The EMBO Journal (2013) 32, 1469–1477
www.embojournal.org

TRANSPARENT PROCESS

THE EMBO JOURNAL

Structure of the *Legionella* effector AnkX reveals the mechanism of phosphocholine transfer by the FIC domain

Valérie Campanacci[1], Shaeri Mukherjee[2], Craig R Roy[2] and Jacqueline Cherfils[1,*]

[1]Laboratoire d'Enzymologie et Biochimie Structurales, Centre de Recherche de Gif, CNRS, Gif-sur-Yvette, France and [2]Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT, USA

The FIC motif and the eukaryotic-like ankyrin repeats are found in many bacterial type IV effectors, yet little is known about how these domains enable bacteria to modulate host cell functions. Bacterial FIC domains typically bind ATP and transfer adenosine monophosphate moiety onto target proteins. The ankyrin repeat-containing protein AnkX encoded by the intracellular pathogen *Legionella pneumophila* is unique in that its FIC domain binds to CDP-choline and transfers a phosphocholine residue onto proteins in the Rab1 GTPase family. By determining the structures of unbound AnkX and AnkX with bound CDP-choline, CMP/phosphocholine and CMP, we demonstrate that the orientation of substrate binding in relation to the catalytic FIC motif enables this protein to function as a phosphocholinating enzyme rather than a nucleotidyl transferase. Additionally, the structure reveals that the ankyrin repeats mediate scaffolding interactions that resemble those found in protein–protein interactions, but are unprecedented in intramolecular interactions. Together with phosphocholination experiments, our structures unify a general phosphoryl transferase mechanism common to all FIC enzymes that should be conserved from bacteria to human.

The EMBO Journal (2013) **32,** 1469–1477. doi:10.1038/emboj.2013.82; Published online 9 April 2013
Subject Categories: microbiology & pathogens; structural biology
Keywords: ankyrin repeats; FIC domain; *Legionella*; phosphocholination; Rab GTPase

Introduction

Many intracellular pathogens use specialized secretion systems to deliver proteins into host cells called effectors, which can have biochemical activities that modulate the cytoskeletal and membrane transport pathways of the infected host cell (reviewed in Cossart and Roy, 2010). The FIC domain (filamentation induced by cyclic adenosine monophosphate moiety (AMP)) and the eukaryotic-like ankyrin repeat domain are modules contained in many different bacterial effectors (reviewed in Roy and Mukherjee, 2009; Voth, 2011). Ankyrin repeats consist of strings of helix-helix-turn units, and in eukaryotes are among the most common protein–protein interaction motifs (reviewed in Li *et al*, 2006). It is thought that ankyrin repeats in bacterial effectors will serve as interaction modules involved in the targeting of host factors, but there is no clear evidence demonstrating such a role, and the function of most bacterial Ank proteins remains unknown.

FIC domains represent another homology region found in multiple bacterial effectors. Work on the *Vibrio parahaemolyticus* type III effector VopS showed that the FIC domain in this protein had an enzymatic activity that resulted in the transfer of an AMP onto host Rho GTPases, resulting in the disruption of the actin cytoskeleton (Yarbrough *et al*, 2009). This adenylylation reaction mediated by VopS was termed as AMPylation. The protein IbpA from the respiratory pathogen *Histophilus somni* was also shown to have an activity that mediated adenylylation of Rho family GTPases to disrupt their function (Kinch *et al*, 2009; Worby *et al*, 2009). Thus, the FIC domain, which is conserved outside the prokaryotic kingdom (Kinch *et al*, 2009; Worby *et al*, 2009), was thought to have a conserved function in mediating the post-translational addition of AMP onto hydroxyl-containing residues of target proteins (Kinch *et al*, 2009; Roy and Mukherjee, 2009).

Mechanisms for ATP cleavage and AMP transfer have been proposed recently based on structural and biochemical analysis of representative bacterial FIC domain-containing enzymes, which include VopS (Luong *et al*, 2010), IbpA bound to AMPylated Cdc42 (Xiao *et al*, 2010), *Bartonella* BepA with bound pyrophosphate (PPi) (Palanivelu *et al*, 2011) and *Neisseria* NmFic bound to AMP and ATP (Engel *et al*, 2012). These studies highlighted the conserved fold of the FIC domain and the catalytic role of its HPFx[D/E]GN[G/K]R signature in cleaving the Pα-Oαβ bond of ATP (Luong *et al*, 2010; Xiao *et al*, 2010; Palanivelu *et al*, 2011), as well as the contribution of adjacent elements in binding the adenine (Xiao *et al*, 2010; Engel *et al*, 2012) and possibly in regulating the AMPylation activity (Engel *et al*, 2012). It should be noted that although all characterized FIC enzymes seem to have a narrow protein target specificity, most undergo auto-modification (Kinch *et al*, 2009; Xiao *et al*, 2010; Engel *et al*, 2012; Feng *et al*, 2012; Goody *et al*, 2012). Although the significance of auto-modification remains unknown, it has been proposed to represent an intermediate in the phosphoryl transfer reaction mediated by the FIC domain or possibly be important in regulating the activity of the enzyme (Luong *et al*, 2010; Goody *et al*, 2012).

The paradigm that all FIC domain-containing effectors function as AMPylating enzymes was challenged by the recent discovery that *Legionella* type IV effector protein AnkX uses

*Corresponding author. Laboratoire d'Enzymologie et Biochimie Structurales, CNRS, 1, Avenue de la Terrasse, Gif-sur-Yvette 91198, France. Tel.: +33 169823492; Fax: +33 169823129;
E-mail: cherfils@lebs.cnrs-gif.fr

Received: 9 September 2012; accepted: 12 March 2013; published online: 9 April 2013

CDP-choline as a substrate to transfer phosphocholine onto serine or threonine in conserved positions in the switch II region of Rab1 family GTPases (Mukherjee *et al*, 2011). Additionally, it has been shown that *Xanthomonas* AvrAC can use UTP to transfer UMP onto plant receptor-like kinases (Feng *et al*, 2012). Thus, FIC domains are able to catalyse chemical reactions others than AMPylation.

All of the FIC domain-containing effectors that have been characterized mediate a phosphoryl transfer reaction requiring a nucleotide-based substrate. What makes the reaction mediated by the FIC domain in AnkX unique is that the transferred moiety is not the nucleotide monophosphate (CMP) but is instead the phosphocholine moiety. AnkX is also unique in that it carries multiple ankyrin repeats in addition to the FIC domain on a single polypeptide. When expressed in mammalian cells, AnkX modifies Rab1 and Rab35 proteins to disrupt membrane transport processes mediates by these GTPases, and this activity requires a conserved histidine residue found in the catalytic site of the FIC motif and several predicted ankyrin repeat domains located nearby in the amino-acid sequence (Pan *et al*, 2008; Mukherjee *et al*, 2011). The defects in membrane transport are likely due to impairments in the interaction of phosphocholinated Rab1 proteins with RabGDIs, RabGEFs, and cellular effectors (Mukherjee *et al*, 2011; Goody *et al*, 2012).

Here, we report the crystallographic structures of apo-AnkX and of AnkX bound to its substrate, CDP-choline, and to its products, CMP and phosphocholine. These structural snapshots, together with further analysis of the phosphocholination reaction, provide important insight into how FIC and ankyrin domains have been co-opted by bacteria, and suggest a general model for the substrate recognition and phosphoryl transferase mechanism common to FIC enzymes from bacteria to human.

Results

The crystal structure of AnkX reveals that the ankyrin repeats scaffold constrains the FIC and substrate-binding domains

Limited proteolysis studies indicated that the smallest protease-resistant domain that retained Rab1 phosphocholination activity contained residues 1–484 in AnkX, so this region was selected for protein crystallization studies. We solved the crystal structures of *L. pneumophila* $AnkX_{1–484}$ in the apo form, in complex with its substrate, CDP-choline (using the inactive $AnkX^{H229A}$ mutant), and in complexes with CMP and phosphocholine and with CMP alone (Figure 1A; Table I). AnkX has an atypical FIC domain comprised of two sub-domains, of which the second carries a canonical FIC motif (Supplementary Figure 1A and B). The FIC sub-domains are separated by a 70-residue domain, which inserts within a β-hairpin that is conserved in all known FIC domain structures (Supplementary Figure 1A and C). Unlike the β-hairpin of the FIC domain of IbpA,

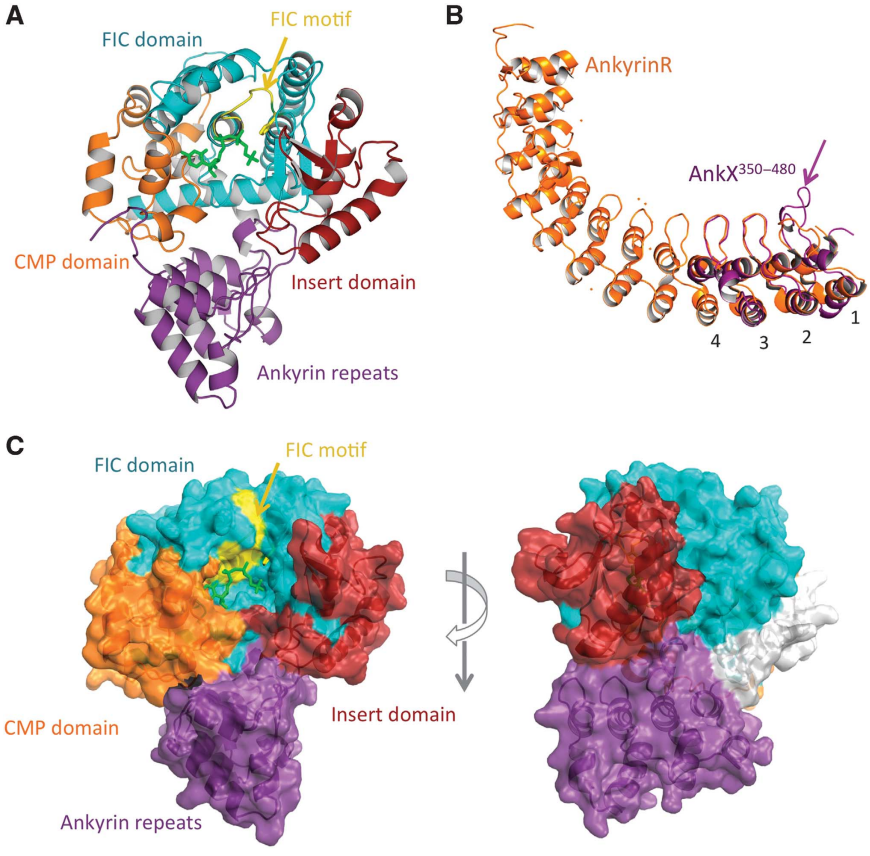

Figure 1 Structure of *L. pneumophila* AnkX. (**A**) Structure of AnkX with bound CDP-choline (in green). The sequences of the CMP binding, FIC, insert and ankyrin repeat domains are given in Supplementary Figure 1A. (**B**) Overlay of the four ankyrin repeats of $AnkX_{1–484}$ (in magenta) to an archetypal ankyrin repeats protein (AnkyrinR, in orange Michaely *et al*, 2002). The longer loop in the first ankyrin motif is shown by an arrow. (**C**) The ankyrin repeats (in magenta) form extensive intramolecular contacts with the CMP-binding domain (in orange), the FIC domain (in cyan), and the insert domain (in red). Intramolecular contacts are shown in Supplementary Figure 3.

Table I Data collection and refinement statistics

	$AnkX_{1-484}$	$AnkX_{1-484}$	AnkX/CMP	AnkX/CMP/ phosphocholine	$AnkX^{H229A}$/ CDP-choline
Data collection	Form I	Form II	Form III	Form IV	
Phasing method	MR-SAD	SAD	MR	MR	MR
Space group	$P2_12_12_1$	$P2_1$	$P222_1$	$P2_12_12$	$P2_1$
Cell constants (Å)	55.3, 91.6, 239.1	59.1, 198.4, 110.0 $\beta = 98.5^\circ$	58.3, 103.6, 224.1	84.3, 122.3, 55.0	66.1, 122.1, 75.7 $\beta = 107.3^\circ$
Molecule/asu	2	4	2	1	2
Wavelength (Å)	0.9792	0.9790	0.9801	0.9801	0.9801
Resolution range[a] (Å)	44.0–3.14 (3.33–3.14)	48.63–3.60 (3.69–3.60)	46.28–2.60 (2.76–2.60)	43.10–2.54 (2.68–2.54)	43.89–2.55 (2.70–2.55)
R_{sym}[a,b] (%)	9.2 (52.3)	7.6 (68.3)	10.5 (101.1)	10.1 (74.7)	9.1 (83.7)
$I/\sigma I$[a]	11.33 (2.59)	11.06 (1.99)	16.16 (2.30)	13.42 (2.50)	12.02 (1.75)
Completeness[a] (%)	98.9 (96.9)	99.2 (92.1)	99.6 (98.8)	99.7 (98.8)	98.5 (97.4)
Redundancy[a]	3.8 (3.8)	4.9 (3.4)	7.3 (7.4)	5.8 (5.7)	4.7 (4.7)
Refinement					
Traced residues					
Chain A	5–482		1–328, 331–477	6–85, 96–484	5–484
Chain B	5–86, 96–477		1–85, 90–479		6–484
No. of reflections	20 682		42 690	19 367	36 935
R_{work}/R_{free} (%)	21.18/25.64		19.58/23.26	19.43/25.71	19.23/24.32
No. of atoms/waters	7434		7582/73	3738/54	7643/114
Mean B-factor (Å^2)					
Protein	43.1		65.8	55.9	69.5
Ligands	—		59.7 (CMP:A)	59.7 (CMP)	59.3 (CDC:A)
	—		62.5 (CMP:B)	99.4 (PC)	52.31 (CDC:B)
r.m.s.d.					
Bond lengths (Å)	0.011		0.010	0.010	0.010
Bond angle (deg)	1.58		1.18	1.21	1.24

[a]Values in parentheses refer to the highest resolution shell.
[b]$R_{sym} = \sum_h \sum_i |I_{hi} - (I_h)| / \sum_h \sum I(I_h)$, where I_i is the ith observation of reflection h and (I_h) is the weighted average intensity for all observations l of reflection h.

which forms the binding site for AMPylated Cdc42 (Xiao *et al*, 2010), the β-hairpin of AnkX is covered by the insert domain and is unavailable for binding Rab1, suggesting that Rab1 has a different binding modality (Supplementary Figure 2). Next to the FIC motif, AnkX features a unique domain comprised of the N-terminal 50 residues and of 40 residues located downstream of the FIC domain, which binds the CMP moiety of CDP-choline (Supplementary Figure 1A and D). Residues beyond Pro351 form four consecutive ankyrin repeats (Figure 1A and B), two of which were not predicted using standard methods (Pan *et al*, 2008). All domains establish multiple interactions with each other, yielding a close-packed structure (Figure 1C). Remarkably, the ankyrin repeats are involved in a large interface with the CMP binding, FIC, and insert domain (about 2200 Å^2 buried surface area, intramolecular contacts in Supplementary Figure 3), which resembles the interfaces found in protein–protein interactions but is seen for the first time in an intramolecular interaction. Consistent with this unprecedented role of the ankyrin repeats in constraining the structure of AnkX, a recombinant AnkX construct lacking the ankyrin repeats did not yield soluble expression in *E. coli*. Thus, our structures reveal that ankyrin repeats are not only protein–protein interaction domains, but can also contribute to intramolecular interactions.

The orientation of CDP-choline binding to AnkX provides a mechanism for transfer of phosphocholine by the FIC motif

It was unknown why AnkX mediates the transfer of phosphocholine rather than CMP given that it carries the canonical catalytic FIC motif that functions in nucleotide transfer in AMPylating toxins. Our structures show that the AnkX FIC motif has the same conformation as in known AMPylating FIC enzymes (Supplementary Figure 4A). CDP-choline shares a nucleoside diphosphate group with ATP, but its binding to AnkX with the orientation of ATP as seen in AMPylating FIC domains would result in a head-to-tail chemical reaction with inversion of the transferred and leaving groups. Unprocessed CDP-choline was captured in complex with the inactive $AnkX^{H229A}$ mutant (Figure 2A), revealing that choline, rather than cytidine, overlays the adenosine moiety of ATP (Figure 2B). The positive charge of the trimethylammonium group of the choline moiety is recognized by the electron-rich aromatic ring of Phe107 in the β-hairpin and the carboxyls of two acidic residues, Asp265 and Glu226 in the FIC motif-containing sub-domain (Figure 2C). Phe107, Ile109, and Asp265 are located in the vicinity of the choline moiety and should exclude the binding of the adenine base through steric hindrance (Figure 2B), which would explain why AnkX cannot function as an AMPylating enzyme. The other unique feature of the active site pocket of AnkX is that it snugly fits the cytidine group, as shown by its well-defined structure in the complexes containing $AnkX^{H229A}$/CDP-choline (Figure 2A), AnkX/CMP, and AnkX/CMP/phosphocholine. Specific recognition is mediated by the CMP domain unique to AnkX, which forms numerous interactions with the base (Tyr 41, Arg 44, and Cys 48) and the sugar (Asp28 and Arg30) (Figure 2C). Tyr41, notably, stacks its aromatic ring on the cytosine ring of CDP-choline. Mutations in the predicted choline-binding residues (Phe107Gly, Glu226Ala, and Asp265Ala) and of a cytosine-binding residue (Tyr41Ala), all of which are located outside the FIC motif, impaired Rab1 phosphocholination *in vitro* (Figure 3A; Supplementary Figure 5A), consistent with these residues playing an important role in enzyme function. The

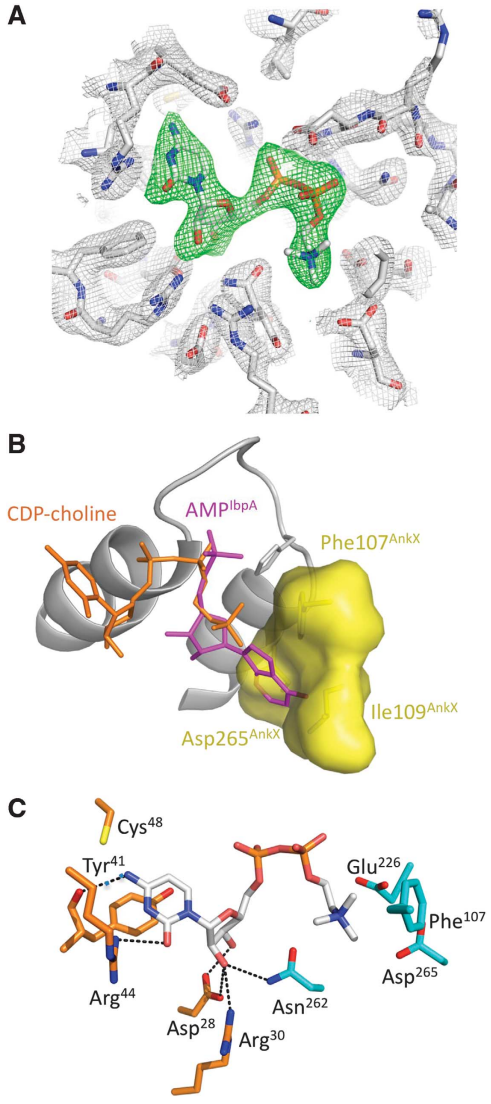

Figure 2 Structure of the CDP-choline binding site of AnkX. (**A**) Fo-Fc electron density omit map of CDP-choline bound to AnkX^{H229A} contoured at 2.5σ. (**B**) Superposition of CDP-choline bound to AnkX (in orange) and AMP bound to cdc42 in the IbpA-cdc42 complex (in magenta, PDB 3N3V). Note that the AnkX active site (yellow surface) cannot accommodate the base of AMP. (**C**) Interactions of the CMP domain (in orange) and FIC domain (in cyan) residues with the cytidine and choline moieties of CDP-choline. Hydrogen bonds are shown in dotted lines.

binding site of CDP-choline to AnkX shares a common use of aromatic and negatively charged residues to recognize the choline moiety with CTP-phosphocholine cytidylyltransferases from bacteria (Kwak *et al*, 2002) and mammals (Lee *et al*, 2009), but is otherwise completely different from that of these enzymes (Supplementary Figure 6). Thus, these structural snapshots provide molecular details on how AnkX uses its CMP domain and FIC domain to uniquely recognize CDP-choline as a substrate and exclude ATP, which makes AnkX a phosphocholine transferase exclusively.

AnkX catalyses the same phosphoryl transfer reaction as AMPylating FIC domains

The structures of substrate- and product-bound AnkX (Figure 4A and B) suggest that the catalytic mechanism of AnkX phosphoryl transfer would be similar to that of AMPylating FIC enzymes. FIC motif-based overlay of AnkX/CDP-choline to IbpA/AMPylated Cdc42 shows that the β-phosphate of CDP-choline superimposes with the α-phosphate of AMP (Figure 4C), indicating that the scissile bond in CDP-choline is the Oαβ-Pβ bond, rather than the Pα-Oαβ of ATP that is cleaved by AMPylating enzymes. Furthermore, the Oαβ-Pβ bond of CDP-choline is exactly co-linear with the phosphotyrosine bond of AMPylated Cdc42 with inversion of configuration of the phosphate (Figure 4C), as would be expected if both enzymes would have the same catalytic mechanisms. Consistently, mutations of conserved residues in the FIC motif in AnkX (His229Ala, Asp233Ala, and Arg237Glu) all severely impaired both auto-phosphocholination and Rab1 phosphocholination *in vitro* (Figure 3B; Supplementary Figure 5B). Thus, a major consequence arising from CDP-choline binding head-to-tail with respect to ATP is that AMPylation and phosphocholination can be achieved through the same chemical reaction, with conservation of the positions of the transferred and leaving groups relative to the FIC motif and with the same use of catalytic residues.

The catalytic mechanism of CDP-choline cleavage by AnkX

Our structures reveal a diphosphate-containing substrate both before and after its cleavage by an FIC domain, allowing for unprecedented insight into the catalytic mechanism. In these structures, phosphoryl transfer to the hydroxyl-containing serine residue in Rab1 was mimicked by a water molecule. We observe that the CDP-choline-binding site has the same conformation in apo-AnkX, AnkX^{H229A}/CDP-choline, and AnkX/CMP/phosphocholine (Supplementary Figure 4B), from which it can be safely assumed that the reaction does not require a disorder-to-order transition or an induced fit. This suggests that the transition state of CDP-choline cleavage is structurally close to the ground states observed in the crystals and can be analysed from these structures. An important contribution to the efficiency of biological phosphoryl transfer reactions is the stabilization of partial negative charges that develop at the transition state, which is provided by positive charges from enzyme active sites (reviewed in Lassila *et al*, 2011). AnkX has only one positively charged residue that interacts with the α-phosphate of both the substrate and leaving group, Arg237 from the FIC motif, and none that interacts with the β-phosphate (Figure 4A and B). Additional stabilization could stem from a cationic metal, one candidate being a Mg^{2+} cation. As shown in Figure 3C, auto-phosphocholination was impaired when no cationic metal was present in the mixture, and was recovered by the addition of Mg^{2+}, indicating that the reaction requires a cationic ion. A Mg^{2+} ion has been observed in the structure of the BepA/pyrophosphate complex, where it bridges the phosphates of the ligand (Palanivelu *et al*, 2011). Superposition of this structure with CDP-choline-bound AnkX suggests that Mg^{2+} is coordinated by the conserved FIC motif aspartate (Figure 4D), which would be consistent with the critical role of this residue despite the fact that it does not make direct contacts with CDP-choline.

An intriguing aspect of FIC enzymes, whose mechanism is currently debated, is their robust auto-modifications, which depend on the catalytic histidine and occurs regardless of whether they catalyse AMPylation, UMPylation, or phosphocholination reactions. One explanation that has been put

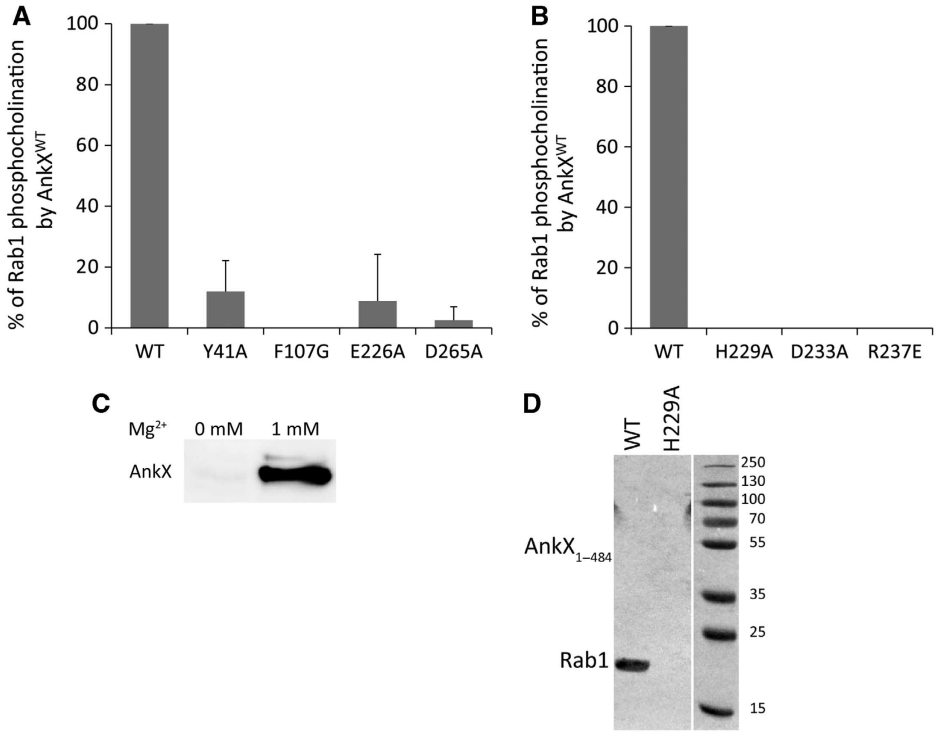

Figure 3 Analysis of Rab1 phosphocholination and AnkX auto-phosphocholination. (**A**) AnkX constructs carrying mutations in the CDP-choline binding site have impaired Rab1 phosphocholination. Rab1 phosphocholination is expressed as the relative percentage of Rab1 phosphocholination by wild-type AnkX. Data are from immunoblots from *in vitro* reactions that contained full-length His-tagged AnkX constructs and Rab1A in the presence of phosphocholination buffer. Blots were probed with an anti-PC antibody. Representative blots are shown in Supplementary Figure 5A. Errors bars represent the standard deviation based on three independent experiments. (**B**) AnkX constructs carrying mutations in the FIC motif have impaired Rab1 phosphocholination. Experimental conditions are as in Figure 3A. Representative blots are shown in Supplementary Figure 5B. Errors bars represent the standard deviation based on three independent experiments. (**C**) Mg^{2+} is necessary for phosphocholination. Full-length His-tagged AnkX was dialysed against a metal-free buffer prior to the experiment, and subsequently assessed for auto-phosphocholination without (left lane) or with (right lane) addition of 1 mM Mg^{2+} in the auto-phosphocholination buffer. (**D**) Auto-phosphocholinated residues are not located in the FIC domain. The $AnkX_{1–484}$ construct (lane 1) phosphocholinates Rab1 but is not auto-phosphocholinated. The $AnkX_{1–484}$ construct carrying the H229A mutation is shown as a control that Rab1 phosphocholination requires the FIC motif histidine (lane 2). Molecular weight markers (in kDa) are shown on the right.

forward is that the auto-modified species represents the accumulation of a phosphohistidine intermediate, which would indicate a ping-pong mechanism of phosphoryl transfer (Goody *et al*, 2012). Comparison of substrate-bound AnkX and product-bound IbpA identifies a conserved geometry for in-line nucleophilic attack of the transferred phosphorus (Figure 4C) that could not be structurally accommodated in a ping-pong mechanism, which would require that at least one of the nucleophilic attack is occurring from the opposite side of the phosphorus. We found that the $AnkX_{1–484}$ protein used in our crystallization studies was not auto-phosphocholinated *in vitro*, even though it retained the ability to phophocholinate Rab1 by a mechanism that required the catalytic histidine (Figure 3D; Supplementary Figure 5C). These data suggest that the auto-phosphocholinated AnkX species does not represent an intermediate in the phosphoryl transfer reaction and that the auto-phosphocholinated residues must be located outside of the enzymatic domain in $AnkX_{1–484}$. Thus, AnkX uses a concerted mechanism of phosphoryl transfer, in which in-line nucleophilic attack of the β-phosphate of CDP-choline mediated by the hydroxyl group of Ser79 in Rab1 is mimicked by a water molecule, according to the transition state summarized in Figure 4E.

Discussion

A large number of bacterial effector proteins contain ankyrin repeats and FIC domains, but understanding how these domains have been adapted for diverse functions remains a critical question. Our structures of AnkX provide new insight into how these domains function. These data show that the ankyrin repeats in AnkX have been co-opted by bacteria to constrain intramolecular interactions, in contrast with most eukaryotic ankyrin repeats where the exposed loops and flexible structure of the repeat functions as a versatile protein–protein interaction module (reviewed in Li *et al*, 2006). Sequence analysis predicts that AnkX may have up to 12 ankyrin repeats (Supplementary Figure 1A), in which the loops have more variable lengths than in eukaryotic ankyrin repeats, as exemplified by the loop of the first motif (Figure 1B). These repeats can be readily modelled by extending the first four repeats without clashes with the rest of the structure (Figure 1B), leaving open the possibility that they form additional interactions, possibly with Rab1 or other partners.

FIC domains have recently been classified as an evolutionary conserved family present in genomes from bacteria, archae, fungi, and metazoan (Kinch *et al*, 2009; Worby *et al*,

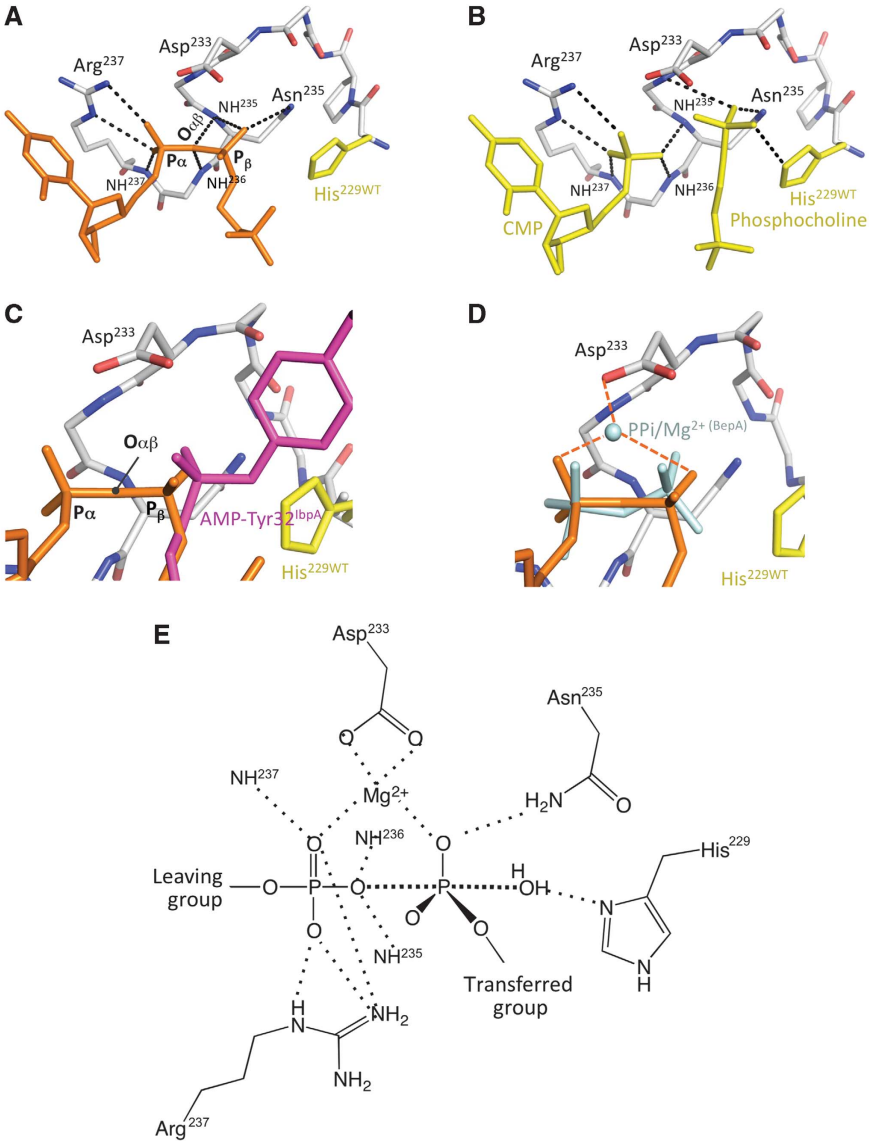

Figure 4 Structural basis for the phosphocholination reaction. (**A**) Interactions of the phosphates of CDP-choline with AnkX^{H229A}. His229 in the FIC motif is overlaid from wild-type AnkX. (**B**) Interactions of the phosphates of CMP and phosphocholine with wild-type AnkX. (**C**) The Oαβ-Pβ bond of CDP-choline aligns with the phosphotyrosine bond of Cdc42 in complex with IbpA. (**D**) Proposed interactions of Mg^{2+} with CDP-choline modelled from *Bartonella* BepA/PPi-Mg^{2+} (PDB 2JK8). PPi bound to BepA is overlaid (in light blue). (**E**) Model for the transition state of phosphoryl transfer by the conserved FIC motif. Hydrogen bonds are in dotted lines, partial bonds are in dashed lines. Superpositions in Figure 4A–D are based on the FIC motif.

2009). Characterizing what makes their *air de famille* and what defines their individual specificities is important to understand their roles in bacterial infections and eukaryotic regulatory functions. Our analysis of substrate- and product-bound AnkX, in perspective with previous analysis of AMPylating FIC proteins, shows that the FIC motif has a general role in catalysing the cleavage of diphosphate-containing substrates and transfer of the resulting phosphoryl-containing moiety onto an entering hydroxyl group, regardless of the specific nature of the substrate. Catalysis would be best explained by in-line nucleophilic attack by the entering hydroxyl group which is activated by the invariant histidine of the HPFx[D/E]GN[G/K]R signature, with stabilization of the developing negative charges by the invariant arginine and by a metal cation bound to the conserved acidic residue, according to the transition state depicted in Figure 4E. We propose that auto-modification, a side reaction reported in all characterized FIC enzymes, occurs using the same catalytic mechanism, but likely target residues located in poorly folded sequences outside of the catalytic region of the enzyme that can readily make it to the active site. This could explain why auto-AMPylation of the C-terminal helix of NmFIC increases upon mutations that render this helix flexible (Engel *et al*, 2012).

When the structure of CDP-choline bound to AnkX is compared with the structure of AMP bound to IbpA (Xiao *et al*, 2010) and ATP bound to NmFIC (Engel *et al*, 2012), the only AMPylating toxins where these ligands have been observed, it becomes clear that the transferred and leaving groups of diphosphate-containing substrates are specifically recognized by variable regions located outside the FIC motif at both ends of the active site. The AMP and phosphocholine

transferred groups bind near the β-hairpin of the FIC domain (Figure 5A), in which specific residues exclude illegitimate ligands (Figure 2B). The structures of AnkX in complex with CDP-choline and CMP also show for the first time how the leaving group is recognized by a native FIC domain, since ATP was observed only in an NmFIC mutant in which an extra C-terminal α-helix had been truncated (Xiao *et al*, 2010), whereas PPi bound to BepA (Palanivelu *et al*, 2011) likely mimicked the Pα-Pβ moiety of ATP rather than the leaving Pβ-Pγ group (see Figure 4D). The CMP domain in AnkX has highly variable counterparts in FIC proteins of known structures (Figure 5A and B). We propose that these extra structural elements determine the nature of the physiological substrate and/or exclude illegitimate ligands. FIC enzymes whose binding of the γ-phosphate of ATP is hindered by an extra α-helix, such as NmFIC, may thus have a physiological substrate other than ATP, a plausible alternative to their proposed auto-inhibition by this helix (Engel *et al*, 2012). According to this model, the fact that *Pseudomonas* AvrB, which is remotely related to FIC proteins (Kinch *et al*, 2009), binds ADP with the same orientation as the CDP moiety of CDP-choline to AnkX (Figure 5C; Supplementary Figure 7), would suggest that it could use ATP to transfer a phosphate group, consistent with its reported protein kinase activity (Desveaux *et al*, 2007).

In conclusion, these data suggest that structurally diverse FIC active sites assemble from the conserved phosphoryl transfer FIC motif and variable substrate-binding regions, suggesting that FIC-containing toxins and eukaryotic proteins have probably evolved to process diverse diphosphate-containing substrates. Thus, future studies on FIC domain-containing proteins should take into consideration that these enzymes can be adapted to transfer a diverse array of phorphoryl-containing moieties.

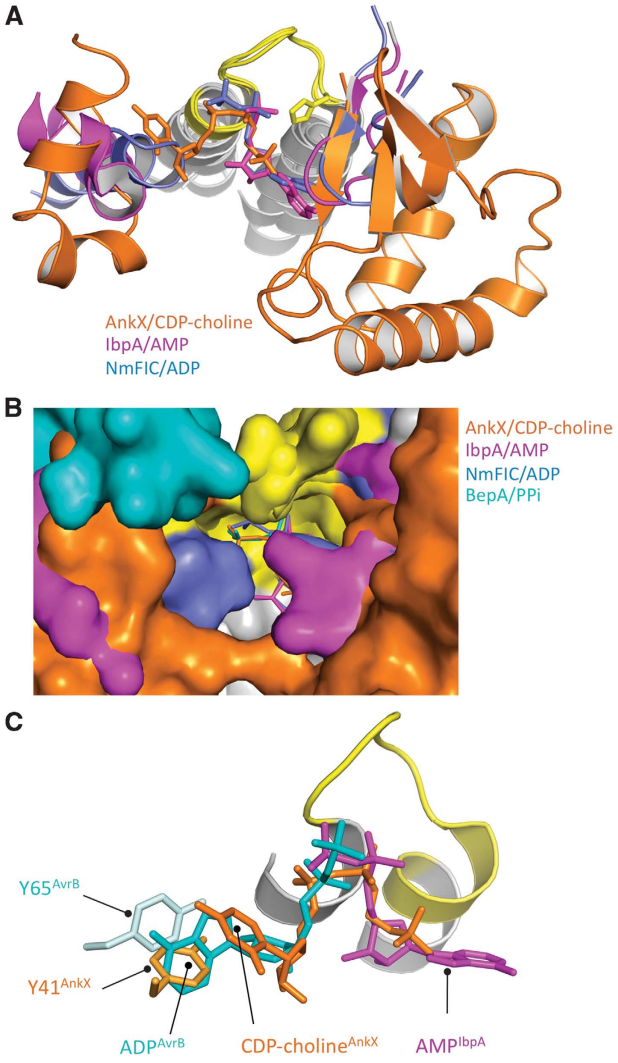

Figure 5 The active sites of FIC domains assemble from the conserved catalytic FIC motif and variable substrate-binding regions. (**A**) Superposition of active sites of phosphocholinating (AnkX, orange), AMPylating (IbpA, magenta) and unknown function (NmFic, blue and BepA, cyan) FIC enzymes. The FIC motif is shown in yellow and is in the same orientation in all views. Structural elements that recognize the leaving group are on the left, and the regions that recognize the transferred phosphoryl group are on the right. The latter regions are probably also involved in binding target proteins, as shown for the β-hairpin of IbpA (Xiao *et al*, 2010). (**B**) Overlaid surfaces of FIC enzymes active sites highlight the variety of their shapes and volumes. The FIC motif is shown in yellow. (**C**) Close-up view showing the common use of an aromatic residue by AnkX to bind the base ring of CDP-choline (in orange) and by the FIC-related AvrB protein (from PDB 2NUN) to bind ADP (in cyan). AMP bound to IbpA is overlaid (in magenta, PDB 3N3V).

Materials and methods

L. pneumophila AnkX subcloning, site-directed mutagenesis, expression, and purification

The *L. pneumophila* AnkX nucleotidic sequence encoding amino acids 1–484 was PCR amplified and cloned by Gateway recombination (Vincentelli *et al*, 2003) into the pDEST17 vector (Invitrogen). The resulting construct, $AnkX_{1-484}$, encoded an N-terminal fusion with a His_6 tag followed by a TEV (tobacco etch virus) protease cleavage site. The H229A mutation was introduced using the Quick Change Site Directed Mutagenesis kit (Stratagene) using the following complementary primers 5′-cgttaagcatattcgaatgtacgaag tattagccccttttcgagatgcc-3′ and 5′-ggcatctcgaaaaggggctaatacttcgtacat tcgaatatgcttaacg-3′ according to manufacturer's instructions. All constructs were checked by sequencing (GATC Biotech). Plasmids were transformed in *E. coli* Rosetta (DE3) pLysS strain (Invitrogen) and expression was induced at 25°C overnight using 0.5 mM IPTG in either a 2YT medium or a minimal medium containing 50 mg l^{-1} seleno-methionine for production of seleno-methionine labelled protein by blocking the methionine biosynthesis pathway (Doublie, 1997). After harvesting, cell lysis was done by a freezing/thawing cycle and sonication in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole) containing a cocktail of antiproteases, 0.25 mg ml^{-1} lysozyme, 5 mM $MgCl_2$ and benzonase. $AnkX_{1-484}$ was purified by a Ni^{2+}-affinity step using 125 mM imidazole for elution followed by gel filtration on a preparative Superdex 200 16/86 column in 10 mM Hepes 150 mM NaCl pH 7.5. Removal of the His_6 tag was performed after the affinity column by incubation with recombinant His_6-TEV protease for one night at 4°C (protease to protein ratio of 1:60 w:w) followed by a second Ni^{2+}-affinity step. The H229A mutant was purified by the same method except that 5 mM CDP-choline was added during lysis and after gel filtration. Purified proteins were concentrated using Vivaspin (Sartorius), flash frozen in liquid nitrogen and stored at −80°C until use.

E. coli XL1-Blue cells producing full-length His-AnkX for the phosphocholination assays were grown in LB broth containing 100 mg ml^{-1} ampicillin at 37°C to A600 = 0.6. IPTG was added to a final concentration of 1 mM and cells were continuously cultured for 3 h at 37°C. Cells were harvested and re-suspended in lysis buffer containing 20 mM Tris pH 7.4, 0.3 M NaCl, 1 mM DTT, 1 mM PMSF and 1% TX-100. Protein supernatant was obtained by sonication, followed by centrifugation at 15 000 *g* for 30 min. The supernatant was loaded onto a Ni-NTA column pre-washed with

lysis buffer. The column was washed extensively with lysis buffer and the His-AnkX protein was eluted with 200 mM imidazole in 20 mM Tris pH 7.4, 0.1 M NaCl, and 1 mM $MgCl_2$.

Crystallization and structure determination

The AnkX-CMP phosphocholine and AnkX-CMP complexes resulted of incubation of wild-type $AnkX_{1-484}$ (0.1 mM) with CDP-choline (CDC, 10 mM) in the presence of 10 mM $MgCl_2$. Initial nano-crystallization screening with the PEGs, PEGs II and JCSG+ suites (QIAGEN) and the Stura Footprint screen (Molecular Dimensions Limited) was performed in 96-well Greiner plates using a Cartesian robot with protein concentration varying between 5 and 9 mg ml^{-1}. Crystal optimization was performed in Limbro plates by mixing 1–2 µl of protein and 1 µl of crystallization well. Crystallization conditions are the following: $AnkX_{1-484}$ forms I and II, 0.2 M potassium thiocyanate 15–25% PEG 3350 0.05 M Hepes pH 7.8–8.2; AnkX/CMP form III, 15% PEG 20000 0.1 M Tris pH 8.5; AnkX/CMP/phosphocholine form IV, 0.2 M ammonium sulphate 30% PEG 5000 MME 0.1 M Mes pH 6.5; $AnkX^{H229A}$/CDC, 0.2 M Lithium sulphate 14% PEG 4000 0.1 M Tris pH 7.5. Crystals were cryoprotected using 20% glycerol. Data were collected on the Proxima-1 beamline at the SOLEIL Synchrotron (Gif-sur-Yvette, France). Data processing and scaling were done with XDS (Kabsch, 2010). Phasing was performed using seleno-methionine-labelled $AnkX_{1-484}$ (crystal form II) by the Single Anomalous Diffraction (Se-SAD) method using the *phenix.autosol* wizard (Adams *et al*, 2010). Resulting phases were the starting point for automatic model building with *phenix.autobuild* (Adams *et al*, 2010). The structure of apo $AnkX_{1-484}$ (crystal form I) was solved by a combination of molecular replacement (MR) using the partial model obtained for crystal form II and Se-SAD (MR-SAD) using the Phenix software (Adams *et al*, 2010). Refinement was performed using 'jelly body' as implemented in Refmac5 (Murshudov *et al*, 2011). Liganded AnkX (forms III and IV) and the $AnkX^{H229A}$ complex were solved by MR with Phaser using crystal form I as a model (McCoy *et al*, 2007) and refined with BUSTER (Blanc *et al*, 2004). Model building was done with Coot (Emsley *et al*, 2010) and validated using Molprobity (Davis *et al*, 2007). TLS groups definition was assisted by the TLSMD server (Painter and Merritt, 2006). Final refinement statistics and quality of the models are summarized in Table I. Figures were generated with Pymol.

Atomic coordinates and structure factors have been deposited with the Protein Data Bank with accession codes: AnkX apo (P212121 form I): 4bep, AnkX/CMP (form III): 4ber, AnkX/CMP/phosphocholine (form IV): 4bes, AnkX-H229A/CDP-choline: 4bet.

Phosphocholination assays

GST-tagged Rab1A attached to glutathione agarose beads was incubated with thrombin to remove the GST tag. Untagged Rab1A (1 µg) was incubated with 1 µg of AnkX proteins as indicated in phosphocholination buffer (20 mM HEPES pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, and 5 mM CDP-choline) for 1 h at 30°C. Samples were boiled in SDS-loading buffer and run on a 15% SDS–PAGE gel, followed by transfer onto a PVDF membrane. Phosphocholination of AnkX and Rab1A was detected by immunoblot analysis using an antibody specific for phosphocholine (TEPC-15, Sigma). Protein loading was assessed by Ponceau staining. To probe the contribution of Mg^{2+} to catalysis, full-length His-tagged AnkX was purified as before and subsequently passed through a PD MidiTrap G-25 column (GE Healthcare) equilibrated in a phosphocholination buffer containing no $MgCl_2$. Blots were quantitated using the Image Quant TL software. All data shown were validated in at least three independent experiments.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

This work was funded by an ANR grant to JC, NIH grant AI041699 to CRR and an Anna Fuller Fellowship award to SM. We thank the staff at PROXIMA 1 beamline at the SOLEIL synchrotron (Gif-sur-Yvette, France) and at the LEBS/IMAGIF cloning and crystallization platforms (CNRS, Gif-sur-Yvette, France) for their help.

Author contributions: JC, VC, and CRR conceived and planned the project. VC and SM performed the experiments. JC and CRR wrote the manuscript with input from the other authors.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221

Blanc E, Roversi P, Vonrhein C, Flensburg C, Lea SM, Bricogne G (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D Biol Crystallogr* **60:** 2210–2221

Cossart P, Roy CR (2010) Manipulation of host membrane machinery by bacterial pathogens. *Curr Opin Cell Biol* **22:** 547–554

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall 3rd WB, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35:** W375–W383

Desveaux D, Singer AU, Wu AJ, McNulty BC, Musselwhite L, Nimchuk Z, Sondek J, Dangl JL (2007) Type III effector activation via nucleotide binding, phosphorylation, and host target interaction. *PLoS Pathog* **3:** e48

Doublie S (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* **276:** 523–530

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501

Engel P, Goepfert A, Stanger FV, Harms A, Schmidt A, Schirmer T, Dehio C (2012) Adenylylation control by intra- or inter-molecular active-site obstruction in Fic proteins. *Nature* **482:** 107–110

Feng F, Yang F, Rong W, Wu X, Zhang J, Chen S, He C, Zhou JM (2012) A Xanthomonas uridine 5′-monophosphate transferase inhibits plant immune kinases. *Nature* **485:** 114–118

Goody PR, Heller K, Oesterlin LK, Muller MP, Itzen A, Goody RS (2012) Reversible phosphocholination of Rab proteins by Legionella pneumophila effector proteins. *EMBO J* **31:** 1774–1784

Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132

Kinch LN, Yarbrough ML, Orth K, Grishin NV (2009) Fido, a novel AMPylation domain common to fic, doc, and AvrB. *PLoS ONE* **4:** e5818

Kwak BY, Zhang YM, Yun M, Heath RJ, Rock CO, Jackowski S, Park HW (2002) Structure and mechanism of CTP:phosphocholine cytidylyltransferase (LicC) from Streptococcus pneumoniae. *J Biol Chem* **277:** 4343–4350

Lassila JK, Zalatan JG, Herschlag D (2011) Biological phosphoryl-transfer reactions: understanding mechanism and catalysis. *Annu Rev Biochem* **80:** 669–702

Lee J, Johnson J, Ding Z, Paetzel M, Cornell RB (2009) Crystal structure of a mammalian CTP: phosphocholine cytidylyltransferase catalytic domain reveals novel active site residues within a highly conserved nucleotidyltransferase fold. *J Biol Chem* **284:** 33535–33548

Li J, Mahajan A, Tsai MD (2006) Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry* **45:** 15168–15178

Luong P, Kinch LN, Brautigam CA, Grishin NV, Tomchick DR, Orth K (2010) Kinetic and structural insights into the mechanism of AMPylation by VopS Fic domain. *J Biol Chem* **285:** 20155–20163

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674

Michaely P, Tomchick DR, Machius M, Anderson RG (2002) Crystal structure of a 12 ANK repeat stack from human ankyrinR. *EMBO J* **21:** 6387–6396

Mukherjee S, Liu X, Arasaki K, McDonough J, Galan JE, Roy CR (2011) Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477:** 103–106

Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67:** 355–367

Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62:** 439–450

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

Pan X, Luhrmann A, Satoh A, Laskowski-Arce MA, Roy CR (2008) Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* **320:** 1651–1654

Roy CR, Mukherjee S (2009) Bacterial FIC proteins AMP Up infection. *Sci Signal* **2:** pe14

Vincentelli R, Bignon C, Gruez A, Canaan S, Sulzenbacher G, Tegoni M, Campanacci V, Cambillau C (2003) Medium-scale structural genomics: strategies for protein expression and crystallization. *Acc Chem Res* **36:** 165–172

Voth DE (2011) ThANKs for the repeat: Intracellular pathogens exploit a common eukaryotic domain. *Cell Logist* **1:** 128–132

Worby CA, Mattoo S, Kruger RP, Corbeil LB, Koller A, Mendez JC, Zekarias B, Lazar C, Dixon JE (2009) The fic domain: regulation of cell signaling by adenylylation. *Mol Cell* **34:** 93–103

Xiao J, Worby CA, Mattoo S, Sankaran B, Dixon JE (2010) Structural basis of Fic-mediated adenylylation. *Nat Struct Mol Biol* **17:** 1004–1010

Yarbrough ML, Li Y, Kinch LN, Grishin NV, Ball HL, Orth K (2009) AMPylation of Rho GTPases by Vibrio VopS disrupts effector binding and downstream signaling. *Science* **323:** 269–272