry* **2015**, *54*, 5209–5224. [CrossRef] [PubMed]

47. Frost, L.; Leplae, R.; Summers, A.; Toussaint, A. Mobile genetic elements: The agents of open source evolution. *Nat. Rev. Genet.* **2005**, *3*, 722–732. [CrossRef] [PubMed]

48. Mebrhatu, M.T.; Cenens, W.; Aertsen, A. An overview of the domestication and impact of the Salmonella mobilome. *Crit. Rev. Microbiol.* **2014**, *40*, 63–75. [CrossRef]

Supplementary Materials

Figure S1. Full-length sequence logo of FIC-BID Beps. Representation as in Fig. 3. Sections corresponding to the BepFIC, OB, and BID domains are indicated at the left. Secondary structure of the FIC and OB domains (helices α1 to α8, α', etc.) has been derived from *Bhe*_BepA (5NH2), secondary structure of the BID domain (helices α1' to α5', etc.) has been derived from *Bro*_Bep6 (4YK1).

1
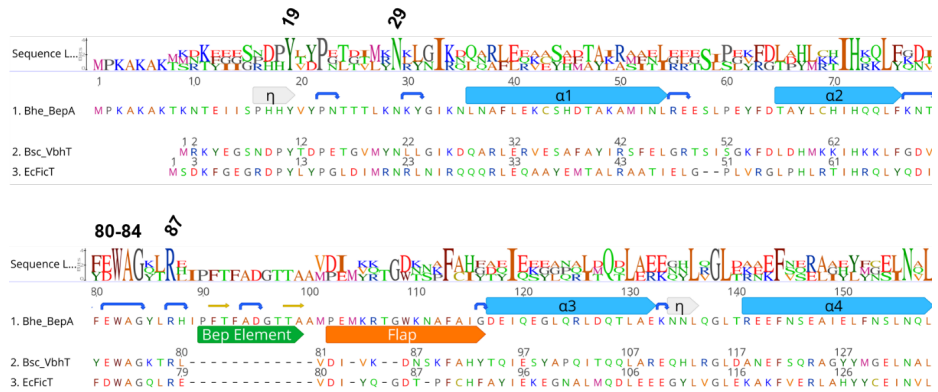
Figure S2. Alignment of *Bhe*_BepA with outgroup members VbhT and *Ec*_FicT.
The alignment shows that the N-terminus is conserved, while the Bep element is
specific to the Beps.
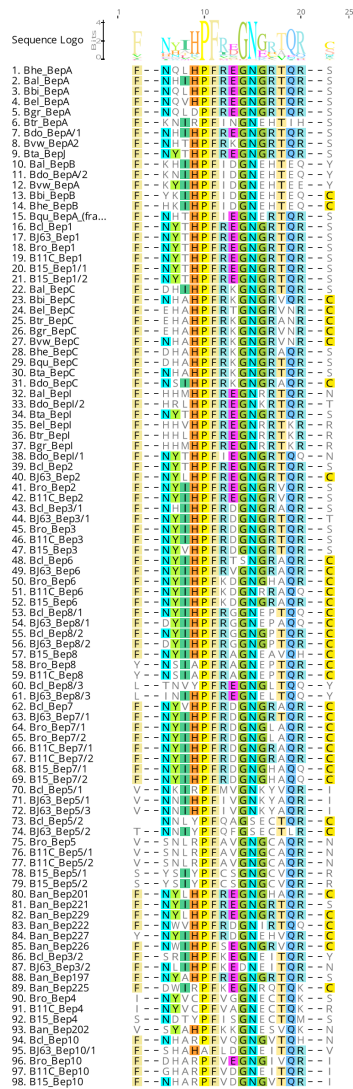
Figure S3. Excerpt of the overall BepFic sequence alignment covering

***Bhe*_BepA active site.** The non-contiguous segments encompass *Bhe*_BepA

residues F113, N156-R170, and S198, see also Fig. 6a. At the right, a green

rectangle indicates a canonical FIC motif, a red rectangle an E in position G2 (groups

b and c, respectively, in Fig. 6b).

3

```
F
N    0.193
H    0.581    0.202
G2   0.446    0.412    0.284
E   -0.382   -0.303   -0.233   -0.784
R1   0.539    0.359    0.543    0.557   -0.589
Q    0.005    0.517    0.060    0.350   -0.212    0.113
R2   0.089    0.193    0.116    0.446   -0.456    0.407    0.200
S    0.539    0.511    0.511    0.450   -0.325    0.544    0.170    0.181
     F        N        H        G2       E        R1       Q        R2       S
```

Figure S4. Correlation (see Methods) between selected residues of the non-contiguous active site motif shown in Fig. 6.
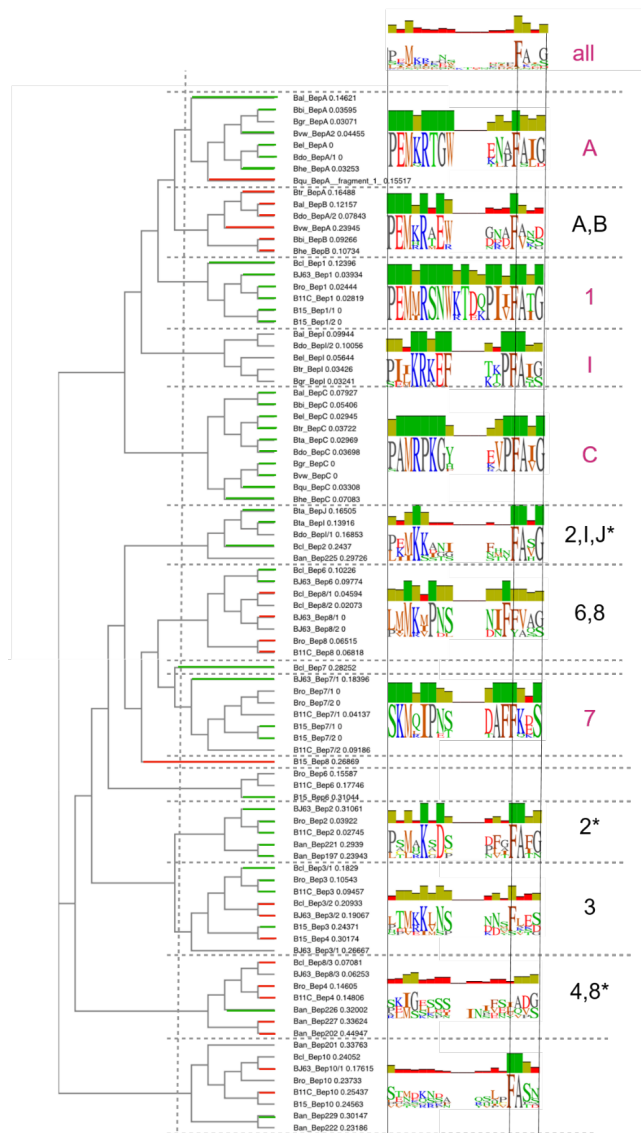
4

Figure S5. Cladogram derived from all BepFIC flap sequences used in this study (with the exception of the flap-less Bep5 sequences). Branches are labeled according to the sub-clades of their members, asterisks indicate the additional presence of *Ban* members. Horizontal lines colored in green or red indicate Beps with canonical FIC motif or with a glutamate in position G2 (groups b and d, respectively, in Fig. 6b). Names in magenta indicate the major branches as defined in the legend to Fig. 8.
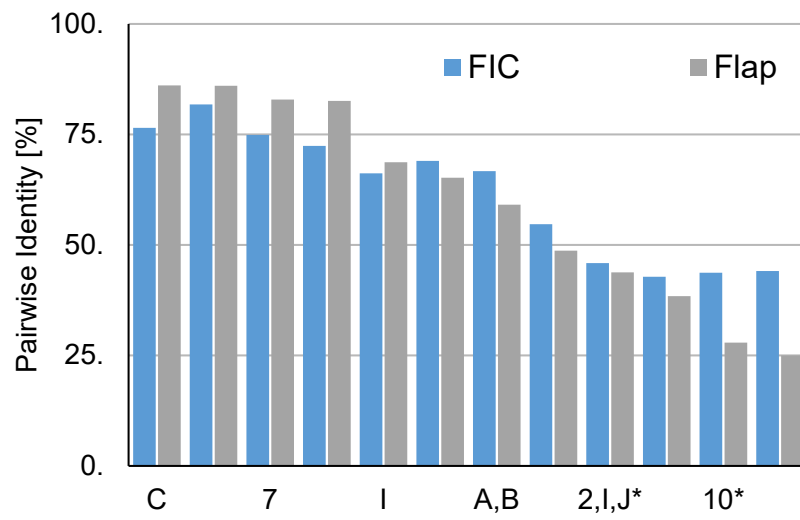
5

Figure S6. Overall (blue) and flap (green) sequence identity of the flap branches as shown in the cladogram of Fig. S5

6
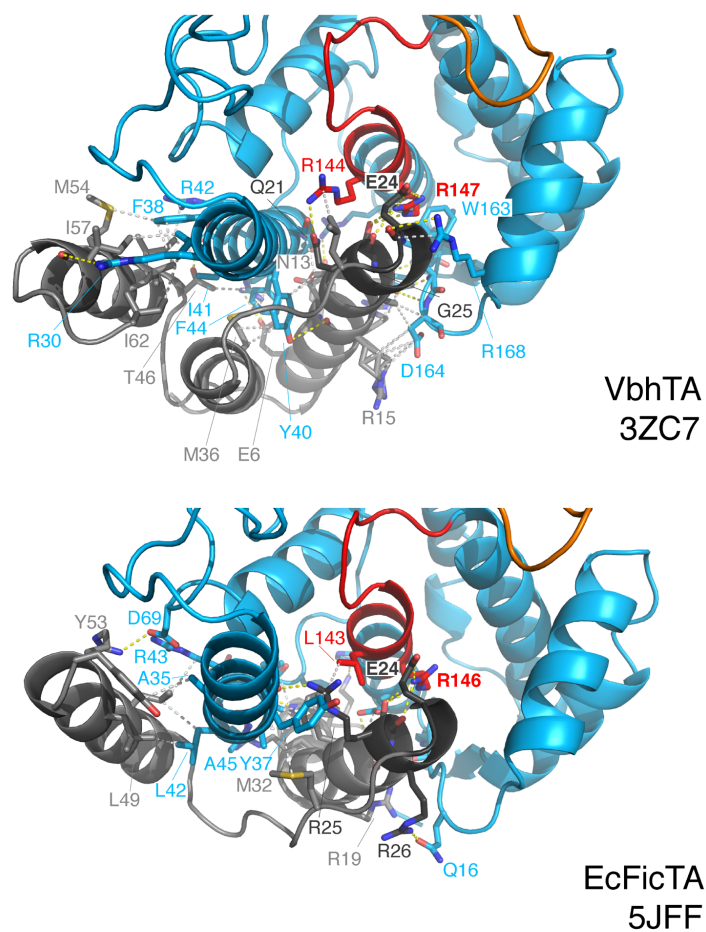
Figure S7. FicTA toxin/antitoxin interaction. The interacting residues between the toxin (blue) and the antitoxin (grey) are shown in light grey indicating apolar interactions and yellow indicating hydrogen bonds. The FIC motif is red and the flap is in orange. (A) VbhTA complex (PBD: 3ZC7). (B) *Ec*_FicAT complex (PBD: 5JFF).

7

Table S1. List of species abbreviations and genome references.

3-letter code	Organism	NCBI taxonomy ID	Accession number(s)
B11C	*Bartonella* sp. 1-1C	NCBI:txid515256	CP019489
BA13	*Bartonella* sp. A1379B	NCBI:txid1933910	CP019780
B15	*Bartonella* sp. AR 15-3	NCBI:txid545617	MUYE00000000
Bal	*Bartonella alsatica* IBS 382	NCBI:txid1094551	AIME01000001.1 - AIME01000021.1
Ban	*Bartonella ancashensis* 20.00	NCBI:txid1318743	KY583505
BbK	*Bartonella bacilliformis* KC583	NCBI:txid360095	NC_008783
BbV	*Bartonella bacilliformis* Ver075	NCBI:txid1293904	ASIV01000001 - ASIV01000010
Bbi	*Bartonella birtlesii* IBS 325	NCBI:txid1095900	AKIP01000001_1.1 - AKIP01000035_1.35
Bbo	*Bartonella bovis* 91-4	NCBI:txid1094491	AGWA01000001 - AGWA01000019
Bcd	*Bartonella* sp. CDCskunk	NCBI:txid1933905	CP019782
Bcl	*Bartonella clarridgeiae* 73	NCBI:txid696125	NC_014932.1
Bco	*Bartonella* sp. Coyote22sub2	NCBI:txid1933911	CP01978
Bdo	*Bartonella doshiae* NCTC 12862	NCBI:txid1094553	AILV01000001 - AILV01000025
Bel	*Bartonella elizabethae* F9251	NCBI:txid109455	AIMF01000001.1 - AIMF01000049.1
Bgr	*Bartonella grahamii* as4aup	NCBI:txid634504	NC_012846 + NC_012847
Bhe	*Bartonella henselae* Houston-1	NCBI:txid283166	NC_005956
Bhd	*Bartonella* sp. Hoopa Dog 114	n.a.	CP019784
Bhf	*Bartonella* sp. Hoopa Fox 11B	n.a.	CP019783
BJ15	*Bartonella* sp. JB15	NCBI:txid1933906	CP019787
BJ63	*Bartonella* sp. JB63	NCBI:txid1933907	CP019788
Bqu	*Bartonella quintana* Toulouse	NCBI:txid283165	NC_005955
BR60	*Bartonella* sp. Raccoon60	NCBI:txid1933912	CP019786
Bro	*Bartonella rochalimae* ATCC BAA-1498	NCBI:txid685782	FN645455 - FN645467
Bsb	*Bartonella schoenbuchensis* R1	NCBI:txid687861	CP019789 + CP019790
BW12	*Bartonella* sp. SikaDeer WD12.1	n.a.	MUBG00000000
BW16	*Bartonella* sp. SikaDeer WD16.2	n.a.	CP01978
Bta	*Bartonella taylorii* IBS 296	n.a.	MUYW00000000
Btr	*Bartonella tribocorum* CIP 105476	NCBI:txid382640	NC_010160 + NC_010161
Bvw	*Bartonella vinsonii* subsp. *berkhoffii* strain Winnie	NCBI:txid1094497	CP003124

8

Table S2. List of FIC-BID Beps used in this study.

Name	UNIPROT
Bal_BepA	J1IX26
Bal_BepB	J0PTT5
Bal_BepC	J1IY69
Bal_BepI	J1IY65
Ban_Bep197	A0A0M3T2K6
Ban_Bep201	A0A0M3T2M9
Ban_Bep202	A0A0M5KSB7
Ban_Bep221	A0A0M3T2N3
Ban_Bep222	A0A0M5KSB8
Ban_Bep225	A0A0M4LIT1
Ban_Bep226	A0A1V0PNG5
Ban_Bep227	A0A0M3T2L2
Ban_Bep229	A0A0M4LHM1
Bbi_BepA	UPI000379BB9C
Bbi_BepB	UPI000367D6D4
Bbi_BepC	UPI0003824625
Bcl_Bep1	E6YFW2
Bcl_Bep2	E6YIF3
Bcl_Bep3/1	E6YHI2
Bcl_Bep3/2	E6YHH2
Bcl_Bep5/1	E6YGF5
Bcl_Bep5/2	E6YGD8
Bcl_Bep6	E6YHH5
Bcl_Bep7	E6YHI3
Bcl_Bep8/1	E6YHH4
Bcl_Bep8/2	E6YHH3
Bcl_Bep8/3	E6YHH8
Bcl_Bep10	E6YIM4
Bdo_BepA/1	J0YV99
Bdo_BepA/2	J0YVD3
Bdo_BepC	J1J5M3
Bdo_BepI/1	J1J4K1
Bdo_BepI/2	J0Q230
Bel_BepA	J0R905
Bel_BepC	J0R175
Bel_BepI	J0ZTN1
Bgr_BepA	C6AES4
Bgr_BepC	C6AES7
Bgr_BepI	C6AET2
Bhe_BepA	Q5QT06
Bhe_BepB	Q5QT04
Bhe_BepC	Q5QT03
Bqu_BepA (fragment 1)	no accession in this strain
Bqu_BepA (fragment 2)	no accession in this strain
Bqu_BepC	A0A0H3LV04
Bro_Bep1	E6YJU0
Bro_Bep2	E6YMI0
Bro_Bep3	E6YLF6
Bro_Bep4	E6YLB1
Bro_Bep5	E6YKB8

Name	UNIPROT
Bro_Bep6	E6YLF3
Bro_Bep7/1	E6YLF1
Bro_Bep7/2	E6YLF7
Bro_Bep8	E6YLF2
Bro_Bep10	E6YMR3
B11C_Bep1	E6YTB2
B11C_Bep2	E6YW78
B11C_Bep3	E6YW82
B11C_Bep4	E6YV21
B11C_Bep5/1	E6YVB0
B11C_Bep5/2	E6YU24
B11C_Bep6	E6YV78
B11C_Bep7/1	E6YV76
B11C_Bep7/2	E6YV83
B11C_Bep8	E6YV77
B11C_Bep10	E6YWF1
B15_Bep1/1	E6YNY8
B15_Bep1/2	E6YNY8
B15_Bep3	E6YQP7
B15_Bep4	E6YQI1
B15_Bep5/1	E6YQT0
B15_Bep5/2	E6YPK5
B15_Bep6	E6YQQ0
B15_Bep7/1	A0A1T3BS91
B15_Bep7/2	E6YQQ2
B15_Bep8	E6YQQ1
B15_Bep10	E6YS53
BJ63_Bep1	UPI0009995D5F
BJ63_Bep2	UPI00099A7C8C
BJ63_Bep3/1	UPI00099B008B
BJ63_Bep3/2	UPI00099A7C8A
BJ63_Bep5/1	UPI0009C1F7B7
BJ63_Bep5/2	UPI00099A9131
BJ63_Bep5/3	UPI0009C348C6
BJ63_Bep6	UPI00099B1F53
BJ63_Bep7/1	UPI0009999BB7
BJ63_Bep8/1	UPI00099A6FF8
BJ63_Bep8/2	UPI000999E53B
BJ63_Bep8/3	UPI00099AB2FD
BJ63_Bep10/1	UPI000999CC63
Bta_BepC	UPI00026E5F09
Bta_BepI	UPI00026E5F06
Bta_BepJ	UPI00026E5F0C
Btr_BepA	A9IWP4
Btr_BepC	A9IWP7
Btr_BepI	A9IWQ5
Bvw_BepA	N6UQE3
Bvw_BepA2	N6VKZ4
Bvw_BepC	N6UWX1

9

Table S3. Data collection and refinement statistics.

	*Bhe*_BepA /BiaA_{L4}	*Bqu*_BepC/ ADP	*Bqu*_BepC	*Btr*_BepC/ AMPPNP	*Bcl*_Bep1	*Bcl*_Bep5	*Bsp15*_Bep8	*B11C*_Bep8
Data collection								
Space group	P $2_1 2_1 2$	P $2_1 2_1 2_1$	P 2_1	P $2_1 2_1 2$	P $2_1 2_1 2$	P $2_1 2_1 2$	P 2_1	P 2_1
a, b, c (Å)	48.03, 56.15, 136.60	40.45, 64.07, 97.60	57.56, 43.63, 88.76	59.33, 92.02, 45.84	73.19, 97.85, 49.11	99.27, 122.86, 143.71	40.24, 47.48, 67.97	55.81, 324.16, 86.13
β (⁰)	90.00	90.00	91.53	90.00	90.00	90.00	105.64	109.24
Resolution (Å)[*]	2.32 (2.40-2.32)	1.55 (1.59-1.59)	2.00 (2.05-2.00)	1.70 (1.74-1.70)	1.90 (1.95-1.90)	2.95 (3.03-2.95)	1.85 (1.90-1.85)	2.35 (2.41-2.35)
Unique reflections	121245	37518	30109	28113	28414	37608	21213	118369
Completeness (%)	94 (84)	99.6 (99.3)	99.7 (99.9)	99.2 (92.1)	99.6 (99.7)	99.7 (100.0)	99.6 (99.8)	99.0 (99.2)
I/σ(I)	18.1 (2.6)	22.7 (3.5)	8.2 (3.1)	20.9 (2.1)	27.8 (3.7)	17.4 (3.7)	18.9 (2.3)	13.3 (2.4)
Redundancy	7.7 (6.6)	6.1 (6.0)	3.7 (3.4)	4.5 (2.1)	5.2 (5.3)	6.1 (6.3)	4.1 (3.8)	3.7 (3.9)
R$^a_{sym}$ (%)	9.5 (66.0)	4.6 (49.0)	11.4 (38.6)	5.0 (46.5)	3.6 (52.2)	7.5 (51.6)	5.9 (55.2)	8.0 (56.9)
Refinement								
R$^b_{work}$/R$^c_{free}$ (%)	26.2 (32.8)	17.2 (19.4)	18.3 (23.5)	16.4 (19.3)	17.2 (21.2)	23.5 (27.9)	16.4 (21.1)	24.2 (25.9)
Rmsd from ideal values								
Bond length (Å)	0.013	0.010	0.017	0.006	0.013	0.002	0.014	0.009
Bond angles (°)	1.2	1.5	1.8	1.0	1.5	0.54	1.5	1.2
No. of atoms								
Protein	2756	1770	3516	1661	2185	8683	1839	16440
Ligands	1	28	14	45	8	12	4	8
Metals	1	1	0	0	0	0	0	0
Water	83	257	325	231	161	8	271	267
Average B-factor (Å^2)								
Protein	48.2	19.9	18.8	19.6	34.5	82.1	23.1	50.8
Ligands	62.2	28.0	31.8	32.2	58.4	89.7	32.3	60.2
Metals	62.2	27.15	NA	NA	NA	NA	NA	NA
Water	45.4	41.3	26.3	32.6	43.4	59.8	31.9	36.7
Ramachandran statistics (%)								
Favored regions	98	100	99.5	100	100	98	99	98
Allowed regions	1.8	0	0.5	0	0	2	1	2
Disallowed regions	0.29	0	.	0	0	0	0	1
PDB ID	5NH2	4N67	4LU4	4WGJ	4NPS	4XI8	4M16	4PY3

[*] The values recorded in parentheses are those for the highest resolution shell
[a] R_{sym} = $\Sigma |I - \langle I \rangle| / \Sigma |I|$, where I is the observed intensity and $\langle I \rangle$ is the average intensity of several symmetry-related observations.
[b] R_{work} = $\Sigma ||Fo| - |Fc|| / \Sigma |Fo|$, where Fo and Fc are the observed and calculated structure factors, respectively.
[c] R_{free} = $\Sigma ||Fo| - |Fc|| / \Sigma |Fo|$ for 7% of the data not used at any stage of the structural refinement.

10

3.2 Research article II

DeAMPylation mediated by the toxin/antitoxin complex VbhT/VbhA of *Bartonella schoenbuchensis*

Stefanie Tamegger[1,2], Christoph Dehio[2] und Tilman Schirmer[1]

[1] Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, Basel, Switzerland

[2] Focal Area Infection Biology, Biozentrum, University of Basel, Basel, Switzerland

3.2.1 Statement of my own contributions

My contribution to research article II was expression and purification of *N. meningitidis* NmFic$_{\Delta8}$, *E. coli* GyrB43 and *B. schoenbuchensis* VbhT(FIC)/VbhA and VbhT(FIC)/VbhA$_{E24G}$. An online assay for nucleotide quantification was previously developed by the group and therefore I adjusted the assay for a Gilson HPLC system to observe toxin induced AMPylation and deAMPylation, showing that GyrB43 gets modified and unmodified by the class I toxin-antitoxin complex VbhT/VbhA.

I participated in designing and performing of the experiments, data analysis and writing of the manuscript.

3.2.2 "DeAMPylation mediated by the toxin/antitoxin complex VbhT/VbhA of *Bartonella schoenbuchensis*"

DeAMPylation mediated by the toxin/antitoxin complex VbhT/VbhA of *Bartonella schoenbuchensis*

Stefanie Tamegger[1,2], Christoph Dehio[2] und Tilman Schirmer[1]

[1] Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, Basel, Switzerland

[2] Focal Area Infection Biology, Biozentrum, University of Basel, Basel, Switzerland

Corresponding author: tilman.schirmer@unibas.ch

Abstract

Proteins containing a conserved FIC (filamentation induced by cyclic AMP) domain modify the function of target proteins via post-translational modifications such as the transfer of AMP called AMPylation. Intensive investigations of several Fic proteins revealed the regulatory mechanism behind AMPylation. The presence of an inhibitory signature motif displayed suppression of toxicity of Fic proteins leading to classification in three classes. Recent studies showed that a representative of class II (FICD) and class III (EfFic) not only perform AMPylation, but act as a bifunctional enzyme mediating deAMPylation. Furthermore it was revealed that the highly conserved glutamate of the inhibition motif plays a significant role in those two reactions.

Here, we use a recently developed nucleotide quantification assay as a method to acquire enzymatic progress curves. The class I FicT VbhT in complex with its cognate antitoxin VbhA catalyzes both AMPylation and deAMPylation of a N-terminal 43kDa fragment of DNA Gyrase subunit B (GyrB43) resulting at steady state in ATP to AMP turnover with accumulation of AMP-GyrB43 intermediate. Mutation of the inhibitory glutamate, which was shown to play a significant role in AMPylation, reveals enhanced AMPylation and deAMPylation indicating a different function of the amino acid compared to class II and class III Fic proteins.

Keywords: AMPylation, deAMPylation, FIC proteins, PTM, toxin-antitoxin, VbhT/VbhA, GyrB43, NmFic

Introduction

Proteins containing a highly conserved FIC (filamentation induced by cyclic AMP (cAMP)) domain can be found in all domains of life. Besides other functions many members of this family were shown to transfer the AMP moiety of ATP onto a threonine or tyrosine residue of a respective target protein. This covalent modification, called AMPylation, then causes different effects on host cells.

A highly conserved stretch of twelve amino acids (HxFx(D/E)GNGRxxR) was shown to be essential for the catalysis of AMPylation [8] [22]. High resolution crystal structures of several Fic proteins revealed six α-helices (α1-α6) as core of the domain, but they can be also decorated with additional helices like in NmFic and HpFic [8] [13]. The connection between helices α4 and α5 and the α5 N-cap forms the catalytic center consisting of the mentioned signature motif [8] [13]. The solved structure of the Fic domain of IbpA (IbpAFic2) in complex with its AMPylated target Cdc42 gave first insights into the catalytic mechanism, which showed a nucleophilic attack of the target hydroxyl side chain onto the α-phosphate of the ATP substrate. The histidine of the signature motif is essential for deprotonating the incoming hydroxyl group and mutation of that amino acid showed suppression of the toxic activity [22] [23] [11]. The glutamate or aspartate is crucial for coordination of the cation (Mg^{2+} or Mn^{2+}) to position the ATP in a suitable way to allow in-line attack. The γ-phosphate of the substrate is interacting with the second arginine of the signature motif as shown in the left panel in figure 1 (**Fig. 1**, left panel) [5] [13].

The crystal structure of the Fic toxin (FicT) VbhT from *Bartonella schoenbuchensis* revealed how the toxin mediated AMPylation reaction of the DNA gyrase subunit B (GyrB). VbhT has a N-terminal Fic domain including the conserved signature motif and experiments revealed strong growth defects which were absent when the histidine was mutated. Toxicity was also abolished when a small protein was co-expressed with the toxin. This protein is encoded upstream of VbhT and interaction leads to healthy cell growth, therefore the protein acts as VbhT's cognate antitoxin called VbhA. The high-resolution crystal structure of the VbhT/VbhA complex (PDB: 3SHG) displayed three α-helices belonging to the antitoxin, which are tightly wrapped around the α1 helix of the toxin [5]. The C-terminal inhibitory α-helix (α_{inh}) of the antitoxin is occupying a position close to the VbhT active site with the side chain of a highly conserved glutamate (E24) directly pointing into the catalytic center of the toxin. Bioinformatic analysis and homology modelling

showed that the signature motif (S/T)xxxE(G/N) of the α_{inh} is highly conserved and found in several putative antitoxin sequences. This motif can be upstream of the toxin like in VbhT/VbhA but can also be part of the Fic core domain itself either at the N-terminus before α1 or at the C-terminus following α7. Thus, most Fic proteins can be divided into three classes based on the location of the α_{inh}. Class I Fic proteins consist of a toxin, which is interacting with its cognate antitoxin like the VbhT/VbhA complex. Class II and class III Fic proteins have the conserved α_{inh} either at the N-terminal part of the toxin like the human Fic protein FICD or at the C-terminus observed for the *Neisseria meningitidis* NmFic [5]. The inhibition motif causes suppression of the toxic effect due to the interaction of the glutamate with the second arginine (R) of the Fic motif. In presence of ATP the γ-phosphate competes with the glutamate for binding to the arginine leading to re-orientation of the α-phosphate (**Fig. 1**, middle panel) [5]. This orientation does not allow in-line attack of the incoming target hydroxyl side chain and therefore makes an AMPylation reaction impossible in presence of the α_{inh}. Structures of Fic proteins, containing a mutation of the inhibitory glutamate, showed the importance of the (D/E)GNGRxxR sequence of the Fic motif, which enables catalytically competent ATP binding due to a large hydrogen-bonding network resulting in favorable orientation of the α-phosphate of ATP [6].

The human FIC domain containing protein FICD was identified as AMP transferase, which modifies the Hsp70 chaperone binding immunoglobulin protein BiP in the endoplasmic reticulum (ER) and is controlled by an α_{inh} on the N-terminus of the protein (class II Fic protein) [5] [17]. Recently it was revealed that wildtype FICD is bifunctional mediating AMPylation and deAMPylation of BiP and the major role of glutamate 234 (E234) in the reactions [18]. The crystal structure of a deAMPylation complex structure suggests that E234 aligns the water molecule for in-line attack of the α-phosphate bound to the threonine side chain and might act as a catalytic base [14]. Previous studies revealed an increase in flexibility of the side chain induced by FICD monomerization allowing ATP binding in presence of Mg^{2+}, accompanied by decrease of the deAMPylation k_{cat} [15] [14]. Lately it was revealed that the presence of Ca^{2+} modulates the deAMPylation by wildtype FICD due to competition with Mg^{2+} [20].

Bifunctional activity was also revealed for the class III Fic protein EfFic from *Enterococcus faecalis* using the same active site. Like FICD both reactions can be regulated by the presence of a second metal competing with the catalytic metal. Ca^{2+} is competing with Mg^{2+} in productive ATP-binding resulting in inhibition of AMPylation [20].

In this study we show that the FIC domain of the class I Fic protein VbhT (VbhT(FIC)) in complex with its cognate antitoxin VbhA has bifunctional activity mediating AMPylation and deAMPylation of its target GyrB43 (Figure 1, right panel). We used an online nucleotide quantification assay to obtain the kinetic characteristics behind the two reactions, which suggest a new role for the antitoxin and its inhibitory glutamate in the deAMPylation process of target proteins.

Material and Methods

Protein expression and purification

The full-length *vbhA* gene and part of the *vbhT* gene encoding the FIC domain (1-198, His$_6$-tagged) from was cloned into the pRSF-Duet1 vector to obtain pFVS0011 (VbhT(FIC)/VbhA). PFVS0065 (VbhT(FIC)/VbhA$_{E24G}$) was generated by introducing a two-base-pair mutation into pFVS0011. The *NmFic* gene from *Neisseria meningitidis* was designed with an N-terminal His$_6$-tag and coding from the amino acid residues 11-167 to acquire pFVS0016 (NmFic$_{\Delta8}$) [5]. The gene of *E. coli* wildtype *gyrB* spanning the residues 1-392 was designed with an N-terminal His$_6$-tag and cloned into the pRSFDuet-1 vector resulting in a N-terminal 43kDa fragment of GyrB (GyrB43) containing the ATPase and transducer domain [1] [19]. Detailed information is listed in **Supplemental table 1 (S1)**.

The vectors pFVS0011, pFVS0016, pFVS0065 and pFVS0109 were transformed into *E. coli* BL21-AI cells (Invitrogen). Bacterial cells were incubated overnight (o/n) at 37°C and 120rpm in 50ml LB medium supplemented with 50µg/ml kanamycin and 1% glucose. Cells of o/n cultures were harvested by centrifugation (4'000rpm (Beckman Coulter Type 70 Ti Titanium Rotor)/20 min/room temperature (RT)) and pellets containing pFVS0011 or pFVS0109 were resuspended in 1L terrific broth medium supplemented with 50µg/ml kanamycin. Pellets of cells containing pFVS0016 or pFVS0065 were resuspended in 750ml LB medium supplemented with 50µg/ml kanamycin and 1% glucose, and incubated o/n at 37°C with 100rpm. Bacterial pellets were harvested (6'000xg/6 min/RT) and resuspended in 1L terrific broth medium containing 50µg/ml kanamycin. Protein expression was induced at 23°C with 0.1% w/v arabinose (Sigma-Aldrich, Germany) and 0.1mM isopropyl-β-D-thiogalactopyranoside (IPTG) (AppliChem GmbH, Darmstadt, Germany) for 24 hours at 100rpm. Bacterial cells containing the overexpressed proteins were harvested by centrifugation at 6'000xg for 6 minutes at 4°C.

Bacterial pellets were resuspended in lysis buffer (20mM Tris-HCl pH7.5, 250mM NaCl, 20mM Imidazole, protease inhibitor cocktail and DNAse I) and disrupted using French press (Thermo Fisher). Cell debris were pelleted via ultracentrifugation at 45'000 rpm for 40 minutes at 4°C. The cleared lysate was applied to a HisTrap HP column (GE Healthcare) pre-equilibrated with wash buffer (20mM Tris-HCl pH7.5, 250mM NaCl, 20mM imidazole). After a washing step the proteins were eluted with a linear gradient of elution buffer (20mM Tris-HCl, 250mM NaCl, 500mM imidazole). The proteins were concentrated and then applied to a HiLoad 16/600 Superdex 75pg size exclusion column (GE Healthcare) pre-equilibrated with SEC buffer (10mM Tris-HCl pH7.5, 100mM NaCl). After another concentration step the pure proteins were used for nucleotide quantification assays.

Due to difficulties with protein solubility the FIC domain (1-198) of VbhT (VbhT(FIC)) was co-expressed and purified with the cognate antitoxin in the wildtype form (VbhT(FIC)/VbhA) or its inhibition-relieved form (VbhT(FIC)/VbhA$_{E24G}$).

Generation of AMPylated GyrB43

To obtain AMPylated GyrB43 (AMP-GyrB43) purified unmodified GyrB43 was incubated with purified NmFic$_{\Delta 8}$, 1mM ATP and novobiocin in AMPylation buffer (50mM Tris-HCl pH8.0, 150mM NaCl, 5mM MgCl$_2$, 5mM DTT) for 4 hours at room temperature. The reaction mix was applied to an ion exchange Resource Q column (GE Healthcare) pre-equilibrated with Buffer A (50mM Tris-HCl pH8.5) and modified GyrB43 was eluted with a linear gradient of Buffer B (50mM Tris-HCl pH8.5, 1M ammonium sulfate). The fractions containing the AMP-GyrB43 were concentrated and loaded on a HisTrap HP column (GE Healthcare) pre-equilibrated with AMPylation buffer to remove nucleotide contamination. AMP-GyrB43 was obtained with a gradient of elution AMPylation buffer (50mM Tris-HCl pH8.0, 150mM NaCl, 5mM MgCl$_2$, 5mM DTT, 500mM Imidazole). Presumable AMP-GyrB43 was concentrated and immediately used for nucleotide quantification assays.

Online ion-exchange chromatography

AMPylation and deAMPylation of GyrB43 was monitored via quantification of nucleotides and proteins at different wavelengths using an online ion exchange chromatography (oIEC) assay on a Gilson HPLC system at room temperature [4] [3]. This system was equipped with a GX-241 II Liquid Handler without pump (Gilson), a GX Syringe Pump (Gilson), a 322 HPLC Pump (Gilson), a 172 Diode Array Detector (Gilson) and a FC 203B Fraction Collector (Gilson). Buffers A (50mM Tris-HCl pH8.5) and B (50mM Tris-HCl pH8.5, 1M ammonium sulfate) were used as mobile phases A and B, respectively, and a 1ml ion exchange Resource Q column (GE Healthcare) was equilibrated with Buffer A. For AMPylation assays purified GyrB43 and novobiocin to inhibit GyrB43's ATP hydrolysis activity, were added to AMPylation buffer.

The reaction was started at t=0min by mixing GyrB43 with 1mM ATP and either VbhT(FIC)/VbhA, VbhT(FIC)/VbhA$_{E24G}$ or NmFic$_{\Delta 8}$. In addition, deAMPylation experiments were performed by mixing AMP-GyrB43 with VbhT(FIC)/VbhA or VbhT/A$_{E24G}$ at t=0min. Subsequently, every 10 minutes 50µl of the reaction were repetitively applied on the column by an autosampler. After each injection a short washing step with Buffer A was applied. The bound nucleotides and proteins were eluted with a linear gradient to 30% Buffer B and chromatograms were acquired at several wavelengths (253nm, 259nm and 280nm). Chromatograms were further processed using proFit7 (QuantumSoft, Uetikon am See, Switzerland) by fitting the peaks corresponding to the reaction compounds to Gaussian functions. Resulting peak areas were converted to concentrations using a scale factor obtained from an ATP calibration curve (**Fig. S1**). The thus obtained progress curves of the reaction compounds were then fitted to a kinetic model (enzymatic group transfer) using the Levenberg-Marquardt algorithm or the Monte-Carlo algorithm as implemented in proFit7.

Results

Kinetic modelling of concurrent target AMPylation/deAMPylation

Previous publications revealed that class II and class III Fic enzymes have bifunctional activity and mediate AMPylation and deAMPylation of target proteins or surrogate targets [18] [20]. A kinetic scheme was set-up to model the enzymatic behavior

of a bifunctional enzyme with antagonistic activities such as FICD, in which two different reactions are catalyzed (**Fig. 2A**, left panel) [18]. The first reaction is an enzyme catalyzed group transfer with S1 as ATP and S2 as target (1).
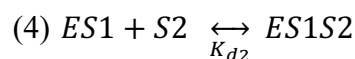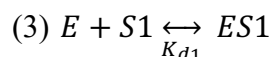
$$(1)\ S1 + S2\ \rightarrow P1 + P2$$
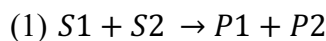
The second reaction is a cleavage (hydrolytic reaction) with P3 as AMP (2).
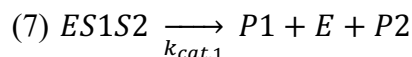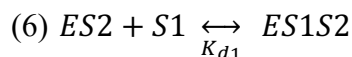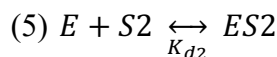
$$(2)\ P2\ \rightarrow S2 + P3$$
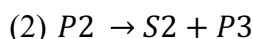
Based on this simple scheme a Michaelis-Menten model was developed to obtain the enzymatic characteristics for a bifunctional Fic enzyme. In an AMPylation reaction the enzyme (E) mediates the transfer of an AMP moiety from the ATP, depicted as substrate 2 (S2), onto the target (S1) resulting in product 2 (P2) accompanied by release of product 1 (P1) (7).

The assumption is that both substrates bind independent from each other to the enzyme and the affinity of the substrate to the enzyme does not have any impact (3-6).

$$(3)\ E + S1 \underset{K_{d1}}{\longleftrightarrow} ES1$$

$$(4)\ ES1 + S2 \underset{K_{d2}}{\longleftrightarrow} ES1S2$$

or

$$(5)\ E + S2 \underset{K_{d2}}{\longleftrightarrow} ES2$$

$$(6)\ ES2 + S1 \underset{K_{d1}}{\longleftrightarrow} ES1S2$$

$$(7)\ ES1S2 \underset{k_{cat,1}}{\longrightarrow} P1 + E + P2$$

The calculation of the deAMPylation reaction only needs the enzyme E and the AMPylated target P2 to result in AMP as product 3 (P3) as depicted in (8) + (9).

$$(8)\ E + P2 \underset{K_{d3}}{\longleftrightarrow} EP2$$
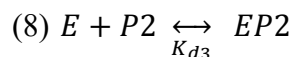
The complex EP2 is assumed to be immediately converted to E and P3.

$$(9)\ EP2 \underset{k_{cat,2}}{\longrightarrow} E + P3$$
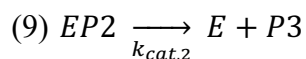
P2 functions as intermediate and substrate for deAMPylation mediated by the same E. This reaction is proposed to take place due to the hydrophilic, in-line attack of a water molecule, which is coordinated by a magnesium ion. The divalent cation is held in place by the conserved glutamate or aspartate of the FIC motif. The reaction leads to the release of bound AMP, which is depicted as product 3 (P3) in the scheme, and the unmodified target S1 is available again for AMPylation (**Fig. 2A**). The right panel shows the same

reaction mechanism but includes the reaction species, which are measured in kinetic assays in this study.

We have simulated four different experiments using the shown equations in which either one or both reactions are happening (**Fig. 2B-E**). The same concentrations for enzyme (20μM), ATP (500μM) and target (80μM) were used in all simulations. The four modelled time courses comprise the time in seconds on the x-axis and the peak area in milli absorbance units (mAU)*milliliter (ml) on the y-axis. ATP is displayed in turquoise, the modified and unmodified target GyrB43 is red, and AMP is shown in yellow.

The first simulation is an AMPylation reaction with $k_{cat,1}$ of $0.07s^{-1}$. The progress curve (**Fig. 2B**) shows how ATP is decreasing to due to the transfer of AMP onto the target. Note that the production of P2 (PPi) is not considered as it cannot be measured. The reaction comes to a static endpoint due to substrate (S1) exhaustion when GyrB43 has been fully AMPylated (**Fig. 2B**). The second simulation (**Fig. 2C**) contains a pool of fully AMPylated target and the enzyme at the start of the reaction, and no ATP. The enzyme mediates deAMPylation with a defined $k_{cat,2}$ of $0.07s^{-1}$ resulting in removal of the AMP moiety and production of free AMP and unmodified target. The reaction reaches its end point once all modified target is converted into a pool of unmodified GyrB43 (**Fig. 2C**).

Simulations with both (AMPylation and deAMPylation) activities are shown in **Figs. 2D** and **2E**. A coupled reaction with $k_{cat,1}$ being $0.07s^{-1}$ and $k_{cat,2}$ is $0.007s^{-1}$ depicts turnover of ATP to AMPylated target and AMP. Modified target is built up at the beginning of the reaction until a steady state is reached with equal AMPylation and deAMPylation rates (**Fig. 2D**). If the AMPylation is the rate-limiting reaction with $k_{cat,1}$ being $0.007s^{-1}$ and deAMPylation having a $k_{cat,2}$ of $0.07s^{-1}$, no intermediate AMPylated target is produced due to the faster deAMPylation (**Fig. 2E**). The produced simulations give an insight about how the reactions might look in the actual experiments.

Analysis of GyrB43 AMPylation by NmFic$_{\Delta 8}$ using online ion-exchange chromatography

AMPylation of target proteins was so far analyzed via end-point assays such as autoradiography using radioactive-labelled ATP to visualize the reaction, but due to several limitations it is not possible to observe the behavior of all reaction compounds at the same time [16] [7]. Online ion-exchange chromatography (oIEC) is a quantitative assay, which

allows to monitor the reaction in real-time being a major advantage compared to autoradiography assays. At defined time points aliquots of a reaction mix containing enzyme, target and substrate are injected onto an anion-exchange column, to which the negative charged compounds can bind. After a short washing step, they are eluted with a linear gradient of ammonium sulfate and with each elution a chromatogram is recorded, which can be further processed to obtain a time curve of the monitored reaction [3] [4].

Using autoradiography Engel *et. al.*, 2012 previously described AMPylation of the 43kDa fragment of GyrB43 by the *Neisseria meningitidis* class III Fic protein NmFic$_{\Delta 8}$, which lacks the complete inhibitory α-helix α$_{inh}$,. Here we used the oIEC assay to test the suitability of the assay and to quantitatively determine the velocity of this reaction and further investigations.

A reaction mix containing GyrB43, ATP and NmFic$_{\Delta 8}$ was loaded onto an anion exchange column. It is known that GyrB has ATPase activity for its essential cell cycle functions [10] [2] [21]. Therefore, the gyrase inhibitor novobiocin was added at sufficient concentration in all experiments [9]. An excerpt overlay of chromatograms displayed in **Fig. 3A** shows three peaks. The peaks can be assigned to AMP, NmFic$_{\Delta 8}$, and GyrB43, as verified by control chromatograms. The GyrB43 peak, however, is not pure, since it contains an overlapping minor ADP contamination (40µM as measured in the control **Fig. S2D**) Due to the high concentration of ATP used in the experiment traces of AMP and ADP could be found in the chromatograms. Modified and unmodified GyrB43 are co-eluting and therefore they are labeled as GyrB43* in all presented chromatograms and time courses. Furthermore, the modification is accompanied by a shift to the right of the target peak as shown in the overlay at 259nm (**Fig. S2A**). The chromatograms after processing using Gaussian peak calculations are depicted in **Fig. 3B** showing the time course of the reaction including all detected reaction species. Each dot represents an injection and elution of the reaction mix. The sum over all measured peak areas is portrayed in grey and should be stable during the reaction to guarantee proper injection and elution of the sample mix. The overall progress curve at 280nm is shown in **Fig. S2B**. After deconvolution of the peak areas of unmodified GyrB43 (GyrB43) and modified GyrB43 (AMP-GyrB43) both reaction species were depicted in a time course containing both wavelengths. The progress curve shows how GyrB43 is fully turned into AMP-GyrB43 after 4000 seconds (**Fig. 3C**). This process is also reflected in the original progress curve (**Fig. S2C**), which shows a

change in the OD ratio between 259nm and 280nm. The plot also demonstrates that no AMP is produced during the reaction.

This experiment showed that oIEC can be successfully used to observe AMPylation of GyrB43 at room temperature in real-time. The enzyme $NmFic_{\Delta 8}$ only catalyzed AMPylation but not deAMPylation of GyrB43 as no free AMP was detected. The data points were fitted using the mentioned model (**Fig 2**). The used conditions and obtained parameters are listed in Table 1 and Table 2, and $k_{cat,1}$ was calculated with $(0.77 \pm 0.040)s^{-1}$ using assumed K_{d1} and K_{d2}, which will be further reviewed.

Concurrent AMPylation and deAMPylation by VbhT(FIC)/VbhA

After validating oIEC as method for monitoring AMPylation of GyrB43, the deAMPylation activity of the class I Fic toxin VbhT was investigated. The FIC domain comprising the residues 1 to 198 of VbhT was co-expressed with its cognate antitoxin VbhA, which contains a conserved inhibition motif.

Injection of a sample mix containing the VbhT(FIC)/VbhA toxin/antitoxin complex, GyrB43 and ATP revealed similar retention times for the VbhT(FIC)/VbhA complex and GyrB43*. Therefore, all three proteins were labelled as protein sum in the time course measured at 259nm displayed in the left panel in **Fig. 4A**. The progress curve showed a slightly decrease of ATP concentration and an increase of the protein sum indicating that the target was getting modified (**Fig4A**, left panel), depicted also in a more detailed view of GyrB43* (**Fig. 4A**, right panel). The peak area of GyrB43* was increasing at 259nm, while the increase at 280nm was only small, due to the higher absorbance of the added AMP at 259nm. The progression curve also shows the accumulation of AMP, which changes into a linear phase after a delay of about 2000 seconds. This indicates deAMPylating activity and is consistent with the slow buildup of the intermediate (AMP-GyrB43) (cf. with the simulation in **Fig. 2D**).

All data points were used for calculations to obtain the kinetic parameters. The conditions are listed in Table 1 and Table 2, and include the concentrations, which were refined during analysis. The k_{cat} for the AMPylation ($k_{cat,1}$) was determined to be $(0.0019 \pm 0.0007)s^{-1}$, while the k_{cat} for the deAMPylation reaction ($k_{cat,2}$) was found to be 6 times higher with $(0.011 \pm 0.0007)s^{-1}$, revealing the AMPylation reaction as the rate-limiting step.

Previous publications have shown that the glutamate in the inhibition motif of the antitoxin plays a significant role in suppressing the activity of the toxin due to interference of the side chain with productive ATP binding [5] [6]. Strong AMPylation by VbhT was shown in autoradiography when the toxin was either expressed alone as full-length protein (VbhT) or in complex with an inhibition-relieved antitoxin (VbhT(FIC)/A$_{E24G}$) [5]. Using oIEC the overall time course at 259nm depicted in **Fig. 4B** (left panel) reveals the expected decrease of ATP due to GyrB43* modification and increase in the protein sum. Significant AMP, again occurring after a certain delay, was detected in the experiment indicating deAMPylation of GyrB43 (**Fig. 4B**, left panel). This is shown at a higher magnification in the right panel. GyrB43* is increasing at both wavelengths, while AMP is already present with only a short delay of less than 1000 seconds (**Fig. 4B**, right panel). Only the first six data points were included in the data processing due to disagreement of the model with the currently used fitting model. K$_{cat,1}$ for the AMPylation reaction was $(0.011 \pm 0.0008)s^{-1}$ and k$_{cat,2}$ is $(0.032 \pm 0.0025)s^{-1}$.

Summarizing, VbhT(FIC)/VbhA revealed minimal AMPylation compared to NmFic$_{\Delta8}$ or the inhibition-relieved T-AT complex, but was deAMPylating GyrB43 with a slight delay. VbhT(FIC)/VbhA$_{E24G}$ showed strong AMPylation and deAMPylation indicating that the glutamate has an inhibitory effect on the second reaction.

DeAMPylation of GyrB43 mediated by VbhT(FIC)/VbhA

Both toxin-antitoxin complexes revealed AMPylation and deAMPylation activity. To investigate the second activity in isolation, AMP-GyrB43 was incubated with the VbhT(FIC)/VbhA variants for the oIEC assay. AMP-GyrB43 was produced by incubation with ATP, novobiocin and NmFic$_{\Delta8}$ for four hours at room temperature and purification using oIEC and affinity chromatography (AC). The AMPylated GyrB43 was then added to VbhT(FIC)/VbhA or VbhT(FIC)/VbhA$_{E24G}$ and chromatograms were acquired over 61 minutes to yield the time course of the deAMPylation reaction (**Fig. 5**).

The left panel of **Fig. 5A** shows the progress curve at 259nm of AMP-GyrB43 mixed with VbhT/VbhA. As expected, the peak area of GyrB43* is decreasing due to the removal of the AMP moiety resulting in a concomitant increase of AMP. This is also depicted in a more detailed comparison between 280nm and 259nm. The peak area of GyrB43* at 259nm (GyrB43*$_{259}$) was decreasing stronger compared to GyrB43*$_{280}$ due to

the removed AMP, which was concurrently increasing at both wavelengths (**Fig. 5A**, right panel). All data points were included in the calculation of the kinetic parameters with the statement that there was still unmodified GyrB43 present, resulting in $k_{cat,2}$ being 0.0066s^{-1}. Unfortunately, the fitting procedure was unstable when using the Levenberg algorithm, such that we had to resort to Monte-Carlo fitting, which does not yield error estimates. The used conditions and parameters are listed in Table 1 and Table 2.

The time course of VbhT(FIC)/VbhA$_{E24G}$ catalyzed deAMPylation obtained at 259nm also showed a decrease of GyrB43* and an increase of AMP (**Fig. 5B**, left panel). This is also depicted in detail in the right panel at 259nm and 280nm (**Fig. 5B**, right panel). The peak areas at 280nm were only decreasing (GyrB43*$_{280}$) or increasing (AMP$_{280}$) slightly, while the peak areas at 259nm revealed strong deAMPylation in the first 1000 seconds. Global fitting of all data points resulted in a $k_{cat,2}$ of 0.045s^{-1}, i.e. a value about seven-fold larger as obtained for the VbhT(FIC)/VbhA reaction.

Discussion

In this study we analyze AMPylation and deAMPylation reactions mediated by a bifunctional class I FicT-AT complex using a nucleotide quantification assay. In contrast to previously used radioactive endpoint assays this setup provides a sensitive and quantitative method to obtain more information about AMPylation of GyrB43 and allows simultaneous monitoring of proteins or nucleotides involved in the reaction [16] [7]. Here the method was validated using the class III Fic protein NmFic$_{\Delta 8}$, which revealed strong AMPylation with a $k_{cat,1}$ of 0.77s^{-1} as depicted in **Table 2** and **Figure 6** (lane 1, light grey) due to the removal of the whole inhibitory α-helix $\alpha_{inh.}$ Additionally, no AMP was detected indicating that no deAMPylation was happening.

No AMPylation of GyrB has been seen hitherto for wildtype VbhT in complex with its antitoxin VbhA (VbhT/VbhA) when measured using autoradiography. Generally, VbhT causes growth inhibition due to target AMPylation, whereas co-expression with its cognate antitoxin suppresses AMPylation and therefore leads to healthy cell growth [5] [7].
Using oIEC revealed small, but significant AMPylation activity of GyrB43 with a $k_{cat,1}$ of 0.0019s^{-1} (**Table 2**) (**Fig. 6**) compared to NmFic$_{\Delta 8}$ or other AMPylating Fic proteins such as Bep1 or IbpA [3] [12]. Additionally, deAMPylation was observed as indicated directly by the delayed production of AMP with a $k_{cat,2}$ of 0.011s^{-1} being 6-fold higher compared to VbhT(FIC)/VbhA's AMPylation reaction (**Fig. 6**). Thus, both activities are of the same

order of magnitude, but AMPylation is the slower reaction. Thus, deAMPylation is delayed until a steady state, with a certain concentration of the AMP-GyrB43 intermediate, is reached.

To analyze solely the deAMPylation activity the pre-AMPylated target, AMP-GyrB43, was employed. VbhT(FIC)/VbhA catalyzed deAMPylation was found to proceed with a rate $k_{cat,2}$ of $0.0066s^{-1}$ (**Fig. 6**), which is similar to the deAMPylation rate of wildtype FICD with $k_{cat,2}$ being $0.01s^{-1}$ [14].

To gain insight into the deAMPylation mechanism, which is hydrolytic reaction, a model of VbhT(FIC)/VbhA in complex with an AMPylated target side chain was built and the water positioned by the Mg^{2+} (**Fig. 7**, green sphere) [18]. The model shows the position of the putative hydrolytic water ready to attack the tyrosine-AMP phosphodiester bond. It is coordinated by interactions with the amides of the N-terminal end of helix $\alpha5$, the ion and surrounding water molecules (**Fig. 7**, pink spheres). Mg^{2+} is known to be crucial for AMPylation since its coordination by glutamic acid (E140 in VbhT) results in finetuning the localization of the α- and β-phosphate of the ATP in the active site [5] [6]. Recent studies revealed that AMPylation and deAMPylation can be regulated by a second metal, which competes with the catalytic metal. Ca^{2+} interacts with ATP and therefore inhibits binding of Mg^{2+} into the active site of EfFic causing suppression of AMPylation activity, while the presence of Ca^{2+} in human FICD leads to upregulation of AMPylation due to competition with Mg^{2+} [20]. A recently published structure of FICD in complex with AMPylated BiP displayed that the phosphate of the AMP-BiP is coordinated by the Mg^{2+} ion. The ion is arranged by the metal-coordinating aspartate while the inhibitory glutamate tightly engages the water molecule within the active site [14].

Early studies revealed that the glutamate of the conserved inhibition motif prevents productive ATP binding due to interaction of the side chain with arginine R147 thus interfering with productive ATP binding [5]. Published data of a mutation from glutamate to glycine revealed a negative effect on cell growth due to AMPylation of GyrB43, which led to its description as inhibition-relieved antitoxin [5] [6] [7]. OIEC experiments using VbhT(FIC)/VbhA$_{E24G}$ revealed as expected AMPylation of GyrB43 with increased turnover numbers compared to wildtype VbhT(FIC)/VbhA (**Table 2**) (**Fig. 6**). $K_{cat,1}$ of VbhT(FIC)/VbhA$_{E24G}$ mediated AMPylation is $0.011s^{-1}$ and 6-fold higher compared to VbhT(FIC)/VbhA. The deAMPylation rate of VbhT(FIC)/VbhA$_{E24G}$ $k_{cat,2}$ is $0.032s^{-1}$, which is 3-fold higher compared to its own AMPylation rate and 7-fold higher to the

isolated VbhT(FIC)/VbhA mediated deAMPylation as depicted in **Figure 6**. In the progress curve shown in **Figure 4B** an increase of AMPylated GyrB43 is detected after 4000 seconds suggesting a second slower AMPylation reaction. Separating the deAMPylation reaction from the AMPylation revealed a $k_{cat,2}$ of $0.045 s^{-1}$, which is similar to the coupled reaction. Studies showed the importance of the glutamate in deAMPylation such as no reaction happened when FICD or EfFic contain a glutamate to glycine mutation, therefore this inhibitory residue is essential in class II and III deAMPylation [18] [20]. Surprisingly, in the case of VbhT(FIC)/VbhA$_{E24G}$ removal of the side chain does not abolish deAMPylation activity but enhances the reaction. The structure model suggests that absence of the side chain leads to a different geometry, which might lead to easier access of the water molecules into the active site.

In conclusion, using a recently developed nucleotide quantification assay to overcome the limits of radioactive endpoint assays we could demonstrate that the class I Fic protein VbhT in complex with its antitoxin VbhA has bifunctional activity with faster deAMPylation than AMPylation. Mutation of the highly conserved inhibitory glutamate led to enhanced deAMPylation suggesting that the side chain of E24 leads to less efficient positioning of the water molecules during the reaction.

Author contributions

S.T. expressed and purified protein constructs and performed the experiments. All authors participated in experimental design. S.T. and T.S. participated in data analysis. S.T. wrote the manuscript with minor contributions by T.S.

References

1. **Ali, J.A., Jackson, A.P., Howells, A.J. and Maxwell, A.**, *The 43-kilodalton N-terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs.* Biochemistry, 1993. **32**(10): p. 2717-24.

2. **Brown, P.O., Peebles, C.L. and Cozzarelli, N.R.**, *A topoisomerase from Escherichia coli related to DNA gyrase.* Proc Natl Acad Sci U S A, 1979. **76**(12): p. 6110-4.

3. **Dietz, N., Huber, M., Sorg, I., Goepfert, A., Harms, A., et al.**, *Structural basis for selective AMPylation of Rac-subfamily GTPases by Bartonella effector protein 1 (Bep1).* Proc Natl Acad Sci U S A, 2021. **118**(12).

4. **Dubey, B.N., Agustoni, E., Bohm, R., Kaczmarczyk, A., Mangia, F., et al.**, *Hybrid histidine kinase activation by cyclic di-GMP-mediated domain liberation.* Proc Natl Acad Sci U S A, 2020. **117**(2): p. 1000-1008.

5. **Engel, P., Goepfert, A., Stanger, F.V., Harms, A., Schmidt, A., et al.**, *Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins.* Nature, 2012. **482**(7383): p. 107-10.

6. **Goepfert, A., Stanger, F.V., Dehio, C. and Schirmer, T.**, *Conserved inhibitory mechanism and competent ATP binding mode for adenylyltransferases with Fic fold.* PLoS One, 2013. **8**(5): p. e64901.

7. **Harms, A., Stanger, F.V., Scheu, P.D., de Jong, I.G., Goepfert, A., et al.**, *Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology.* Cell Rep, 2015. **12**(9): p. 1497-507.

8. **Kinch, L.N., Yarbrough, M.L., Orth, K. and Grishin, N.V.**, *Fido, a novel AMPylation domain common to fic, doc, and AvrB.* PLoS One, 2009. **4**(6): p. e5818.

9. **Lamour, V., Hoermann, L., Jeltsch, J.M., Oudet, P. and Moras, D.**, *Crystallization of the 43 kDa ATPase domain of Thermus thermophilus gyrase B in complex with novobiocin.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 8): p. 1376-8.

10. **Liu, L.F. and Wang, J.C.**, *Interaction between DNA and Escherichia coli DNA topoisomerase I. Formation of complexes between the protein and superhelical and nonsuperhelical duplex DNAs.* J Biol Chem, 1979. **254**(21): p. 11082-8.

11. **Luong, P., Kinch, L.N., Brautigam, C.A., Grishin, N.V., Tomchick, D.R., et al.**, *Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain.* J Biol Chem, 2010. **285**(26): p. 20155-63.

12. **Mattoo, S., Durrant, E., Chen, M.J., Xiao, J., Lazar, C.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, 2011. **286**(37): p. 32834-42.

13. **Palanivelu, D.V., Goepfert, A., Meury, M., Guye, P., Dehio, C., et al.**, *Fic domain-catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA.* Protein Sci, 2011. **20**(3): p. 492-9.

14. **Perera, L.A., Preissler, S., Zaccai, N.R., Prévost, S., Devos, J.M., et al.**, *Structures of a deAMPylation complex rationalise the switch between antagonistic catalytic activities of FICD (14/96/109).* bioRxiv, 2021: p. 2021.04.20.440599.

15. **Perera, L.A., Rato, C., Yan, Y., Neidhardt, L., McLaughlin, S.H., et al.**, *An oligomeric state-dependent switch in the ER enzyme FICD regulates AMPylation and deAMPylation of BiP.* EMBO J, 2019. **38**(21): p. e102177.

16. **Pieles, K., Glatter, T., Harms, A., Schmidt, A. and Dehio, C.**, *An experimental strategy for the identification of AMPylation targets from complex protein samples.* Proteomics, 2014. **14**(9): p. 1048-52.

17. **Preissler, S., Rato, C., Chen, R., Antrobus, R., Ding, S., et al.**, *AMPylation matches BiP activity to client protein load in the endoplasmic reticulum.* Elife, 2015. **4**: p. e12621.

18. **Preissler, S., Rato, C., Perera, L., Saudek, V. and Ron, D.**, *FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP.* Nat Struct Mol Biol, 2017. **24**(1): p. 23-29.

19. **Stanger, F.V., Dehio, C. and Schirmer, T.**, *Structure of the N-terminal Gyrase B fragment in complex with ADPPi reveals rigid-body motion induced by ATP hydrolysis.* PLoS One, 2014. **9**(9): p. e107289.

20. **Veyron, S., Oliva, G., Rolando, M., Buchrieser, C., Peyroche, G., et al.**, *A Ca(2+)-regulated deAMPylation switch in human and bacterial FIC proteins.* Nat Commun, 2019. **10**(1): p. 1142.

21. **Wigley, D.B., Davies, G.J., Dodson, E.J., Maxwell, A. and Dodson, G.**, *Crystal structure of an N-terminal fragment of the DNA gyrase B protein.* Nature, 1991. **351**(6328): p. 624-9.

22. **Worby, C.A., Mattoo, S., Kruger, R.P., Corbeil, L.B., Koller, A., et al.**, *The fic domain: regulation of cell signaling by adenylylation.* Mol Cell, 2009. **34**(1): p. 93-103.

23. **Yarbrough, M.L., Li, Y., Kinch, L.N., Grishin, N.V., Ball, H.L., et al.**, *AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling.* Science, 2009. **323**(5911): p. 269-72.

Table 1. Conditions obtained after global fitting of AMPylation and/or deAMPylation reactions.

reaction	enzyme (E)		substrate (S1)			target (S2)			target (P2)		
	name	conc. (μM)	name	conc. (μM)	eps$_{259}$	name	conc. (μM)	eps$_{259}$	name	conc. (μM)	eps$_{259}$
AMPylation	NmFic$_{\Delta 8}$	1.00	ATP	1000.00	15400	GyrB43	14.79 ± 0.26	17400	-	-	-
AMPylation / deAMPylation	VbhT(FIC)/ VbhA	15.00	ATP	1000.00	15400	GyrB43	20.34 ± 0.15	17400	-	-	-
AMPylation / deAMPylation[1]	VbhT(FIC)/ VbhA$_{E24G}$	15.00	ATP	1000.00	15400	GyrB43	23.05 ± 0.41	17400	-	-	-
deAMPylation	VbhT(FiC)/ VbhA	4.00	-	-	-	GyrB43	14.90	17400	AMP-GyrB43	9.12	32400
deAMPylation	VbhT(FIC)/ VbhA$_{E24G}$	4.00	-	-	-	GyrB43	7.84	17400	AMP-GyrB43	8.55	32400

[1] only the first six data points were used for data fitting

[2] due to unknown contamination the start concentration of GyrB43 was set higher

Table 2. Parameters obtained after global fitting of AMPylation and/or deAMPylation reactions.

reaction	enzyme	$k_{cat,1}$ (s^{-1})	$k_{cat,2}$ (s^{-1})	K_{D1} (μM)	K_{D2} (μM)	K_{D3} (μM)
AMPylation	NmFic$_{\Delta 8}$	0.77 ± 0.040	0.00	1000	100	50
AMPylation / deAMPylation	VbhT(FIC)/ VbhA	0.0019 ± 0.0007	0.011 ± 0.0007	1000	100	50
AMPylation / deAMPylation[1]	VbhT(FIC)/ VbhA$_{E24G}$	0.011 ± 0.0008	0.032 ± 0.0025	1000	100	50
deAMPylation	VbhT(FIC)/ VbhA	0.00	0.0066[2]	1000	100	50
deAMPylation	VbhT(FIC)/ VbhA$_{E24G}$	0.00	0.045[2]	1000	100	50

[1] only the first six data points were used for data fitting

[2] Monte - Carlo Algorithm was used for data fitting (no errors obtained)
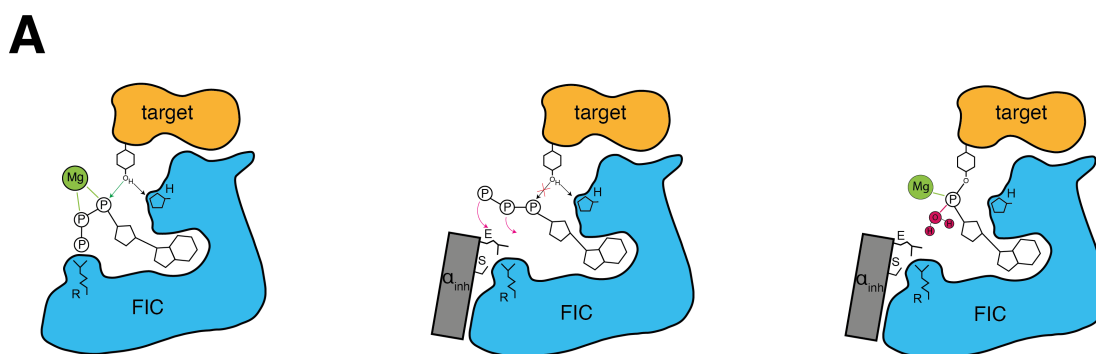
A

Figure 1: AMPylation and deAMPylation reactions can be mediated by bifunctional Fic enzymes.

(A) Scheme of the general, previously published AMPylation mechanism (left and middle) and a model for the deAMPylation process (right) [5]. AMPylation of a target protein (yellow) is mediated due to the interaction of the γ-phosphate of ATP with the second arginine (R) of the Fic motif on the toxin (blue), while a cation for example Mg^{2+} (green) holds the α- and β-phosphate in place. This leads to an orientation suitable for an in-line attack of the target hydroxyl side chain (green arrow) on the α-phosphate (P) after proton transfer (black arrow) to the conserved histidine (H) of the Fic protein. The presence of an inhibitory α-helix $α_{inh}$ with a conserved SxxxEG motif can lead to active site obstruction causing disruption of the interactions (pink arrows) and therefore ATP might not be able to bind into the active site of the enzyme (Engel et. al., 2012). Removal of the bound AMP (deAMPylation) is proposed to involve the presence of a hydrogen molecule (pink spheres) positioned by Mg^{2+}, which is able to perform an in-line attack onto the bound phosphate causing the release of free AMP and unmodified target independent of the presence of the $α_{inh}$.
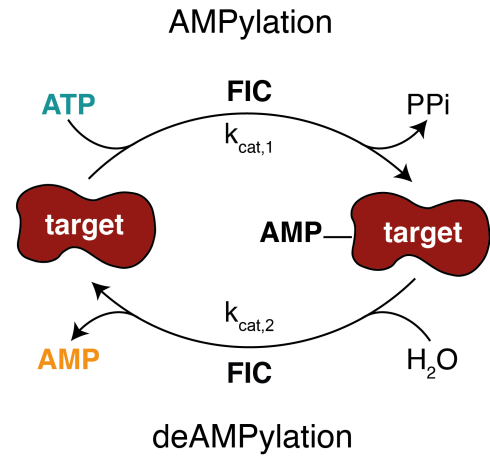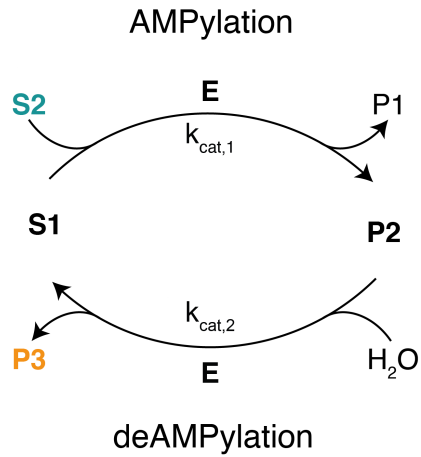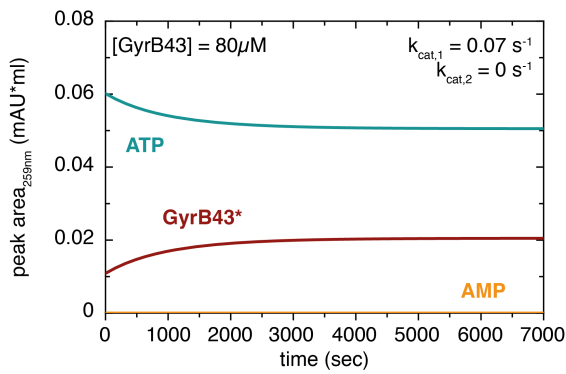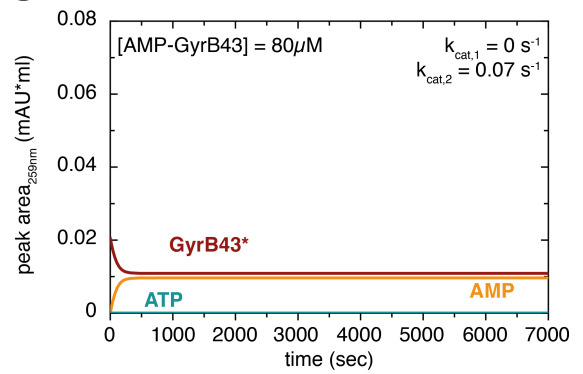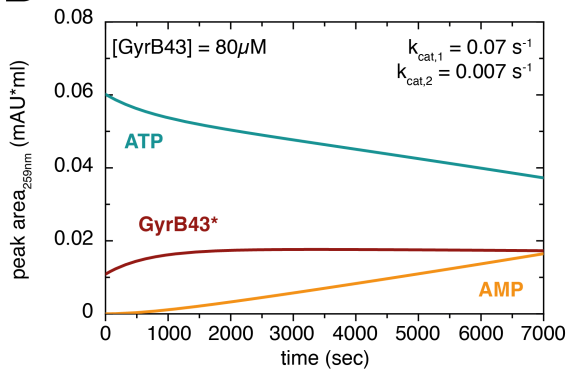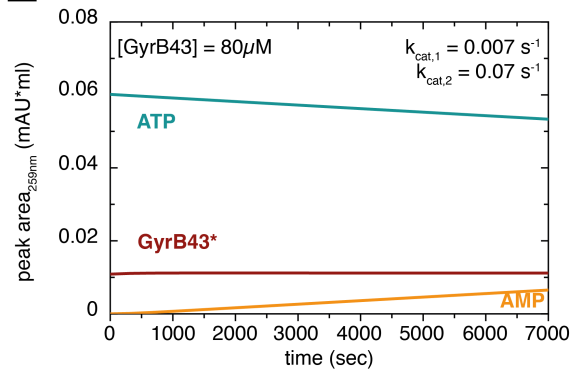
Figure 2: Kinetic scheme and progress curves as simulated for a bifunctional enzyme with AMPylation/deAMPylation activity.

(A) Left: Reaction cycle of a bifunctional Fic enzyme mediating AMPylation (turn-over number $k_{cat,1}$) and deAMPylation ($k_{cat,2}$). Substrate 1 (S1) gets AMPylated to product 2 (P2) due to the transfer of an AMP moiety from substrate 2 (S2) (turquoise) onto its hydroxyl side chain accompanied by the release of pyrophosphate (PPi). DeAMPylation proceeds via enzyme (E) catalyzed hydrolysis of the phosphodiester bond resulting in S2 and product 3.

Right: The reaction cycle with the components used for experiments displayed in this study.

(B – E) Simulated progress curves for ATP (turquoise), GyrB43+AMP-GyrB43 (GyrB43*, red), and AMP (yellow). The concentration of ATP was 500µM, the enzyme concentration was 20µM. K_{cat} values are also indicated in the panels.

(B) AMPylation simulation showing buildup of GyrB43*.

(C) DeAMPylation simulation starting with full AMPylated GyrB.

(D) AMPylation/deAMPylation simulation with rate-limiting deAMPylation. AMP production sets in with a certain delay, the concentration of the reaction intermediate AMP-GyrB43 amounts to about 70 % of GyrB43 at steady-state.

(E) AMPylation/deAMPylation simulation with rate-limiting AMPylation. AMP production sets in immediately, buildup of the reaction intermediate AMP-GyrB43 is almost negligible.
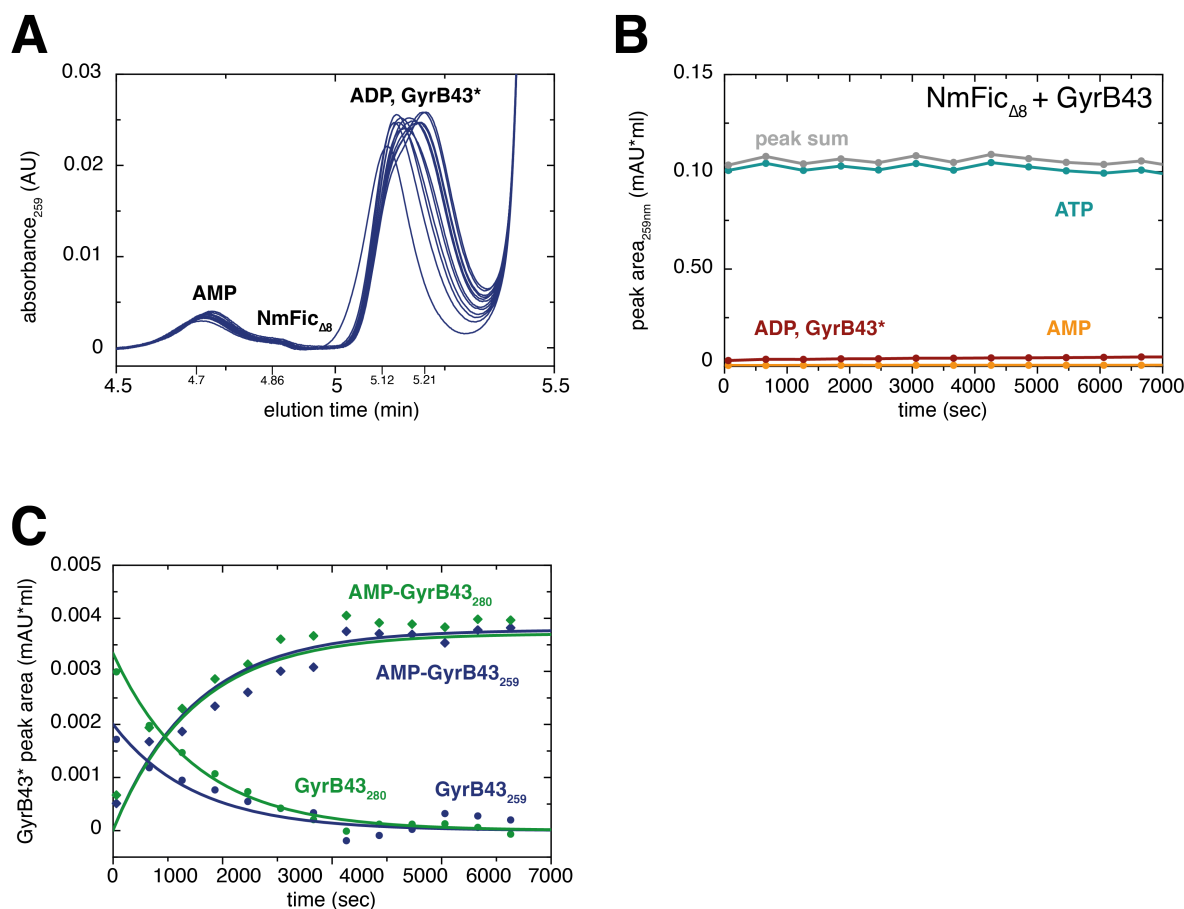
Figure 3: NmFic$_{\Delta 8}$ catalyzed GyrB43 AMPylation as monitored by online ion-exchange chromatography.

(A) Chromatograms (excerpts, measured at 259nm) acquired at incubation times t=7260 seconds (left to right, shown until 7000 seconds). GyrB43 AMPylation is manifested by a shift and area increase of the respective peak. Elution times of AMP, NmFic$_{\Delta 8}$, GyrB43 and AMP-GyrB43 are shown on the x-axis.

(B) Progress curves for ATP (turquoise), GyrB43+AMP-GyrB43 (GyrB43*, red)+ADP, AMP+NmFic$_{\Delta 8}$ (orange), and the sum of these peak areas (grey).

(C) Progress curves for the deconvoluted peak areas of GyrB43 (diamonds) and AMP-GyrB43 (circles) measured at 259nm (blue) and 280nm (green) as derived from the chromatograms shown in panel A. Lines represent the global fit to the data using the kinetic model of Fig. 2A and the parameters and conditions of tables 1 and 2.
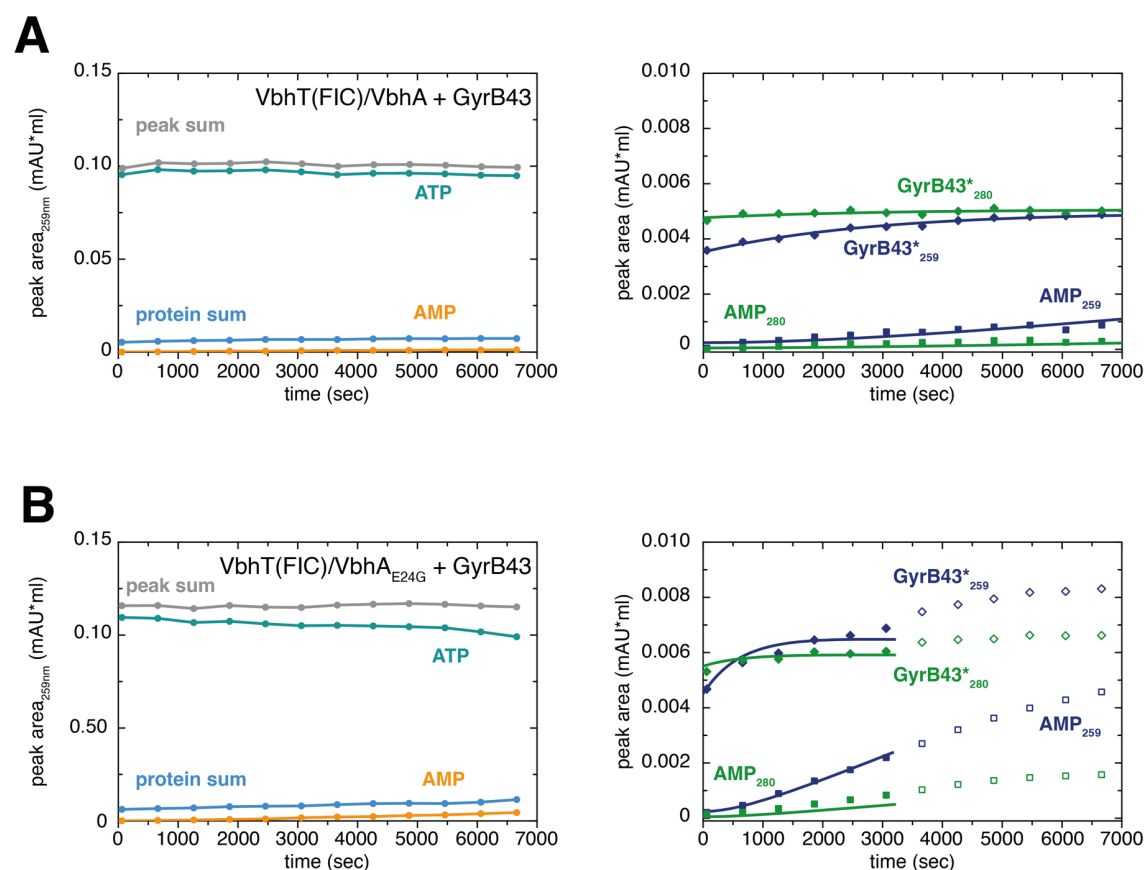
Figure 4: Time course of VbhT(FIC)/VbhA catalyzed GyrB43 ATP to AMP turnover and AMP-GyrB43 build-up.

Left: Progress curve for ATP (turquoise), VbhT(FIC)/VbhA+GyrB43+AMP-GyrB43 (protein sum, blue), AMP (orange) and the sum of peak areas (grey) at 259nm.

Right: Progress curve for GyrB43+AMP-GyrB43 (GyrB43*) (diamonds) and AMP (squares) at 259nm (blue) and 280nm (green). Lines represent the global fit of the data using the kinetic model of Fig. 2A and the parameters and conditions of tables 1 and 2.

(A) AMPylation/deAMPylation reaction catalyzed by VbhT(FIC)/VbhA.

(B) AMPylation/deAMPylation reaction catalyzed by VbhT(FIC)/VbhA$_{E24G}$.
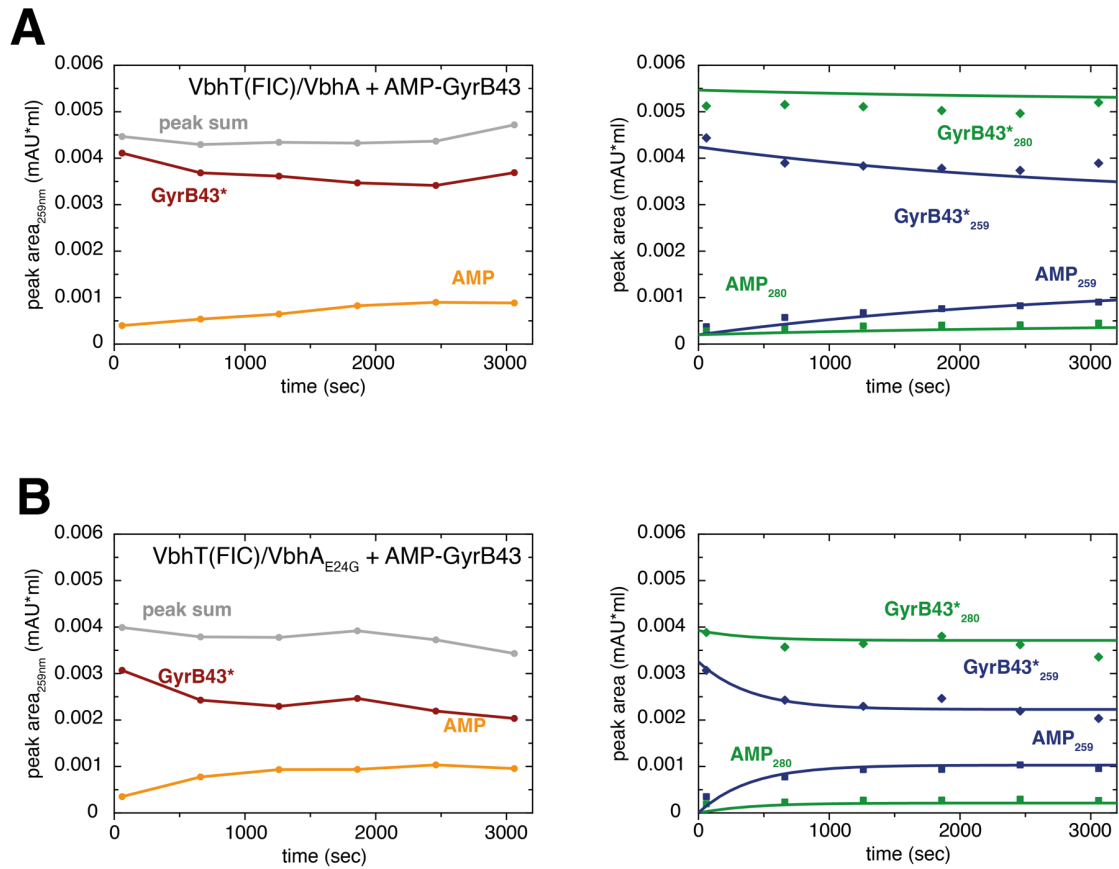
Figure 5: A closer look on the VbhT(FIC)/VbhA mediated deAMPylation of AMP-GyrB43

Left: Progress curve for GyrB43+AMP-GyrB43 (GyrB43*, red), AMP (orange) and the sum of peak areas (grey) at 259nm.

Right: Progress curve for GyrB43+AMP-GyrB43 (GyrB43*) (diamonds) and AMP (circles) at 259nm (blue) and 280nm (green). Lines represent the global fit to the data using the kinetic model of Fig. 2A and the parameters and conditions of tables 1 and 2.

(A) DeAMPylation reaction mediated by VbhT(FIC)/VbhA.

(B) DeAMPylation reaction mediated by VbhT(FIC)/VbhA$_{E24G}$.
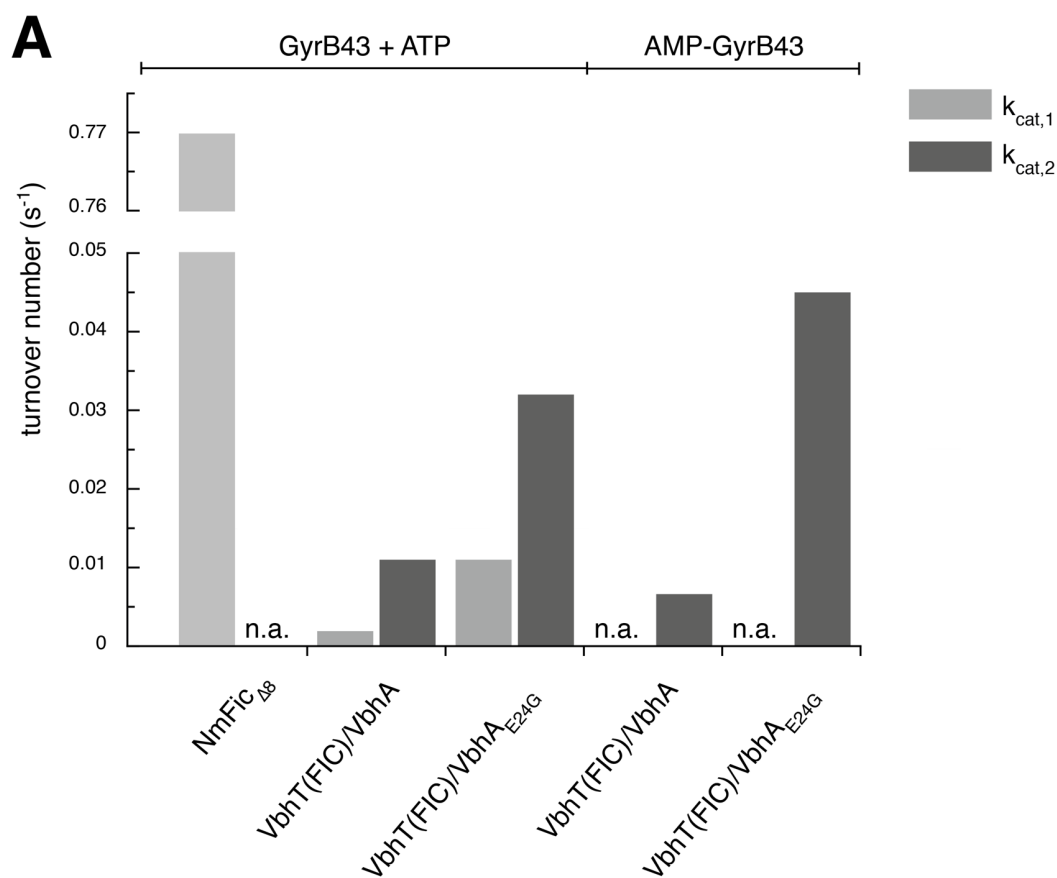
Figure 6: Comparison of AMPylation and deAMPylation rates.

All shown time courses were fitted using an enzyme catalyzed group transfer to obtain the kcat of the reactions. The details are depicted in table 2 and 3. The calculated k_{cat} is shown here in a bar plot. The $k_{cat,1}$ (light grey) represents the AMPylation reaction, $k_{cat,2}$ (dark grey) is the deAMPylation. The used targets are shown over the plot, whereas the enzymes are depicted on the y-axis.
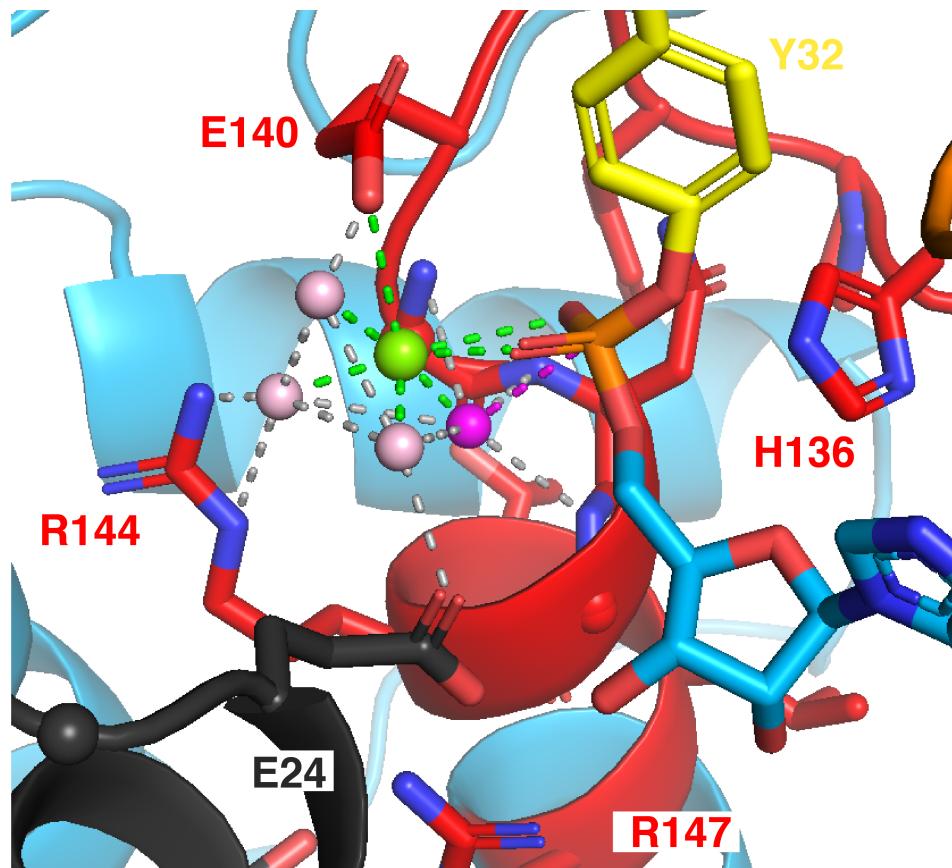
Figure 7: 3D model of the proposed location of the crucial water molecule necessary for deAMPylation.

Model of VbhT(FIC)/VbhA (PDB: 3ZC7) in complex with AMPylated Cdc42 (PDB: 4ITR). VbhT is shown in blue, with the fic loop in sticks and spheres in red. VbhA is grey with important residues of the α_{inh} shown in sticks. The AMPylated Y32 of Cdc42 is yellow and in sticks. Interaction of the modelled Mg^{2+} (green sphere) are shown in green dashes, interactions of the simulated water molecules (light pink spheres) with the fic motif are depicted in grey dashes. The essential water molecule for the hydrophilic attack is depicted as magenta sphere and its interaction with the bound AMP is in magenta dashes.

Table S1. List of plasmids used in this study.

Plasmid	Backbone	Description	Primer (fw)	Primer (rv)	Source
pRSF-Duet1	pRSF-Duet1	Empty vector (RSF 1030 *ori*, P*T7*)			Novagen
pFVS0011	pRSF-Duet1	VbhA (MCS1, HA-tagged) + VbhT 1-198 (MCS2, His$_6$-tagged)	prAG037	prFVS001	[5]
pFVS0016	pRFS-Duet1	NmFic 11-167 Δ8 (MCS2, His$_6$-tagged)	prFVS007	prFVS009	[5]
pFVS0065	pRSF-Duet1	VbhA Glu24Gly (MCS1, HA-tagged) + VbhT 1-198 (MCS2, His$_6$-tagged)	prFVS0063	prFVS0064	[5]
pFVS0109	pRSF-Duet1	*E. coli* GyrB 1-392 (MCS2, N-ter His$_6$-tagged)	prFVS107	prFVS114	[7]
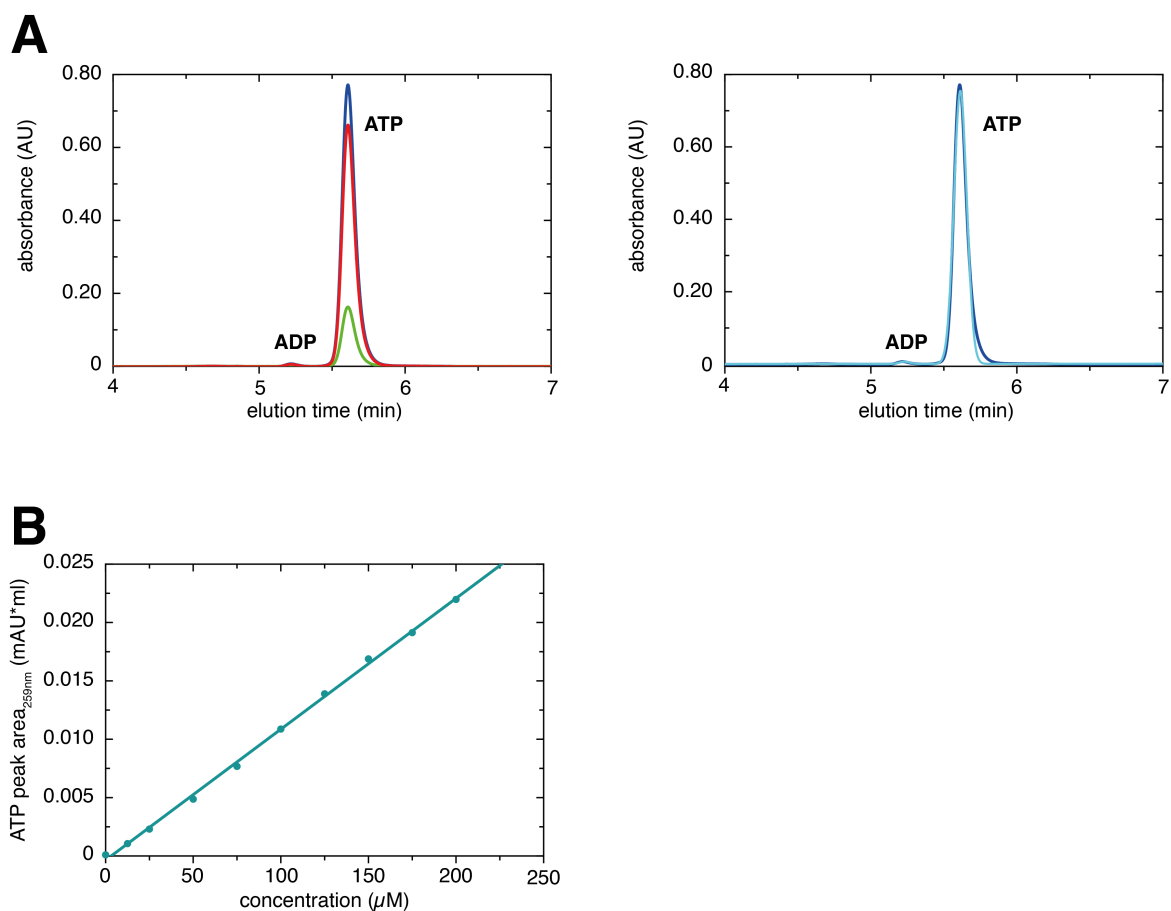
Figure S1: ATP Calibration Curve.

(A) Left: Chromatogram of 1mM ATP obtained at 259nm (blue), 253nm (red) and 280nm (green).

Right: Gaussian fit (light blue) of 1mM ATP measured at 259nm (blue).

(B) Peak areas obtained from chromatograms using ATP concentrations from 0 to 250μM to calculate the ATP related scale factor, which was 7.812 mAU*ml/mM and used for the further global fitting of progress curves.
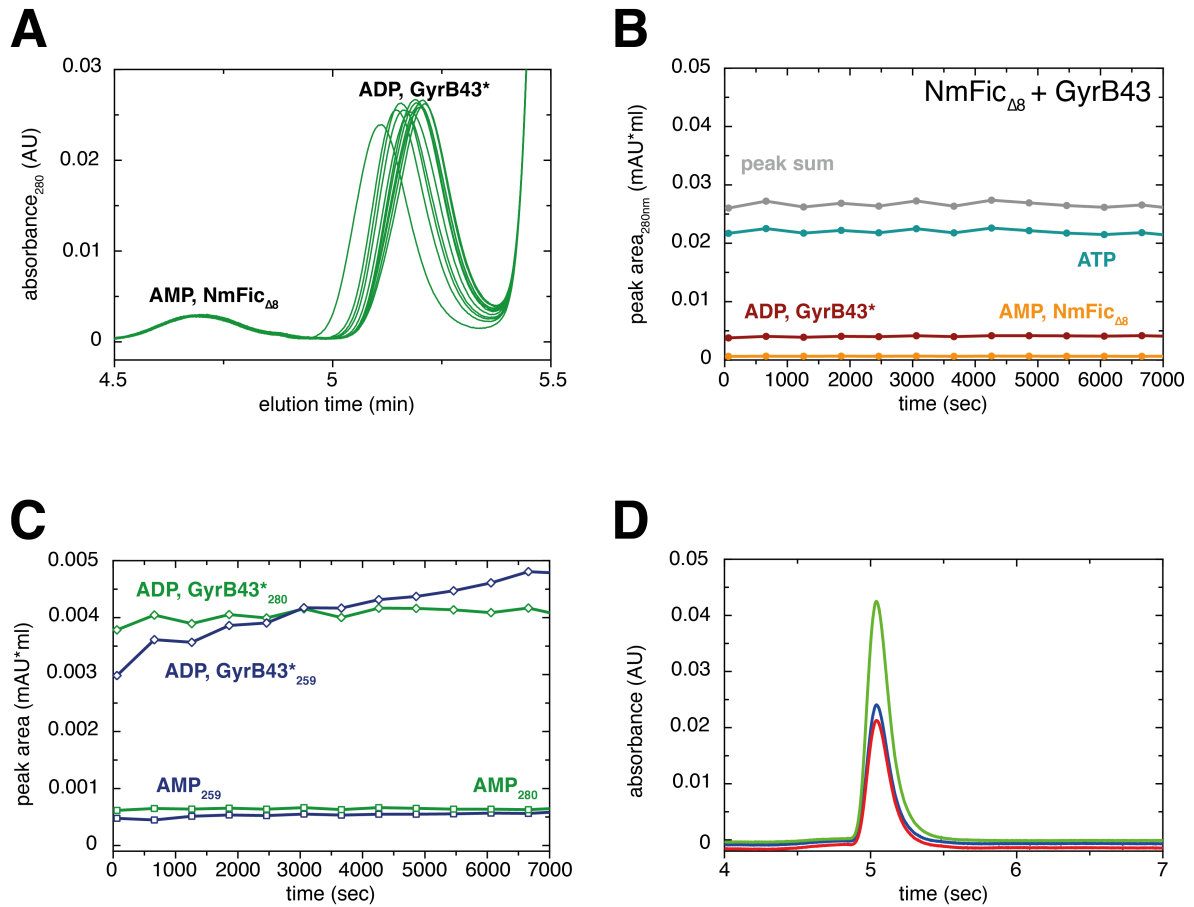
Figure S2: Chromatograms and time course at 280nm of NmFic$_{\Delta 8}$ catalyzed GyrB43 AMPylation.

(A) Chromatograms (excerpts, measured at 259 nm) acquired at incubation times t = 7260 seconds (left to right, shown until 7000 seconds). GyrB43 AMPylation is manifested by a shift and area increase of the respective peak. Elution times of AMP, NmFic$_{\Delta 8}$, GyrB43 and AMP-GyrB43 are shown on the x-axis.

(B) Progress curves for ATP (turquoise), GyrB43+AMP-GyrB43 (GyrB43*, red)+ADP, AMP+NmFic$_{\Delta 8}$ (orange), and the sum of these peak areas (grey).

(C) Progress curve of the peak areas of ADP and GyrB43*, and AMP at 259nm (dark blue) and 280nm (dark green).

(D) Chromatogram of GyrB43 in complex with Novobiocin.

3.2.3   Supplementary results

3.2.3.1   Application of online ion-exchange chromatography for HPLC

GyrB43 AMPylation was detected using radioactive end point assays where *in vitro* reactions of the target (GyrB43) and α^{32}P-ATP catalyzed by an enzyme were incubated for various time points. After stopping the reactions they were loaded onto a SDS-PAGE gel and visualized using autoradiography (Pieles et al. 2014). Since this assay has several limitations and cannot be used to characterize reactions in a quantitative way, another assay called online ion-exchange chromatography (oIEC) was developed by the lab to allow real time measurement of reactions and quantification of several reaction species to obtain enzymatic progress curves.

In this assay a reaction mix containing the reaction buffer and the target protein are prepared in a 96 deep well plate (Figure 24A). The reaction is started by adding substrate and target, and an aliquot of the reaction mix is immediately injected onto an anion exchange column pre-equilibrated with mobile phase A buffer (Figure 24B). The time span between the start of the reaction and the first load onto the column is around 60 seconds and represents the first time point (t0) in case of the used system. Proteins and negative charged nucleotides can bind to the column, whereas positive charged compounds flow through in the first washing step. The bound components are eluted with increasing concentration of ammonium sulfate present in the mobile phase B buffer (blue line). After a short equilibration step, the next injection is performed (t_i). During every injection and accompanied elution a chromatogram of the reaction at various wavelengths is measured. The chromatogram can be further processed to obtain a full-time course of the observed reaction and acquire the kinetic characteristics (Dubey et al. 2020) (Dietz et al. 2021). The assay was originally developed for two ÄKTA purification systems located in the laboratory but was further adapted for a Gilson HPLC system. This system contains of several separated components like a GX-241 II Liquid Handler without pump (Gilson), a GX Syringe Pump (Gilson), a 322 HPLC Pump (Gilson), a 172 Diode Array Detector (Gilson) and a FC 203B Fraction Collector (Gilson), which was used for the experiments in *research article II* and in the following chapters.
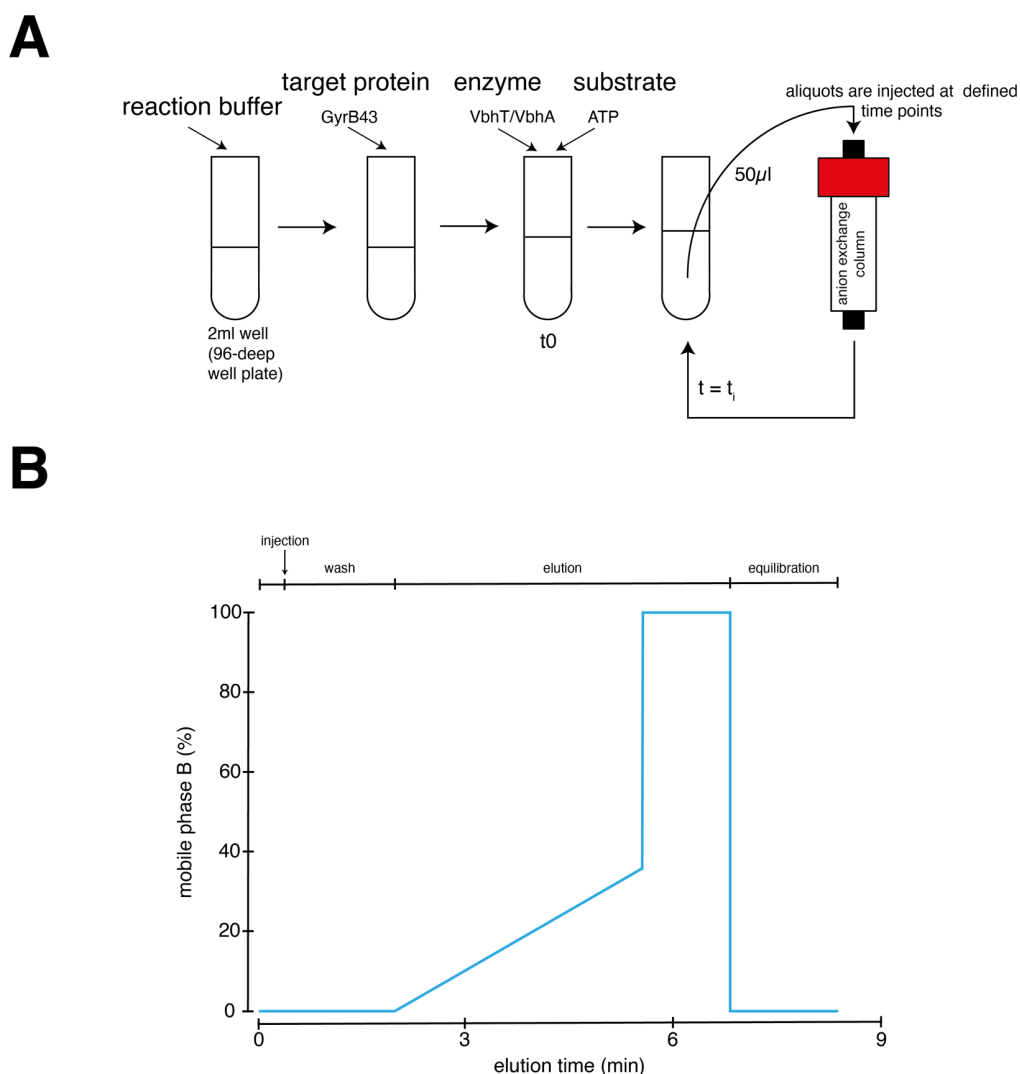
Figure 24: Pipetting scheme and reaction setup used for online ion-exchange chromatography (oIEC).
(A) A reaction mix is prepared in a deep well plate and the reaction is started by adding the substrate and the enzyme. An aliquot is immediately injected onto an anion-exchange column (time point 0 = t0). After elution of the reaction compounds the next injection can take place leading to further time points of the ongoing reaction (t=t_i).
(B) The reaction compounds can be eluted with a gradient of mobile phase B buffer (blue line), which contains ammonium sulphate. After the injection a short washing step is performed to remove unbound compounds. The bound reaction compounds are eluted with a linear gradient to 36% of buffer B ending in a short 100% step to remove all compounds from the column. After equilibration the next injection is happening.

The elution is dependent on the charge of the used proteins and nucleotides. Therefore, the mobile phase B gradient shown in figure 24B was set up to guarantee separation of AMP, GyrB43 and ATP (Figure 25A). Since ATP is used as a substrate in

AMPylation reactions, the coumarine drug novobiocin is needed to inhibit the ATPase activity of GyrB43 in kinetic experiments where modification of GyrB43 should be observed (Lamour et al. 2002). If the activity of the target is not completely inhibited a peak overlap is caused between the target and the produced ADP. This overlap is shown in figure 25B displaying chromatograms obtained 1 minute and 41 minutes after start of the reaction.

A

B
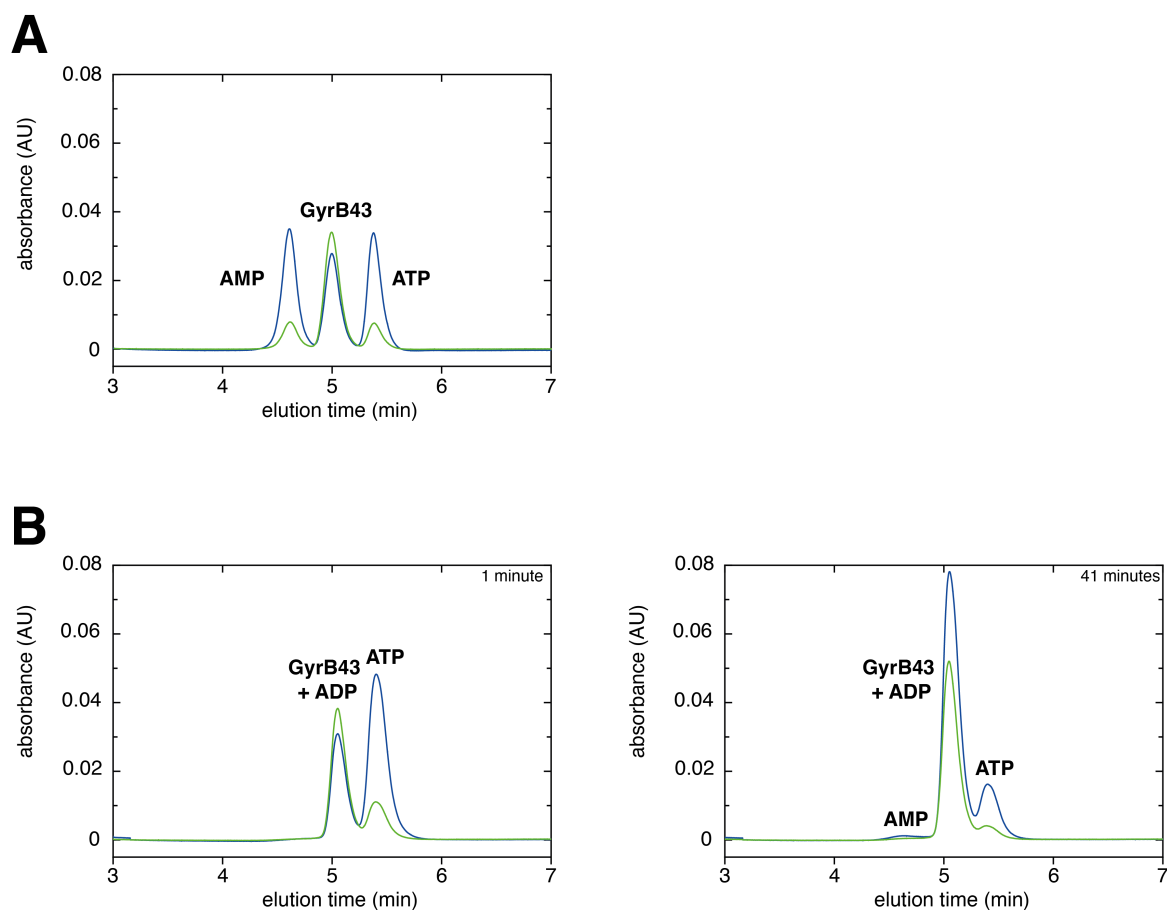
Figure 25: Peak separation of GyrB43 and nucleotides.

The shown gradient separates the used target and nucleotides. Chromatograms were obtained at 280nm (green) and 259nm (blue). The elution time in minutes is shown at the x-axis, the absorbance in absorbance units (AU) is depicted on the y-axis.

(A) Chromatogram displaying the peak separation of AMP, GyrB43 and ATP.

(B) Chromatograms obtained from a reaction containing 20µM GyrB43 and 100µM ATP 1 minute and 41 minutes after the start of the reaction. GyrB43's ATP hydrolysis activity leads to production of ADP, which is overlapping with GyrB43 for example shown after 41 minutes.

The acquired chromatograms of the measured reactions are then processed to obtain the peak areas of the detected reaction species using Gaussian peak integration and acquire a time course of the experiment. The data can be used to obtain the kinetic characteristics of the reaction using an enzyme catalyzed group transfer as shown in *research article II*. This pipeline was also used for the experiments shown in the next chapters. The kinetic conditions were not calculated due to the differences in the experimental setup and fitting programs.

3.2.3.2 GyrB43 ATPase activity and its inhibition by novobiocin

One of the first experiments to validate the oIEC assay using the Gilson HPLC system was the injection of GyrB43 without inhibitor, substrate or target. A time course of the experiment showed straight lines at 259nm (blue line) and 280nm (green line) (Figure 26A) indicating that the injected sample volume is constant over time and the protein is stable for experiments using the described method.

The 43kDa fragment of the GyrB subunit contains the transducer domain and the ATPase domain performing ATP hydrolysis, which is needed for its function in the cell cycle (Wigley et al. 1991) (Ali et al. 1993). We wanted to test if it is possible to observe this activity with oIEC. The chromatograms were processed, and the peak areas of the detected reaction species are depicted in a time course at 259nm (Figure 26B). ATP is decreasing due to the ATPase activity of GyrB43, while ADP, which is overlapping with the peak of the protein, and AMP are produced over time. The ATPase activity of GyrB43 can be inhibited by the coumarin drug novobiocin, which binds into the active site of the protein and prevents binding of the nucleotide (Lamour et al. 2002). The time course in figure 26 shows that novobiocin is successfully inhibiting Gyrb43 activity as no ATP is consumed.

GyrB43's ATPase activity can also be decreased by VbhT(FIC)/VbhA mediated AMPylation. Addition of the AMP moiety onto the side chain of Y109, which is part of the active site loop of GyrB43 causes suppression of ATP hydrolysis (Harms et al. 2015). The corresponding oIEC experiment is depicted in figure 26D.

The FIC domain of VbhT (VbhT(FIC)) ranging from residue 1 to 198 was used for the displayed experiments. In early studies of my thesis expression of VbhT(FIC)/VbhA$_{E24G}$ was not successful, therefore VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$ was used for several experiments as indicated in the figure legends. The VbhT(FIC)/VbhA complex has a similar retention time as the target GyrB43 and the peak is therefore labelled as

protein sum in the time courses. OIEC experiments showed that ATP is decreasing as expected while ADP and AMP are accumulating since both reactions are happening at the same time.
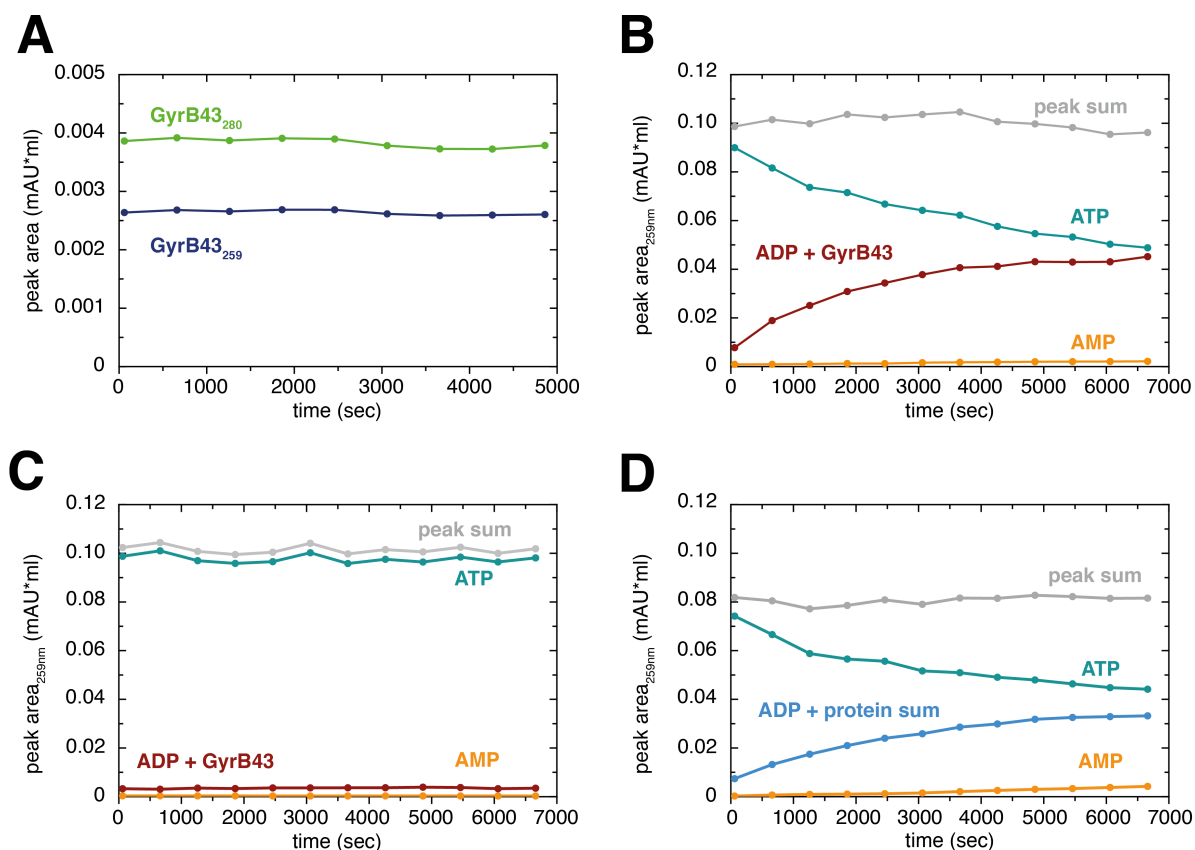
Figure 26: GyrB43 ATP hydrolysis activity and its inhibition by novobiocin or VbhT(FIC)/VbhA.
Experiments with GyrB43 were used to validate the Gilson HPLC system as a tool for oIEC. The time courses were obtained from chromatograms acquired at 259nm (in A 280nm as well). The reaction time in seconds is shown on the x-axis, the calculated peak areas in mAU*ml are on the y-axis. Each point depicts an injection onto the column. (B-D) ATP is shown in turquoise and AMP is yellow. Due to similar retention time ADP and GyrB43 are depicted together in red. The sum over all peak areas is grey.

(A) Time course of 20µMGyrB43 at pH 8.0 without other components obtained from chromatograms at 259nm (blue) and 280nm (green).

(B) Time course of 40µMGyrB43 mediated ATP hydrolysis (1mM ATP) at pH 8.0.

(C) Time course of 40µM GyrB43 at pH 8.0 with 1mM ATP and 200µM novobiocin, which inhibits GyrB43 ATPase activity.

(D) Time course of 40µM GyrB43 ATP hydrolysis (1mM ATP) at pH8.5 in presence of VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$ (20µM) mediated inhibition. Due to overlapping peaks of enzyme and target the peak is labelled as protein sum.

3.2.3.3 AutoAMPylation of VbhT(FIC)/VbhA

Radioactive endpoint assays also display autoAMPylation of VbhT(FIC)/VbhA, which is boosted after mutation of the glutamate 24 (E24) to glycine in the inhibitory α-helix (VbhA$_{E24G}$) (Engel et al. 2012) (Goepfert et al. 2013). Using oIEC to detect automodification of the enzyme should be revealed in increase of the VbhT(FIC)/VbhA peak area and decrease of the substrate ATP (Fig. 14). The time courses were obtained using different concentrations of ATP and enzyme, and at various pH values as displayed in figure 27.

The wildtype toxin-antitoxin complex VbhT(FIC)/VbhA showed no automodification consistent with end-point assay results (Goepfert et al. 2013) (Fig. 27A). The used oIEC assay conditions also did not show any autoAMPylation when the glutamate in the antitoxin was mutated (VbhT(FIC)/VbhA$_{E24G}$), which was detected in autoradiography (Goepfert et al. 2013) (Fig. 27B). Mutation of glutamate 140 (E140) in the Fic motif (VbhT(FIC)$_{E140Q}$) or of serine 175 (S175) (VbhT(FIC)$_{S175C}$) in complex with the mutated antitoxin VbhA$_{E24G}$, also revealed no autoAMPylation under the used conditions. This indicates that either the assay conditions are not suitable to detect autoAMPylation or the time frame is not suitable. OIEC assays were performed at room temperature and the reaction measured directly after adding the substrate whereas in radioactive AMPylation assays the reaction mix was incubated for one hour at 30°C, which could have an impact on the folding of the protein making it more accessible for ATP binding and resulting in automodification.
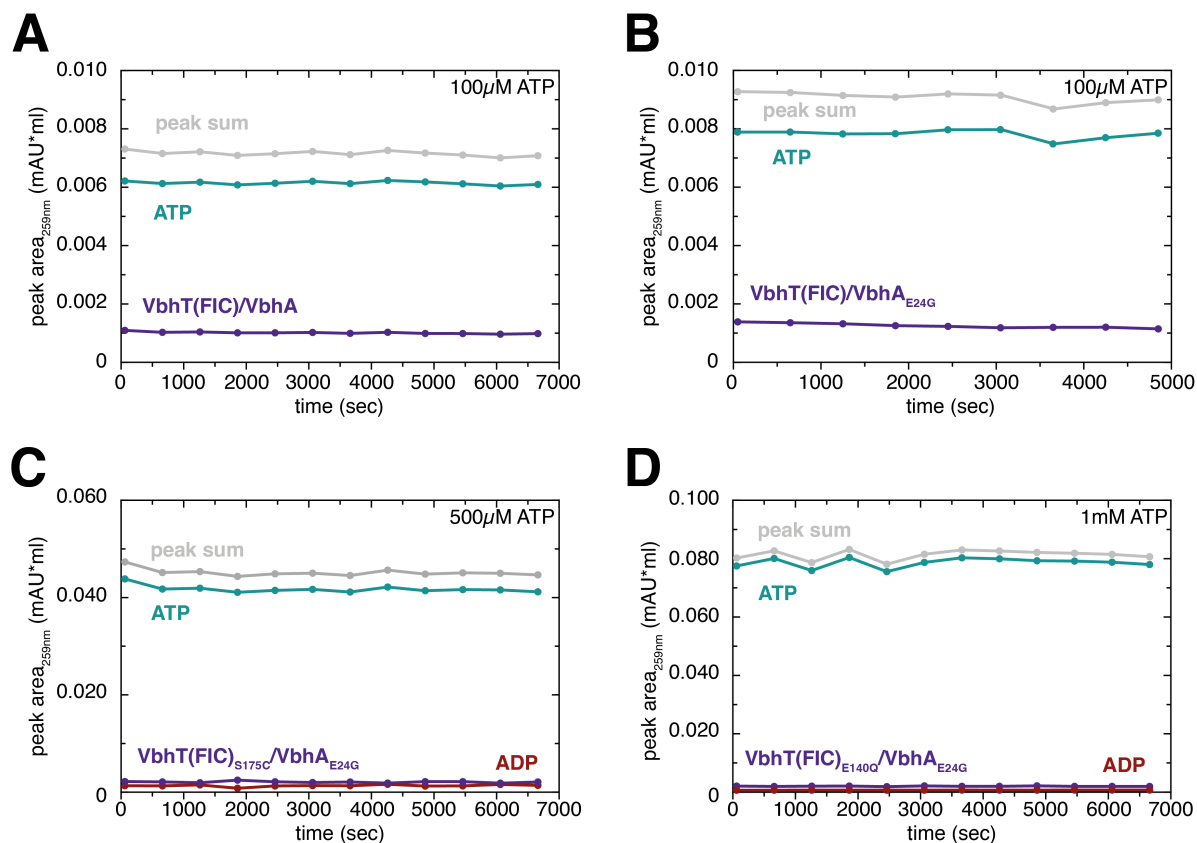
Figure 27: AutoAMPylation of different VbhT(FIC)/VbhA toxin-antitoxin complexes.

20μM of VbhT(FIC)/VbhA complexes were added to different amounts of ATP at different pH and room temperature. Every ten minutes chromatograms were obtained, and the peak areas were calculated. The enzyme is shown in violet, the ATP in turquoise, ADP in red. The sum of all peak areas is displayed in grey. (A) Time course of wildtype VbhT(FIC) in complex with the wildtype VbhA (VbhT(FIC)/VbhA) with 100μM ATP at pH 8.5.

(B) Time course of VbhT(FIC)/VbhA$_{E24G}$ with 100μM ATP at pH 8.5.

(C) Time course of VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$ with 500μM ATP at pH 8.5.

(D) Time course of VbhT(FIC)$_{E140Q}$/VbhA$_{E24G}$ with 1mM ATP at pH 8.0.

3.2.3.4 Supplementary results on GyrB43 AMPylation and deAMPylation by VbhT(FIC)/VbhA

AMPylation of GyrB43 requires ATP as a substrate, therefore novobiocin was added in the following experiments to inhibit GyrB43's ATP hydrolysis activity.

First experiments analyzing the deAMPylation of GyrB43 by VbhT(FIC)/VbhA revealed increase of AMP only at the later time points under the used condition. Therefore, the reaction time was increased to six hours instead of two hours (Figure 28). As stated in *research article II* the wildtype toxin-antitoxin complex shows AMPylation and

deAMPylation activity. AMPylation is taking place due to the decrease of ATP over time, while the protein sum, containing the modified and unmodified target and the enzyme, is increasing because of the addition of AMP. Delayed accumulation of AMP reveals deAMPylation of the target protein indicating a bifunctional toxin-antitoxin complex (Figure 28A). While most of the tested VbhT(FIC)/VbhA complexes show both functions VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$ only has AMPylation activity as depicted in figure 28B. ATP is consumed until the available GyrB43 pool is modified and no AMP is produced indicating that the serine is required for deAMPylation of GyrB43. A reaction mix containing the wildtype VbhT(FIC)/VbhA complex as well as VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$ revealed strong AMPylation and deAMPylation of GyrB43 (Figure 28C). After six hours nearly all the ATP is either added to the target because of AMPylation or converted into AMP due to deAMPylation. Mixing of the two complexes also might indicate that the antitoxins could switch toxins resulting in a highly efficient bifunctional VbhT(FIC)/VbhA$_{E24G}$ complex and an AMPylation inactive VbhT(FIC)$_{S175C}$/VbhA complex.
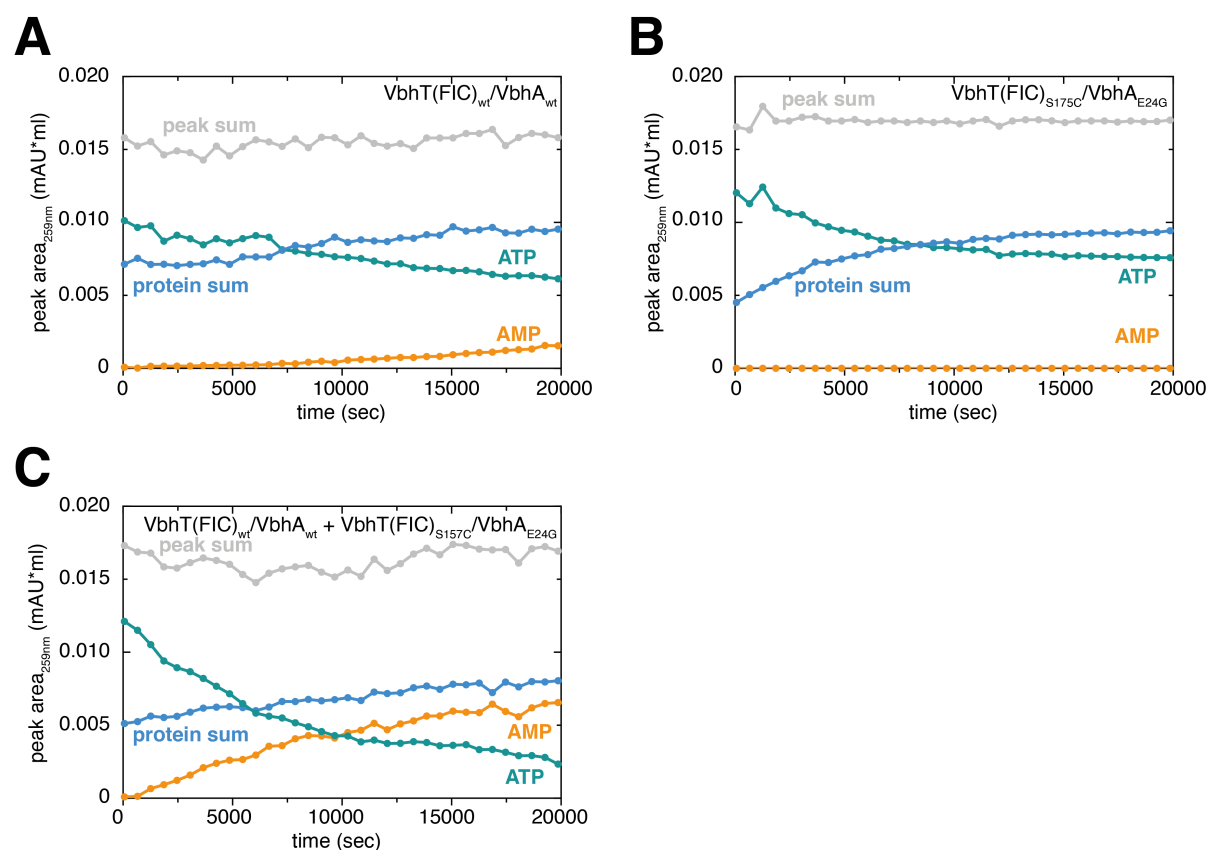
Figure 28: VbhT(FIC)/VbhA mediates AMPylation and deAMPylation in 6 hours' time courses.

Every ten minutes chromatograms of the indicated reactions containing 40µM GyrB43 (inhibited by addition of novobiocin), 100µM ATP and 20µM enzyme at pH 8.5 and room temperature were obtained over a period of six hours. The peak areas were calculated and shown in the time courses. Due to similar retention time the enzyme and modified/unmodified target (GyrB43*) are shown as protein sum (blue line). ATP is displayed in turquoise, AMP in yellow. The sum over all peak areas is shown in grey.

(A) Time course of an AMPylation/deAMPylation reaction mediated by the wildtype VbhT(FIC)/VbhA complex.

(B) Time course of an AMPylation reaction by VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$.

(C) Time course of an AMPylation/deAMPylation reaction mediated by a mix of VbhT(FIC)/VbhA and VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$.

Two VbhT(FIC)/VbhA complexes showed no AMPylation of GyrB43. Mutation of the catalytic histidine 136 (H136, VbhT(FIC)$_{H136A}$) in the Fic motif is known to suppress the effect of the toxins in plating experiments and no modification of GyrB43 is detected in autoradiography (Engel et al. 2012) (Goepfert et al. 2013). A double mutant with the modified antitoxin (VbhA$_{E24G}$) displayed no AMPylation of GyrB43 in oIEC experiments as indicated by the lack ATP consumption in a time span of 60 minutes (Figure 29A). VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$ displays AMPylation of GyrB43 while the VbhT(FIC)$_{S175C}$ toxin in complex with wildtype VbhA reveals no modification due to the presence of the glutamate side chain preventing productive substrate binding (Figure 29B).
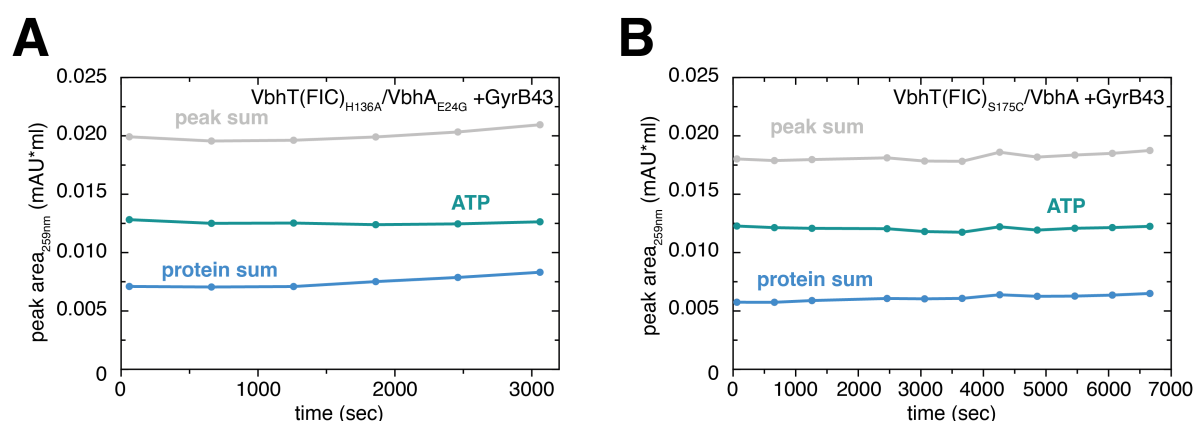
Figure 29: Time courses of AMPylation-incompetent VbhT(FIC)/VbhA complexes.

Chromatograms at 259nm were obtained from experiments containing 40µM GyrB43 (ATPase activity inhibited with novobiocin), 100µm ATP and 20µM enzyme at pH 8.5 and room temperature. The peak areas were calculated using Gaussian peak fitting and time courses were produced. ATP is shown in turquoise, the

protein sum containing the enzyme and the modified/unmodified GyrB43 (GyrB43*) is depicted in blue, and the sum over all detected peaks in grey.

(A) Time course at 259nm of an experiment with VbhT(FIC)$_{H136A}$/VbhA$_{E24G}$.

(B) Time course at 259nm of an experiment with VbhT(FIC)$_{S175C}$/VbhA.

Solved crystal structures of several Fic proteins containing an E>G mutation in the inhibition motif revealed the presence of a magnesium. The metal links the α- and the β-phosphate of the ATP substrate and is coordinated by the conserved aspartate (D) or glutamate (E) in the Fic active site motif. This cation is only observed in the mutated complexes, but not in the wildtype Fic proteins and is crucial for the toxin's AMPylation activity due to finetuning of the phosphate orientation of the ligand (Goepfert et al. 2013).

Experiments with the class II human Fic protein FICD and the class III Fic protein EfFic showed the importance of the metal ions Mg^{2+} and Ca^{2+} in AMPylation and deAMPylation. Ca^{2+} can bind to EfFic but competes with Mg^{2+} resulting in downregulation of AMPylation activity. Meanwhile Ca^{2+} downregulates deAMPylation of FICD by hindering Mg^{2+} binding. Mutation of the metal coordinating glutamate 115 (E115) in EfFic to alanine suppresses deAMPylation in presence of Ca^{2+} (Veyron et al. 2019). In VbhT(FIC)/VbhA glutamate 140 (VbhT(FIC)$_{E140}$) is responsible for coordinating the Mg^{2+} during AMPylation reactions (Goepfert et al. 2013). To investigate the importance of that specific residue it was mutated either to aspartate (VbhT(FIC)$_{E140D}$), which contains a negative charge but shorter side chain, or to glutamine (VbhT(FIC)$_{E140Q}$) having no charged side chain.

Experiments using VbhT(FIC)$_{E140D}$ in complex with the wildtype VbhA show no modification of GyrB43 as depicted in figure 30A, which might appear due to the glutamate side chain in the conserved inhibition motif preventing substrate binding. The combination of VbhT(FIC)$_{E140D}$ and VbhA$_{E24G}$ revealed AMPylation of GyrB43 and accumulation of AMP indicating deAMPylation (Fig. 30B). This suggests that in addition to the productive substrate binding the aspartate is still able to coordinate the Mg^{2+} resulting in suitable positioning of the cation. DeAMPylation of GyrB43 by VbhT(FIC)$_{E140D}$/VbhA$_{E24G}$ was also shown in experiments using AMPylated GyrB43 (AMP-GyrB43) as a substrate and accumulation of AMP was displayed over time (Fig. 30C). AMP-GyrB43 was produced using the same protocol as described in research article II. Mutation of E140 to glutamine (VbhT(FIC)$_{E140Q}$) in complex with VbhA$_{E24G}$ revealed no modification of GyrB43 and

might be caused due to the missing negative charge of the side chain, which cannot orientate the Mg^{2+}. This therefore results in unproductive ATP binding.
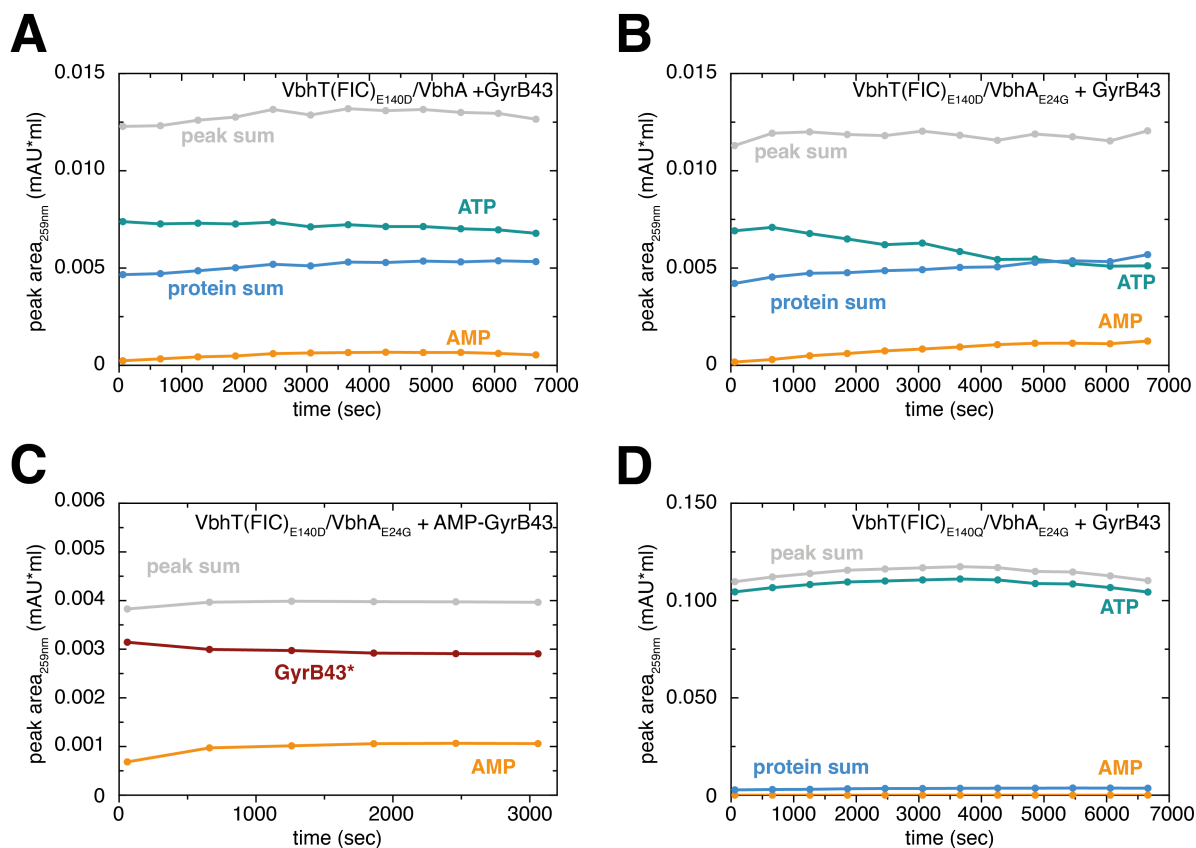
Figure 30: Role of the metal coordinating E140 of VbhT(FIC) in AMPylation and deAMPylation of GyrB43.

(A, B, D) Reaction mixes containing 40μM GyrB43 with novobiocin were prepared and the reaction started by adding 20μM of the indicated enzymes and ATP at pH 8.0 and room temperature. The chromatograms were obtained every ten minutes, the peak areas at 259nm were calculated and displayed in the shown time courses. ATP is depicted in turquois, AMP in yellow, the protein sum in blue and the peak sum over all detected peaks is shown in grey.

(A) Time course of VbhT(FIC)$_{E140D}$/VbhA and 100μM ATP.

(B) Time course of VbhT(FIC)$_{E140D}$/VbhA$_{E24G}$ and 100μM ATP.

(C) Time course of 40μM AMP-GyrB43 (preparation of AMP-GyrB43 as stated in research article II) shown in red and 4μM VbhT(FIC)$_{E140D}$/VbhA$_{E24G}$. The accumulated AMP is shown in yellow and the peak area is depicted in grey.

(D) Time course of VbhT(FIC)$_{E140Q}$/VbhA$_{E24G}$ 1mM ATP.

3.2.4   Concluding remarks

The toxin-antitoxin complex VbhT(FIC)/VbhA has bifunctional activity, which is enhanced when the inhibitory glutamate of the antitoxin is mutated and no side chain is present as stated in research article II. This leads to the assumption that the network of hydrogen bonds is altered and the hydrophilic attack on the bound AMP of the target can be better coordinated. Additional data using a mutation of the catalytic histidine (VbhT(FIC)$_{H136A}$) in complex with VbhA$_{E24G}$ displayed no AMPylation/deAMPylation activity due to the missing deprotonation site for the incoming target hydroxyl necessary for modification. Interestingly, VbhT(FIC)$_{S175C}$/VbhA$_{E24G}$, which was originally used due to poor expression of the single mutant VbhT(FIC)/VbhA$_{E24G}$, only AMPylates the target but does not induce deAMPylation. While the double mutant displays modification of the target, the single mutation VbhT(FIC)$_{S175C}$/VbhA does not mediate AMP-transfer. This leads to the assumption that the serine, which forms a hydrogen bond to the 2'-OH of the ATP ribose, is necessary to mediate bifunctionality of the toxin-antitoxin complex, while mutation to cysteine abolishes deAMPylation activity. This suggests that the different chemical properties of that position play a role in binding of the substrate ribose.

Class II FICD and class III EfFic are regulated by the presence of Mg^{2+} or Ca^{2+}, which are coordinated by the aspartate or glutamate of the Fic motif. While AMPylation activity of EfFic is enhanced in the presence of Mg^{2+}, Ca^{2+} leads to increased deAMPylation. DeAMPylation activity is abolished when the metal coordinating glutamate is mutated to alanine. Ca^{2+} downregulates deAMPylation of FICD leading to active AMPylation of BiP (Veyron et al. 2018) (Engel et al. 2012). Preliminary experiments of a VbhT(FIC)/VbhA complex with mutations of the metal coordinating glutamate 140 (E140) to aspartate (VbhT(FIC)$_{E140D}$) or glutamine (VbhT(FIC)$_{E140Q}$) already show some insight into the metal coordination by E140 of VbhT(FIC). VbhT(FIC)$_{E140D}$ in complex with VbhA$_{E24G}$ displays bifunctionality indicating that the metal ion can still be coordinated with a shorter side chain. This was also shown for FICD, which has an aspartate in its FIC motif. DeAMPylation activity mediated by this VbhT(FIC)/VbhA complex is confirmed after incubation of the purified AMP-GyrB43 with VbhT(FIC)$_{E140D}$/VbhA$_{E24G}$. Removal of the negative charge of E140 with VbhT$_{E140Q}$ mutant showed no AMPylation and deAMPylation, which confirms that the ion might not be positioned properly and the

substrate is not in a suitable conformation revealing the metal ion being crucial for GyrB43 modification.

Further investigations regarding the role of the antitoxin glutamate are needed because of its role in the bifunctional activity of the toxin-antitoxin complex. Additionally, the importance of the metal ion in VbhT(FIC)/VbhA mediated GyrB43 AMPylation and deAMPylation needs to be analyzed using different metal ions such as Mn^{2+} and Ca^{2+} in future oIEC experiments and crystallization experiments.

4 References

Adler, S. P., D. Purich and E. R. Stadtman (1975). "Cascade control of Escherichia coli glutamine synthetase. Properties of the PII regulatory protein and the uridylyltransferase-uridylyl-removing enzyme." J Biol Chem **250**(16): 6264-6272.

Ali, J. A., A. P. Jackson, A. J. Howells and A. Maxwell (1993). "The 43-kilodalton N-terminal fragment of the DNA gyrase B protein hydrolyzes ATP and binds coumarin drugs." Biochemistry **32**(10): 2717-2724.

Almassy, R. J., C. A. Janson, R. Hamlin, N. H. Xuong and D. Eisenberg (1986). "Novel subunit-subunit interactions in the structure of glutamine synthetase." Nature **323**(6086): 304-309.

Arnez, J. G. and D. Moras (1997). "Structural and functional considerations of the aminoacylation reaction." Trends Biochem Sci **22**(6): 211-216.

Berg, P. and E. J. Offengand (1958). "An Enzymatic Mechanism for Linking Amino Acids to RNA." Proc Natl Acad Sci U S A **44**(2): 78-86.

Bernard, P., K. E. Kezdy, L. Van Melderen, J. Steyaert, L. Wyns, M. L. Pato, P. N. Higgins and M. Couturier (1993). "The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase." J Mol Biol **234**(3): 534-541.

Brantl, S. and N. Jahn (2015). "sRNAs in bacterial type I and type III toxin-antitoxin systems." FEMS Microbiol Rev **39**(3): 413-427.

Brick, P., T. N. Bhat and D. M. Blow (1989). "Structure of tyrosyl-tRNA synthetase refined at 2.3 A resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate." J Mol Biol **208**(1): 83-98.

Brino, L., A. Urzhumtsev, M. Mousli, C. Bronner, A. Mitschler, P. Oudet and D. Moras (2000). "Dimerization of Escherichia coli DNA-gyrase B provides a structural mechanism for activating the ATPase catalytic center." J Biol Chem **275**(13): 9468-9475.

Brown, P. O., C. L. Peebles and N. R. Cozzarelli (1979). "A topoisomerase from Escherichia coli related to DNA gyrase." Proc Natl Acad Sci U S A **76**(12): 6110-6114.

Bunney, T. D., A. R. Cole, M. Broncel, D. Esposito, E. W. Tate and M. Katan (2014). "Crystal structure of the human, FIC-domain containing protein HYPE and implications for its functions." Structure **22**(12): 1831-1843.

Cappadocia, L. and C. D. Lima (2018). "Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism." Chem Rev **118**(3): 889-918.

Christensen-Dalsgaard, M., M. G. Jorgensen and K. Gerdes (2010). "Three new RelE-homologous mRNA interferases of Escherichia coli differentially induced by environmental stresses." Mol Microbiol **75**(2): 333-348.

Cusack, S., C. Berthet-Colominas, M. Hartlein, N. Nassar and R. Leberman (1990). "A second class of synthetase structure revealed by X-ray analysis of Escherichia coli seryl-tRNA synthetase at 2.5 A." Nature **347**(6290): 249-255.

Daniels, N. A., L. MacKinnon, R. Bishop, S. Altekruse, B. Ray, R. M. Hammond, S. Thompson, S. Wilson, N. H. Bean, P. M. Griffin and L. Slutsker (2000). "Vibrio parahaemolyticus infections in the United States, 1973-1998." J Infect Dis **181**(5): 1661-1666.

Dehio, C., C. Lanz, R. Pohl, P. Behrens, D. Bermond, Y. Piemont, K. Pelz and A. Sander (2001). "Bartonella schoenbuchii sp. nov., isolated from the blood of wild roe deer." Int J Syst Evol Microbiol **51**(Pt 4): 1557-1565.

Dietz, N., M. Huber, I. Sorg, A. Goepfert, A. Harms, T. Schirmer and C. Dehio (2021). "Structural basis for selective AMPylation of Rac-subfamily GTPases by Bartonella effector protein 1 (Bep1)." Proc Natl Acad Sci U S A **118**(12).

Dubey, B. N., E. Agustoni, R. Bohm, A. Kaczmarczyk, F. Mangia, C. von Arx, U. Jenal, S. Hiller, I. Plaza-Menacho and T. Schirmer (2020). "Hybrid histidine kinase activation by cyclic di-GMP-mediated domain liberation." Proc Natl Acad Sci U S A **117**(2): 1000-1008.

Engel, P., A. Goepfert, F. V. Stanger, A. Harms, A. Schmidt, T. Schirmer and C. Dehio (2012). "Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins." Nature **482**(7383): 107-110.

Engel, P., W. Salzburger, M. Liesch, C. C. Chang, S. Maruyama, C. Lanz, A. Calteau, A. Lajus, C. Medigue, S. C. Schuster and C. Dehio (2011). "Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen Bartonella." PLoS Genet **7**(2): e1001296.

Eriani, G., M. Delarue, O. Poch, J. Gangloff and D. Moras (1990). "Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs." Nature **347**(6289): 203-206.

Faber, P. W., G. T. Barnes, J. Srinidhi, J. Chen, J. F. Gusella and M. E. MacDonald (1998). "Huntingtin interacts with a family of WW domain proteins." Hum Mol Genet **7**(9): 1463-1474.

Fernandez De Henestrosa, A. R., T. Ogi, S. Aoyagi, D. Chafin, J. J. Hayes, H. Ohmori and R. Woodgate (2000). "Identification of additional genes belonging to the LexA regulon in Escherichia coli." Mol Microbiol **35**(6): 1560-1572.

Fineran, P. C., T. R. Blower, I. J. Foulds, D. P. Humphreys, K. S. Lilley and G. P. Salmond (2009). "The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair." Proc Natl Acad Sci U S A **106**(3): 894-899.

Francklyn, C. S. and P. Mullen (2019). "Progress and challenges in aminoacyl-tRNA synthetase-based therapeutics." J Biol Chem **294**(14): 5365-5385.

Ghosh, P. (2004). "Process of protein transport by the type III secretion system." <u>Microbiol Mol Biol Rev</u> **68**(4): 771-795.

Giege, R. (2006). "The early history of tRNA recognition by aminoacyl-tRNA synthetases." <u>J Biosci</u> **31**(4): 477-488.

Goepfert, A., F. V. Stanger, C. Dehio and T. Schirmer (2013). "Conserved inhibitory mechanism and competent ATP binding mode for adenylyltransferases with Fic fold." <u>PLoS One</u> **8**(5): e64901.

Guy, L., B. Nystedt, Y. Sun, K. Naslund, E. C. Berglund and S. G. Andersson (2012). "A genome-wide study of recombination rate variation in Bartonella henselae." <u>BMC Evol Biol</u> **12**: 65.

Harms, A., D. E. Brodersen, N. Mitarai and K. Gerdes (2018). "Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology." <u>Mol Cell</u> **70**(5): 768-784.

Harms, A. and C. Dehio (2012). "Intruders below the radar: molecular pathogenesis of Bartonella spp." <u>Clin Microbiol Rev</u> **25**(1): 42-78.

Harms, A., M. Liesch, J. Korner, M. Quebatte, P. Engel and C. Dehio (2017). "A bacterial toxin-antitoxin module is the origin of inter-bacterial and inter-kingdom effectors of Bartonella." <u>PLoS Genet</u> **13**(10): e1007077.

Harms, A., F. H. Segers, M. Quebatte, C. Mistl, P. Manfredi, J. Korner, B. B. Chomel, M. Kosoy, S. Maruyama, P. Engel and C. Dehio (2017). "Evolutionary Dynamics of Pathoadaptation Revealed by Three Independent Acquisitions of the VirB/D4 Type IV Secretion System in Bartonella." <u>Genome Biol Evol</u> **9**(3): 761-776.

Harms, A., F. V. Stanger and C. Dehio (2016). "Biological Diversity and Molecular Plasticity of FIC Domain Proteins." <u>Annu Rev Microbiol</u> **70**: 341-360.

Harms, A., F. V. Stanger, P. D. Scheu, I. G. de Jong, A. Goepfert, T. Glatter, K. Gerdes, T. Schirmer and C. Dehio (2015). "Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology." <u>Cell Rep</u> **12**(9): 1497-1507.

Hedberg, C. and A. Itzen (2015). "Molecular perspectives on protein adenylylation." <u>ACS Chem Biol</u> **10**(1): 12-21.

Hershko, A. and A. Ciechanover (1998). "The ubiquitin system." <u>Annu Rev Biochem</u> **67**: 425-479.

Hochstrasser, M. (2009). "Origin and function of ubiquitin-like proteins." <u>Nature</u> **458**(7237): 422-429.

Itzen, A., W. Blankenfeldt and R. S. Goody (2011). "Adenylylation: renaissance of a forgotten post-translational modification." <u>Trends Biochem Sci</u> **36**(4): 221-228.

Jaggi, R., W. C. van Heeswijk, H. V. Westerhoff, D. L. Ollis and S. G. Vasudevan (1997). "The two opposing activities of adenylyl transferase reside in distinct homologous domains, with intramolecular signal transduction." EMBO J **16**(18): 5562-5571.

Janson, C. A., P. S. Kayne, R. J. Almassy, M. Grunstein and D. Eisenberg (1986). "Sequence of glutamine synthetase from Salmonella typhimurium and implications for the protein structure." Gene **46**(2-3): 297-300.

Kawamukai, M., H. Matsuda, W. Fujii, T. Nishida, Y. Izumoto, M. Himeno, R. Utsumi and T. Komano (1988). "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**(9): 3864-3869.

Kawano, M., L. Aravind and G. Storz (2007). "An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin." Mol Microbiol **64**(3): 738-754.

Khater, S. and D. Mohanty (2015). "Deciphering the Molecular Basis of Functional Divergence in AMPylating Enzymes by Molecular Dynamics Simulations and Structure Guided Phylogeny." Biochemistry **54**(33): 5209-5224.

Kinch, L. N., M. L. Yarbrough, K. Orth and N. V. Grishin (2009). "Fido, a novel AMPylation domain common to fic, doc, and AvrB." PLoS One **4**(6): e5818.

Kingdon, H. S., B. M. Shapiro and E. R. Stadtman (1967). "Regulation of glutamine synthetase. 8. ATP: glutamine synthetase adenylyltransferase, an enzyme that catalyzes alterations in the regulatory properties of glutamine synthetase." Proc Natl Acad Sci U S A **58**(4): 1703-1710.

Lamour, V., L. Hoermann, J. M. Jeltsch, P. Oudet and D. Moras (2002). "Crystallization of the 43 kDa ATPase domain of Thermus thermophilus gyrase B in complex with novobiocin." Acta Crystallogr D Biol Crystallogr **58**(Pt 8): 1376-1378.

Lamour, V., L. Hoermann, J. M. Jeltsch, P. Oudet and D. Moras (2002). "An open conformation of the Thermus thermophilus gyrase B ATP-binding domain." J Biol Chem **277**(21): 18947-18953.

Lehnherr, H., E. Maguin, S. Jafri and M. B. Yarmolinsky (1993). "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**(3): 414-428.

Liu, L. F. and J. C. Wang (1979). "Interaction between DNA and Escherichia coli DNA topoisomerase I. Formation of complexes between the protein and superhelical and nonsuperhelical duplex DNAs." J Biol Chem **254**(21): 11082-11088.

Liu, L. F. and J. C. Wang (1987). "Supercoiling of the DNA template during transcription." Proc Natl Acad Sci U S A **84**(20): 7024-7027.

Lu, C., E. S. Nakayasu, L. Q. Zhang and Z. Q. Luo (2016). "Identification of Fic-1 as an enzyme that inhibits bacterial DNA replication by AMPylating GyrB, promoting filament formation." Sci Signal **9**(412): ra11.

Lu, X., S. K. Olsen, A. D. Capili, J. S. Cisar, C. D. Lima and D. S. Tan (2010). "Designed semisynthetic protein inhibitors of Ub/Ubl E1 activating enzymes." J Am Chem Soc **132**(6): 1748-1749.

Makarova, K. S., Y. I. Wolf and E. V. Koonin (2009). "Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes." Biol Direct **4**: 19.

Marlaire, S. and C. Dehio (2021). "Bartonella effector protein C mediates actin stress fiber formation via recruitment of GEF-H1 to the plasma membrane." PLoS Pathog **17**(1): e1008548.

Masuda, H., Q. Tan, N. Awano, K. P. Wu and M. Inouye (2012). "YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in Escherichia coli." Mol Microbiol **84**(5): 979-989.

Mattoo, S., E. Durrant, M. J. Chen, J. Xiao, C. S. Lazar, G. Manning, J. E. Dixon and C. A. Worby (2011). "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**(37): 32834-32842.

Maxwell, A. and A. J. Howells (1999). "Overexpression and purification of bacterial DNA gyrase." Methods Mol Biol **94**: 135-144.

Merrell, D. S. and S. Falkow (2004). "Frontal and stealth attack strategies in microbial pathogenesis." Nature **430**(6996): 250-256.

Montero, M., M. Brini, R. Marsault, J. Alvarez, R. Sitia, T. Pozzan and R. Rizzuto (1995). "Monitoring dynamic changes in free Ca2+ concentration in the endoplasmic reticulum of intact cells." EMBO J **14**(22): 5467-5475.

Muller, M. P., H. Peters, J. Blumer, W. Blankenfeldt, R. S. Goody and A. Itzen (2010). "The Legionella effector protein DrrA AMPylates the membrane traffic regulator Rab1b." Science **329**(5994): 946-949.

Neunuebel, M. R., Y. Chen, A. H. Gaspar, P. S. Backlund, A. Yergey and M. P. Machner (2011). "De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila." Science **333**(6041): 453-456.

Page, R. and W. Peti (2016). "Toxin-antitoxin systems in bacterial growth arrest and persistence." Nat Chem Biol **12**(4): 208-214.

Palanivelu, D. V., A. Goepfert, M. Meury, P. Guye, C. Dehio and T. 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.

Pandey, D. P. and K. Gerdes (2005). "Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes." Nucleic Acids Res **33**(3): 966-976.

Pang, Y. L., K. Poruri and S. A. Martinis (2014). "tRNA synthetase: tRNA aminoacylation and beyond." Wiley Interdiscip Rev RNA **5**(4): 461-480.

Pedersen, K., S. K. Christensen and K. Gerdes (2002). "Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins." Mol Microbiol **45**(2): 501-510.

Perera, L. A., S. Preissler, N. R. Zaccai, S. Prévost, J. M. Devos, M. Haertlein and D. Ron (2021). "Structures of a deAMPylation complex rationalise the switch between antagonistic catalytic activities of FICD (14/96/109)." bioRxiv: 2021.2004.2020.440599.

Perera, L. A., C. Rato, Y. Yan, L. Neidhardt, S. H. McLaughlin, R. J. Read, S. Preissler and D. Ron (2019). "An oligomeric state-dependent switch in the ER enzyme FICD regulates AMPylation and deAMPylation of BiP." EMBO J **38**(21): e102177.

Pieles, K., T. Glatter, A. Harms, A. Schmidt and C. Dehio (2014). "An experimental strategy for the identification of AMPylation targets from complex protein samples." Proteomics **14**(9): 1048-1052.

Postow, L., N. J. Crisona, B. J. Peter, C. D. Hardy and N. R. Cozzarelli (2001). "Topological challenges to DNA replication: conformations at the fork." Proc Natl Acad Sci U S A **98**(15): 8219-8226.

Preissler, S., J. E. Chambers, A. Crespillo-Casado, E. Avezov, E. Miranda, J. Perez, L. M. Hendershot, H. P. Harding and D. Ron (2015). "Physiological modulation of BiP activity by trans-protomer engagement of the interdomain linker." Elife **4**: e08961.

Preissler, S., C. Rato, R. Chen, R. Antrobus, S. Ding, I. M. Fearnley and D. Ron (2015). "AMPylation matches BiP activity to client protein load in the endoplasmic reticulum." Elife **4**: e12621.

Preissler, S., C. Rato, L. Perera, V. Saudek and D. Ron (2017). "FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP." Nat Struct Mol Biol **24**(1): 23-29.

Preissler, S., L. Rohland, Y. Yan, R. Chen, R. J. Read and D. Ron (2017). "AMPylation targets the rate-limiting step of BiP's ATPase cycle for its functional inactivation." Elife **6**.

Reece, R. J. and A. Maxwell (1991). "DNA gyrase: structure and function." Crit Rev Biochem Mol Biol **26**(3-4): 335-375.

Ribas de Pouplana, L. and P. Schimmel (2001). "Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem." Cell **104**(2): 191-193.

Roy, C. R. and S. Mukherjee (2009). "Bacterial FIC Proteins AMP Up Infection." Sci Signal **2**(62): pe14.

Sanyal, A., A. J. Chen, E. S. Nakayasu, C. S. Lazar, E. A. Zbornik, C. A. Worby, A. Koller and S. Mattoo (2015). "A novel link between Fic (filamentation induced by cAMP)-mediated adenylylation/AMPylation and the unfolded protein response." J Biol Chem **290**(13): 8482-8499.

Schmelz, S. and J. H. Naismith (2009). "Adenylate-forming enzymes." Curr Opin Struct Biol **19**(6): 666-671.

Schulein, R., P. Guye, T. A. Rhomberg, M. C. Schmid, G. Schroder, A. C. Vergunst, I. Carena and C. Dehio (2005). "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**(3): 856-861.

Schvartzman, J. B., M. L. Martinez-Robles, P. Hernandez and D. B. Krimer (2013). "The benefit of DNA supercoiling during replication." Biochem Soc Trans **41**(2): 646-651.

Short, F. L., X. Y. Pei, T. R. Blower, S. L. Ong, P. C. Fineran, B. F. Luisi and G. P. Salmond (2013). "Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot." Proc Natl Acad Sci U S A **110**(3): E241-249.

Sissi, C. and M. Palumbo (2010). "In front of and behind the replication fork: bacterial type IIA topoisomerases." Cell Mol Life Sci **67**(12): 2001-2024.

Stadtman, E. R. (2001). "The story of glutamine synthetase regulation." J Biol Chem **276**(48): 44357-44364.

Stanger, F. V., B. M. Burmann, A. Harms, H. Aragao, A. Mazur, T. Sharpe, C. Dehio, S. Hiller and T. Schirmer (2016). "Intrinsic regulation of FIC-domain AMP-transferases by oligomerization and automodification." Proc Natl Acad Sci U S A **113**(5): E529-537.

Stanger, F. V., T. A. P. de Beer, D. M. Dranow, T. Schirmer, I. Phan and C. Dehio (2017). "The BID Domain of Type IV Secretion Substrates Forms a Conserved Four-Helix Bundle Topped with a Hook." Structure **25**(1): 203-211.

Stanger, F. V., C. Dehio and T. Schirmer (2014). "Structure of the N-terminal Gyrase B fragment in complex with ADPPi reveals rigid-body motion induced by ATP hydrolysis." PLoS One **9**(9): e107289.

Stanger, F. V., A. Harms, C. Dehio and T. Schirmer (2016). "Crystal Structure of the Escherichia coli Fic Toxin-Like Protein in Complex with Its Cognate Antitoxin." PLoS One **11**(9): e0163654.

Tan, Y. and Z. Q. Luo (2011). "Legionella pneumophila SidD is a deAMPylase that modifies Rab1." Nature **475**(7357): 506-509.

Thisted, T., N. S. Sorensen, E. G. Wagner and K. Gerdes (1994). "Mechanism of post-segregational killing: Sok antisense RNA interacts with Hok mRNA via its 5'-end single-stranded leader and competes with the 3'-end of Hok mRNA for binding to the mok translational initiation region." EMBO J **13**(8): 1960-1968.

Unterholzner, S. J., B. Poppenberger and W. Rozhon (2013). "Toxin-antitoxin systems: Biology, identification, and application." <u>Mob Genet Elements</u> **3**(5): e26219.

Utsumi, R., Y. Nakamoto, M. Kawamukai, M. Himeno and T. Komano (1982). "Involvement of cyclic AMP and its receptor protein in filamentation of an Escherichia coli fic mutant." <u>J Bacteriol</u> **151**(2): 807-812.

Veyron, S., G. Oliva, M. Rolando, C. Buchrieser, G. Peyroche and J. Cherfils (2019). "A Ca(2+)-regulated deAMPylation switch in human and bacterial FIC proteins." <u>Nat Commun</u> **10**(1): 1142.

Veyron, S., G. Peyroche and J. Cherfils (2018). "FIC proteins: from bacteria to humans and back again." <u>Pathog Dis</u> **76**(2).

Wigley, D. B., G. J. Davies, E. J. Dodson, A. Maxwell and G. Dodson (1991). "Crystal structure of an N-terminal fragment of the DNA gyrase B protein." <u>Nature</u> **351**(6328): 624-629.

Witz, G. and A. Stasiak (2010). "DNA supercoiling and its role in DNA decatenation and unknotting." <u>Nucleic Acids Res</u> **38**(7): 2119-2133.

Woolfolk, C. A. and E. R. Stadtman (1967). "Regulation of glutamine synthetase. 3. Cumulative feedback inhibition of glutamine synthetase from Escherichia coli." <u>Arch Biochem Biophys</u> **118**(3): 736-755.

Worby, C. A., S. Mattoo, R. P. Kruger, L. B. Corbeil, A. Koller, J. C. Mendez, B. Zekarias, C. Lazar and J. E. Dixon (2009). "The fic domain: regulation of cell signaling by adenylylation." <u>Mol Cell</u> **34**(1): 93-103.

Xiao, J., C. A. Worby, S. Mattoo, B. Sankaran and J. E. Dixon (2010). "Structural basis of Fic-mediated adenylylation." <u>Nat Struct Mol Biol</u> **17**(8): 1004-1010.

Xu, Y., P. D. Carr, S. G. Vasudevan and D. L. Ollis (2010). "Structure of the adenylylation domain of E. coli glutamine synthetase adenylyl transferase: evidence for gene duplication and evolution of a new active site." <u>J Mol Biol</u> **396**(3): 773-784.

Yamaguchi, Y., J. H. Park and M. Inouye (2011). "Toxin-antitoxin systems in bacteria and archaea." <u>Annu Rev Genet</u> **45**: 61-79.

Yarbrough, M. L., Y. Li, L. N. Kinch, N. V. Grishin, H. L. Ball and K. Orth (2009). "AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling." <u>Science</u> **323**(5911): 269-272.

Yarbrough, M. L. and K. Orth (2009). "AMPylation is a new post-translational modiFICation." <u>Nat Chem Biol</u> **5**(6): 378-379.

5   Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors Prof. Tilman Schirmer and Prof. Christoph Dehio for giving me the opportunity to work in a very close collaboration between these two groups on projects combining my interest in structural biology and microbiology. Their guidance, incredible knowledge, expertise, as well as their mentorship, patience and help in research and writing is highly appreciated.

I would like to thank Prof. Dirk Bumann, who accepted to be part of my PhD committee and for the helpful discussions and feedback during committee meetings.

I would also like to thank my external supervisor Prof. Markus Wiederstein for accepting to be in my PhD committee and especially for being part of my scientific career since my bachelor studies at the University of Salzburg.

A big thank you to all present and past members of the Schirmer lab and Dehio lab, with whom I have enjoyed working over the last years, especially the current and past members of office 307 (Markus Huber, Johanna Rueher, Elia Agustoni, Raphael Dias Teixeira, Firas Fadel, Nikolaus Dietz) for great discussions in terms of science and a lot of fun times and laughter outside of science.

A very special and huge thank you to Isabel Sorg for all her unconditional support, feedback, and discussions during my PhD, especially in all the meetings during the writing process and for critical reading of my thesis.

I would also like to thank my best friends Florentina Gareiß and Bernadette Brenner, who have been by my side for more than a decade, who I can count on every day and night despite the geographical distance and who are like sisters to me. I'll be eternally grateful for having them in my life.

I am incredible grateful to Anne Grahl for being an incredible colleague and for becoming one of my best friends during my time at Biozentrum. I'm grateful for all the discussions and advice for every scientific or non-scientific problem I encountered.

A special, huge thank you goes to Runi, Tuba, Eva and Sora-chan who I met online a few years ago because of our shared interest in music and who soon became an incredible important part of my daily life and best friends. Words cannot describe how I grateful I am for their unconditional friendship through good and bad times. I would not be here today without them.

Music has always played an important role in my life and I think I would not be at this point in my life without the motivation and inspiration I got from several music groups I am interested in and who are my support in every situation of my life.
Therefore, a huge thanks to the members of Exile Tribe, Sandaime J Soul Brothers III, Fantastics, The Rampage, EXILE, Solidemo, Da-iCE, AAA, X4, SuperM, NCT and WayV
My biggest gratitude goes to the members of Generations, SHINee, EXO, EXO-CBX, EXO-SC and TVXQ, who really pushed me through my PhD thesis, who helped me dealing with some really challenging situations in my life, who make me laugh and motivate me to give my best every day.

I also want to thank Heidi and Felix Leumann, who welcomed me in Basel nearly five years ago, made it a second home and became a 2nd family for me.

Lastly, I want to express my sincere gratitude to my family, especially my grandparents and my parents for their unconditional support and for their believe in me pursuing this path. Everything I have achieved today would not be possible without them and I am forever indebted to them for giving me all opportunities in life.

6 Curriculum Vitae

Personal Details

Name:	Stefanie Tamegger
Address:	Hungerbachweg 13, 4125 Riehen, Basel-Stadt, Switzerland
Phone:	0041 786547206
E-Mail:	stefanie.tamegger@aon.at
Nationality:	Austria
Gender:	Female
Date of birth:	6th of March, 1988
Place of birth:	Raigarh, India

Work experience

07/2021 – 12/2021	PostDoc in the research group of Christoph Dehio (Focal Area of Infection Biology), Biozentrum, University of Basel, Switzerland
03/2020 – 03/2021	Instructor for the Block Course Microbiology and Immunology, Biozentrum, University of Basel, Switzerland
10/2017 – 10/2019	Instructor for the Block Course Structural Biology and Biophysics, Biozentrum, University of Basel, Switzerland
06/2015 – 01/2016	Instructor for the practical course bachelor module In Silico Biology, University of Salzburg, Austria
11/2014 – 11/2015	Instructor for the practical course bachelor module Molecular Biology, University of Salzburg, Austria
02/2015	Workgroup internship in the research group of Prof. Dr. Silja Weßler, Division of Microbiology, Department of Molecular Biology, University of Salzburg, Austria
03/2014	Workgroup internship in the research group of Mag. Dr. Nikolaus Bresgen, Division of Genetics,

Department of Cell Biology, University of Salzburg, Austria

Education

09/2016 – 06/2021	PhD studies of Structural Biology at the Biozentrum, University of Basel, Switzerland PhD thesis: *Structure-function studies on FIC-mediated AMPylation and deAMPylation by class I Fic proteins* (supervisor: Prof. Dr. Tilman Schirmer, Prof. Dr. Christoph Dehio)
10/2013 – 04/2016	Master studies of Biology at the University of Salzburg, Austria Master thesis: *The translocation and phosphorylation of Helicobacter pylori CagA in different epithelial and non-epithelial cells* (supervisor: Prof. Dr. Silja Wessler)
02/2010 – 07/2013	Bachelor studies of Biology at the University of Salzburg, Austria Bachelor thesis: *The pathogenicity factors of Helicobacter pylori* (supervisor: Prof. Dr. Silja Wessler)
10/2007 – 01/2010	Bachelor studies of Biology and Teacher accreditation programme of Biology and English at the University of Vienna, Austria
09/2002 – 06/2007	Bundesbildungsanstalt für Sozialpädgogik, St. Pölten, Austria
09/1998 – 06/2002	BRG/BORG St. Pölten with focus on musical education, St. Pölten, Austria

Conferences

04/2019	EMBO Workshop on Toxin-antitoxin systems in bacteria – Poster Presentation

Additional

qualifications

Languages	German (mother tongue)
	English (C1)
	Spanish (A2)
Computer skills	Good in command of Mac OS X and Microsoft OfficeTM Tools for Mac OS (Word, Excel and PowerPoint)
	Good in command of graphic editor software: Adobe Photoshop CC, Adobe Illustrator CC
	Good in command of software for interactive visualization and molecular modelling: UCSF Chimera, Jmol, PyMol