Structure of a bacterial toxin-activating acyltransferase

Nicholas P. Greene[1], Allister Crow[1], Colin Hughes[2], and Vassilis Koronakis

Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, United Kingdom

Edited by Scott J. Hultgren, Washington University School of Medicine, St. Louis, MO, and approved May 4, 2015 (received for review March 6, 2015)

Secreted pore-forming toxins of pathogenic Gram-negative bacteria such as *Escherichia coli* hemolysin (HlyA) insert into host–cell membranes to subvert signal transduction and induce apoptosis and cell lysis. Unusually, these toxins are synthesized in an inactive form that requires posttranslational activation in the bacterial cytosol. We have previously shown that the activation mechanism is an acylation event directed by a specialized acyl-transferase that uses acyl carrier protein (ACP) to covalently link fatty acids, via an amide bond, to specific internal lysine residues of the protoxin. We now reveal the 2.15-Å resolution X-ray structure of the 172-aa ApxC, a toxin-activating acyl-transferase (TAAT) from pathogenic *Actinobacillus pleuropneumoniae*. This determination shows that bacterial TAATs are a structurally homologous family that, despite indiscernible sequence similarity, form a distinct branch of the Gcn5-like *N*-acetyl transferase (GNAT) superfamily of enzymes that typically use acyl-CoA to modify diverse bacterial, archaeal, and eukaryotic substrates. A combination of structural analysis, small angle X-ray scattering, mutagenesis, and cross-linking defined the solution state of TAATs, with intermonomer interactions mediated by an N-terminal α-helix. Superposition of ApxC with substrate-bound GNATs, and assay of toxin activation and binding of acyl-ACP and protoxin peptide substrates by mutated ApxC variants, indicates the enzyme active site to be a deep surface groove.

hemolysin | acyltransferase | posttranslational modification | X-ray crystallography | acyl carrier protein

Pathogenic bacteria secrete pore-forming protein toxins (PFTs) that target tissue and immune cell membranes to aid colonization and survival during infections, subvert cell signaling, induce apoptosis, and promote cell lysis (1–8). Among Gram-negative bacteria, large PFTs are secreted by pathogenic species of *Pasteurella*, *Actinobacillus*, *Proteus, Morganella, Moraxella*, and *Bordetella*, exemplified by the 110-kDa hemolysin (HlyA) of uropathogenic and enterohemorrhagic *Escherichia coli*. These toxins play important roles in cystitis and pyelonephritis, hemorrhagic intestinal disease, periodontitis, pneumonia, septicemia, whooping cough, and wound infections (4), and unusually they are made as an inactive protoxin, requiring posttranslational activation before export (9–12).

Reconstituting the toxin activation reaction in vitro some time ago demonstrated that the essential modification is a novel fatty acid acylation, affected by a specialized coexpressed toxin-activating acyltransferase, in *E. coli* HlyC, that uses acyl-acyl carrier protein (acyl-ACP) as the fatty acid donor (4, 13, 14). The acyltransferase does not share significant sequence identity with other bacterial and eukaryotic enzymes, and cellular acyltransferases from either the host or pathogen cannot substitute for HlyC in toxin activation. HlyC independently binds two separate 50- to 80-aa transferase recognition domains (15), each encompassing one of the internal target lysines K564 and K690 of *E. coli* protoxin HlyA, which are acylated by amide linkage, heterogeneously with fatty acids containing 14, 15, and 17 carbon chains (16, 17). Loss of the HlyC binding domain or substitution of protoxin K564 and K690 prevents fatty acyl modification and abrogates all toxin activity (14) as does loss of the transferase (18).

Acylation is essential to the entire family of pore-forming toxins, *Bordetella pertussis* proCyaA lysine acylation has also been demonstrated (19), and the toxin-activating acyltransferases (which we now call TAATs) have high sequence similarity and cross-activate other protoxins (4, 20–22). The TAAT activation mechanism is seemingly unique, and extensive site-directed mutagenesis has so far only identified a single potentially catalytic residue, His23 of HlyC (23–25). Structural information is essential to understand the toxin activation mechanism and assess TAATs as a potential target for developing antivirulence compounds that do not affect the host commensal flora. Here, we determine the TAAT crystal structure, solution state, and likely active site.

Results

The ApxC Structure Reveals TAATs as a Distinct, Closely Conserved Branch of the GNAT Superfamily. Multiple TAAT genes were cloned and expressed in *E. coli* and where possible TAAT proteins were purified and screened for their amenability to crystallization. This process yielded the crystal structure of the 172-aa *Actinobacillus pleuropneumoniae* ApxIC (hereafter ApxC) at 2.15-Å resolution by using the selenomethionine single-wavelength anomalous dispersion (SAD) method of phase determination. Crystallographic data and refinement statistics are given in Table S1 and the structure is shown in Fig. 1. Crystals belong to space group $P2_12_12_1$ with 4 molecules per asymmetric unit (Fig. 1*A*). The four monomers are similar and can be superposed pairwise with a mean rmsd of 1.7 Å (Cα positions, residues 4–170). Structural diversity is largely limited to the last 7 residues of the C terminus (mean rmsd for residues 5–165 is 0.84 Å) and is not expected to be of physiological relevance because this region is not well-conserved among TAATs. Each ApxC monomer is composed of a central five-membered β-sheet that is decorated by six α-helices (Fig. 1*B*). A distinctive split in the sheet is located between strands β3 and β4 which are the only strands that run parallel with one another. This split β-sheet is a key

Significance

Secreted pore-forming toxins of pathogenic bacteria such as *Escherichia coli* and *Bordetella pertussis* insert into cell membranes to subvert signaling and cause cell death, facilitating infection of human and animal hosts. These toxins require a unique activation step before secretion, the covalent linkage of lipid groups to specific lysines of the inactive protoxin, directed by a specialized toxin-activating acyl transferase (TAAT). Here, we present the TAAT crystal structure, the soluble dimeric topology, and likely active site, revealing that despite no discernible sequence similarity, TAATs are a structurally and functionally distinct group of the Gcn5-like *N*-acetyl transferase (GNAT) superfamily of modifying enzymes. Our findings open the way to further understanding of the unique toxin activation, and the possibility of inhibiting toxin action.

Author contributions: N.P.G., A.C., C.H., and V.K. designed research; N.P.G. and A.C. performed research; N.P.G., A.C., C.H., and V.K. analyzed data; and N.P.G., A.C., C.H., and V.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4WHN).

[1]N.P.G. and A.C. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: ch@mole.bio.cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1503832112/-/DCSupplemental.

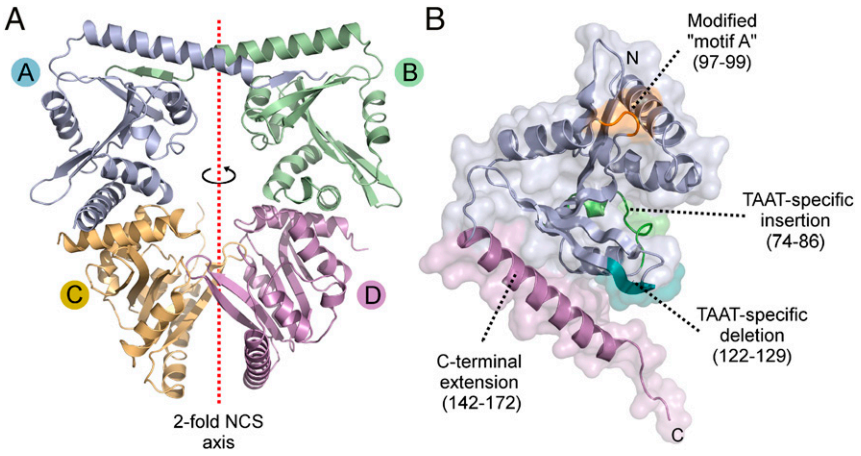

Fig. 1. Crystal structure of the toxin-activating acyl transferase ApxC. (*A*) The asymmetric unit contains four monomers of ApxC. An imperfect twofold noncrystallographic symmetry axis (dotted red line) relates two pairs of monomers. (*B*) Structure of a single monomer indicating structural features that discern TAATs from other members of the GNAT superfamily.

feature of the underlying fold, establishing a deep groove in the enzyme surface.

The overall fold of ApxC is clearly related to the Gcn5-like *N*-acetyl transferase (GNAT) superfamily, confirmed by using structural-similarity searches with Dali (26). GNAT superfamily members mediate a diverse range of prokaryotic and eukaryotic processes ranging over antibiotic detoxification, circadian rhythm regulation, and polyketide biosynthesis (27, 28). The closest structural match is the GNAT-like domain of CurA (rmsd 2.8 Å, Dali Z score = 8.6) (29), but robust structural similarity is also evident with other GNATs including the histone acetyl transferases (HATs), Gcn5 (rmsd 3.2 Å, Z = 7.8) (30), p300/CBP (rmsd 3.1 Å, Z = 7.9) (31), Myst (rmsd 4.0 Å, Z = 3.1) (32), and acyl-amino acid synthase FeeM (rmsd 3.6 Å, Z = 6.6) (33). The ApxC structure is nonetheless distinguished from all other GNATs by the absence of a helical element that typically interacts with the planar nucleotide portion of CoA [this would be located between ApxC residues 122 and 129 (Fig. 1*B*, teal)] and the CoA diphosphate binding motif ["motif A" in GNATs; (27)], that would otherwise be located between strand β3 and helix α3 (residues 97–99; Fig. 1*B*, orange). ApxC appears to be the only GNAT that lacks both of these two features; indeed, the helix that is missing in ApxC is so uniformly conserved among GNATs that, alongside the split β-sheet, it is considered a defining feature of the GNAT fold (27, 28). In addition to the absence of CoA-interacting elements, the long (25-residue) α-helix at the extreme C terminus of ApxC (Fig. 1*B*, violet) is absent from all known GNAT structures, and a short helix and loop (residues 74–86, Fig. 1*B*, green) is present only in CurA (29). The close structural relationship between ApxC and members of the GNAT superfamily is remarkable given that there is no primary sequence similarity between these proteins. Even structure-driven sequence alignments produced with Dali reveal just 6.4% sequence identity between ApxC and its closest structural match, the CurA GNAT domain, barely above that expected by chance (5%).

In contrast to the low sequence identity between TAATs and other GNATs, the TAAT family itself is well conserved, illustrated by the multiple sequence alignment in Fig. 2. ApxC shares 70% primary sequence identity with the best studied TAAT, *E. coli* HlyC, and can directly substitute for HlyC in activation of the *E. coli* hemolysin toxin in vivo (Fig. S1), suggesting the ApxC structure is representative of the TAAT family and allowing us to construct high confidence homology models of homologs *E. coli* HlyC and *Bordetella pertussis* CyaC (Datasets S1 and S2).

TAATs Adopt a Dimeric Arrangement in Solution, Mediated by an N-Terminal Helical Interface. Gel filtration experiments using recombinant ApxC indicate an apparent molecular mass (42 kDa) that is twice that expected for the monomer (21 kDa) (Fig. S2). Examination of the crystal packing in the ApxC structure shows four arrangements that could conceivably represent this dimer (dimers I to IV in Fig. 3*A*). Dimer I, underpinned by the unique C-terminal helix, and dimer III, which appears twice in the asymmetric unit (Fig. S3), were attractive possibilities, but it was not obvious by inspection which interface reflects the physiological dimer and which are crystal contacts.

We assessed the extent and quality of each of the four dimer interfaces by measuring the buried surface area and calculating the predicted stability by using PISA (34) (Fig. 3*A*). Dimer IV had the lowest buried surface area and was predicted to be unstable but dimers I, III, and IV had substantial buried surface area and were all predicted to be stable. Analysis of scattering data from small angle X-ray scattering (SAXS) (Fig. 3*B*) indicated that the particle has an estimated molecular mass of ∼38.8 kDa, a 28 Å radius of gyration and a probable maximum dimension of ∼77 Å. Ab initio shape reconstructions of the solution particle were consistent with a twofold symmetric bilobal globular dimer (Fig. 3*B*). Comparison of the scattering data with theoretical profiles predicted for each dimer was used to assess which arrangement most closely resembles the solution particle (Fig. 3*C*). Both dimer I and dimer III have low χ^2 values, indicating a good fit between prediction and experiment, whereas dimers II and IV do not. Inspection of the predicted P(r) profiles (Fig. 3*C*, black lines) shows that both dimers I and III are close to the experimentally derived profile (blue lines), consistent with bilobal structures with a maximum dimension of ∼80 Å, whereas dimers II and especially IV deviate significantly, suggesting that the maximum dimensions of these dimers are, respectively, 10 Å and 15 Å too short to be consistent with the data. Taken together, the surface areas, predicted stabilities, and SAXS data excluded dimers II and IV but did not satisfactorily discern between dimers I and III.

We therefore applied two further molecular genetic experimental approaches. First, we conducted in vitro cross-linking experiments with the aim of covalently trapping the physiological dimer by using experimentally engineered cysteine residues placed independently at each potential dimerization interface. Second, we used a bacterial two-hybrid system (35) to monitor the intermonomer interaction in vivo after introducing disruptive mutations to each putative interface. The interface structures, genetic interventions, and results are presented in Fig. 4. Both

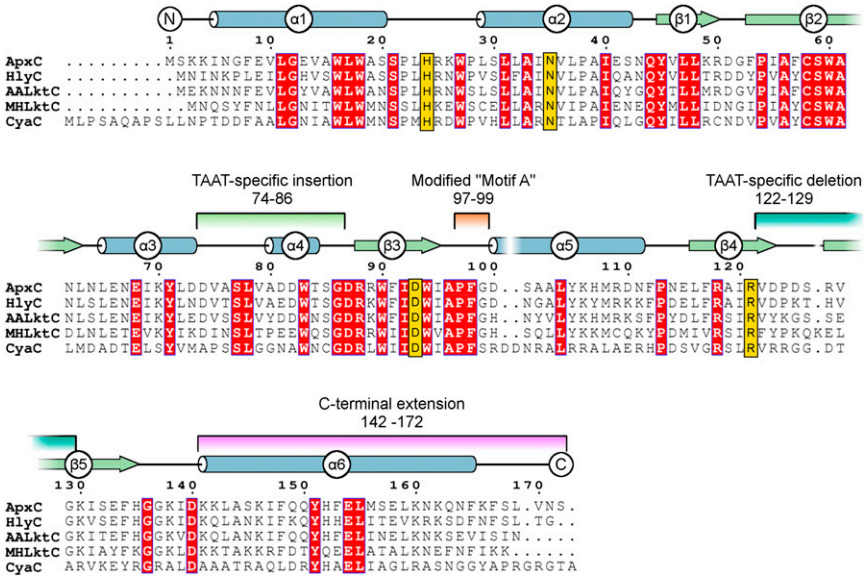

Fig. 2. Multiple sequence alignment for key members of the TAAT family. Secondary structure of ApxC (consensus of the four monomers in the crystal structure) is shown immediately above the alignment of TAAT sequences ApxC (*A. pleuropneumoniae*), HlyC *(Uropathogenic E. coli*), AaLktC (*A. actinomycetemcomitans*), MhLktC (*M. hemolytica*), and CyaC (*Bordetella pertussis*). The α-helices indicated are shown as tubes (blue); β-strands by arrows (green). Conserved residues are highlighted red, and important active site residues identified later (His24, Asn35, Asp93, and Arg121) are shown in gold. Features that differentiate TAATs from other GNATs are highlighted by using the same color scheme as in Fig. 1*B*.

dimers I and III are twofold rotationally symmetric, so it was possible to introduce single cysteines close to the relevant rotation axis to generate potential cross-link sites, namely Gly12Cys within the N-terminal helix to specifically trap dimer III, and Lys160Cys within the C-terminal helix to trap dimer I (Fig. 4*A*). Cross-linking of the dimer was successful with Gly12Cys, but not Lys160Cys, and no cross-linked dimer was generated by the wild-type protein, nor when solvent accessible cysteines were pre-blocked with *N*-ethylmaleimide (NEM) (Fig. 4*B*). In parallel, we used a bacterial two-hybrid assay (35) in which wild-type ApxC dimerization results in reporter activity (β-galactosidase activity, blue spot, top box; Fig. 4*D*). Mutations were engineered to disrupt either the N-terminal helix within the dimer III interface (Gly12Ala, Gly12Glu, Ala15Glu, and Ala19Glu substitutions; Fig. 4*C*) or the intertwined C termini of dimer I by a 10-aa deletion (Fig. 4*E*), all variant fusion proteins expressed comparably (Fig. S4). The Gly12Glu and Ala15Glu mutation disrupted the dimer (white spots; Fig. 4*D*), whereas deletion of the C terminus had no effect (Fig. 4*F*). Taken together these experiments unambiguously define dimer III as the state of the enzyme in solution.

Identification of the TAAT Active Site Within a Deep Surface Groove. The structural similarity between TAATs and the GNATs provoked a search for shared catalytic residues in structural superimpositions of ApxC with key members of the Gcn5 (30), Hat1 (31), Myst (36), p300/CBP (31), FeeM (33), and CurA (29) GNAT subfamilies. Statistics for the structural superpositions are given in Table S2 alongside summaries of the expected active site residues and the ApxC residues to which they most closely align. Our search confirmed that TAATs do not contain catalytic residues established for other GNATs including the buried glutamates found in the Gcn5 and Hat1 families (27, 28), the cysteine and glutamate found in Myst (36), or the phenylalanine and tyrosine residues thought responsible for substrate alignment in the CBP/p300 subfamily (31). To assess potential substrate-binding sites, ApxC was also compared with substrate-bound structures of the GNATs Gcn5 (30) and FeeM (33) (Fig. 5). Gcn5, the archetypal member of GNAT superfamily is a histone acetyltransferase that catalyzes the acetyl CoA-dependent acylation of a sequence-internal histone lysine residue. The crystal structure of the ternary complex (30) formed among Gcn5, CoA, and a peptide representing the histone tail superposes well with ApxC with parts of both substrates overlaying portions of the TAAT active site groove (Fig. 5*A*). FeeM catalyses the acylation of tyrosine by using FeeL, an ACP-like protein, as acyl donor. Structural superposition of FeeM with ApxC places the acyl chain of the *N*-lauroyl tyrosine product (cocrystallized with FeeM) within a hydrophobic portion of the TAAT surface groove (Fig. 5*B*). These superpositions suggest that the deep groove observed for ApxC is the likely substrate binding site.

To confirm the location of the active site and to identify key residues, we generated alanine mutants of residues lining the groove, in particular targeting polar residues conserved in the TAAT family (Fig. 2). These residues included His24 for which the equivalent residue in *E. coli* TAAT HlyC has been shown to be important (24), and Ser21 and Asp93, which together with His24 have a 3D arrangement reminiscent of catalytic triads found in non-GNAT acyltransferases and proteases). The structural context of these targeted residues is shown in Fig. 6*A*, and relative TAAT activities of ApxC alanine variants assessed by erythrocyte lysis of variant-activated toxin are indicated in Fig. 6*B*. The Ser21Ala substitution caused only a marginal (21%) reduction in TAAT activity, inconsistent with a role as a nucleophile within a catalytic triad. A role for the TAAT family conserved cysteine residue can similarly be excluded as the ApxC structure reveals Cys58 is buried within the core of the protein where it is inaccessible to substrates (Fig. S5 *A* and *B*). In agreement with this observation, the Cys58Ala variant has wild-type levels of activity (Fig. S5*C*). In contrast, both Asp93 and Arg121 alanine variants were profoundly affected by mutation with TAAT activities less than 1% of the wild type. The His24 and Asn35 variants were also significantly impaired, although not to the same degree, with measurable activities of c.5% of wild type (Fig. 6*B*).

To probe the roles of these four important residues, we performed bacterial two-hybrid experiments to assess their ability to engage with TAAT substrates (i.e., the protoxin and ACP). All four alanine-substituted variants were able to dimerize efficiently, indicating that each is properly folded (Fig. 6*C*, *Left*), and all

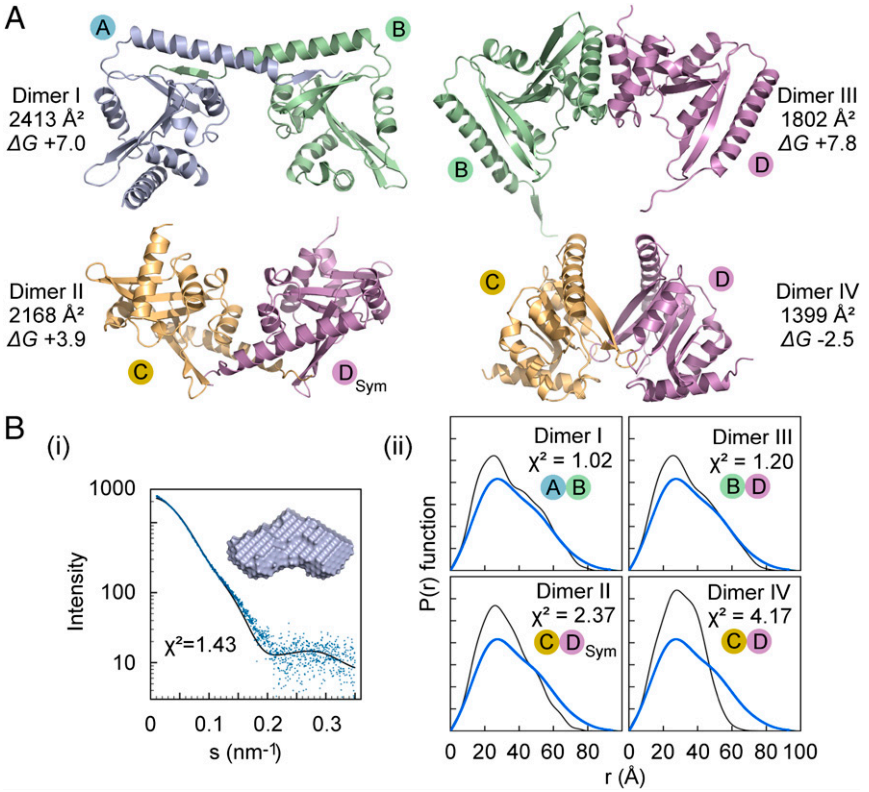

Fig. 3. X-ray analysis of potential ApxC dimers *in crystallo* and in solution. (*A*) The four dimeric arrangements of ApxC extracted from the crystal structure (I–IV) indicating buried surface area and predicted free energy of dissociation (*ΔG*, positive values indicate favorable association) obtained from PISA (34). (*B*) SAXS analysis of ApxC in solution. (*i*) The measured SAXS scattering intensity profile is shown in blue (56 μM ApxC) and the predicted scattering profile of the solution model in *black*, whereas the solution structure of ApxC, a filtered model derived from the average of 20 models, is shown in *Inset*. (*ii*) Distance distribution analysis of each of the candidate ApxC dimers verses the SAXS data. The quality of the fit can be seen by comparing predicted P(r) profile in black with the measured profile in blue, quantified by the χ^2 value shown in *Inset*.

engaged with a 58-aa peptide (residues 679–736 of the protoxin) known to be the recognition sequence for TAAT spanning K690 (15) (Fig. 6*C*, *Center*). Asn35, Asp93, and His24 alanine variants were also able to bind ACP, but the Arg121Ala mutant was not (Fig. 6*C, Right*) despite being expressed at comparable level to the wild type (Fig. 6*B*). We therefore anticipate Arg121 mediates interaction with ACP, whereas the remaining essential residues, His24, Asn35, and Asp93, are likely to have roles in catalysis. The putative active sites encompassing these residues in the surface groove are on opposing faces within the dimer, not in close juxtaposition (Fig. 6*D*).

Discussion

We have described the high-resolution crystal structure (Fig. 1), solution structure (Fig. 3*B*), dimerization interface (Figs. 3 and 4), and active site location (Fig. 6) of ApxC, revealing the close-knit family of TAATs as a previously unidentified subdivision of the GNAT superfamily. Whereas TAATs retain the characteristic GNAT split in the central β-sheet that underlies the putative substrate-binding groove, they lack the structural motifs typically associated with CoA binding, including both the nucleotide-interacting helix and diphosphate-binding motif that are considered core elements of the characteristic GNAT fold. A long C-terminal α-helix and small loop-helix feature further distinguish TAATs from other GNATs (Fig. 1).

In vitro SAXS (Fig. 3) and in vivo two-hybrid analyses (Fig. 4) confirmed indications from gel filtration [previously with HlyC (18) and here with ApxC] that TAATs are dimers in solution. Analysis of biophysically tenable dimers abstracted from the crystal structure using a combination of experimental solution scattering data, targeted mutagenic disruption, and cross-link locking of potential intermonomer contacts firmly established the dimer topology (Fig. 3*A*, dimer III). The dimer interface is largely hydrophobic, with the close approach of the N-terminal helices facilitated by the small side chains of Gly12 and Ala15. Consequently, the proposed active sites of the monomers, each located within a surface groove, are topologically distinct within the context of the dimer and unlikely to act in concert (Fig. 6*D*).

The lack of typical GNAT CoA-binding features in TAATs is understandable from a functional perspective because TAATs use ACP as the substrate acyl donor and not CoA. However, FeeM also depends on an ACP-like acyl donor (FeeL) yet retains the GNAT-typical CoA-interacting helix. The interaction between FeeM and FeeL has been partially mapped by mutation and involves the helical region that is absent in TAATs (33), suggesting the TAAT:ACP interaction must be distinct from that of FeeM:FeeL. The only other structurally characterized GNAT known to interact with ACP is CurA, a large multidomain protein involved in polyketide synthesis. The CurA GNAT domain interacts with both CoA and an ACP-like internal domain facilitating acyl transfer between these substrates (29). CoA binding in CurA occurs in typical GNAT fashion, whereas a distinct ACP binding site has been proposed to lie in a separate tunnel that connects directly to the CoA binding site. A similar binding mode is possible in TAATs because the proposed ACP-binding tunnel of CurA partially overlaps the ApxC surface groove.

Structure-led mutagenesis of the TAAT surface groove, the putative active site, revealed two residues, Asp93 and Arg121, as essential to function and Asn35 and His24 as important. Mutation to alanine did not impair dimerization, and although all four

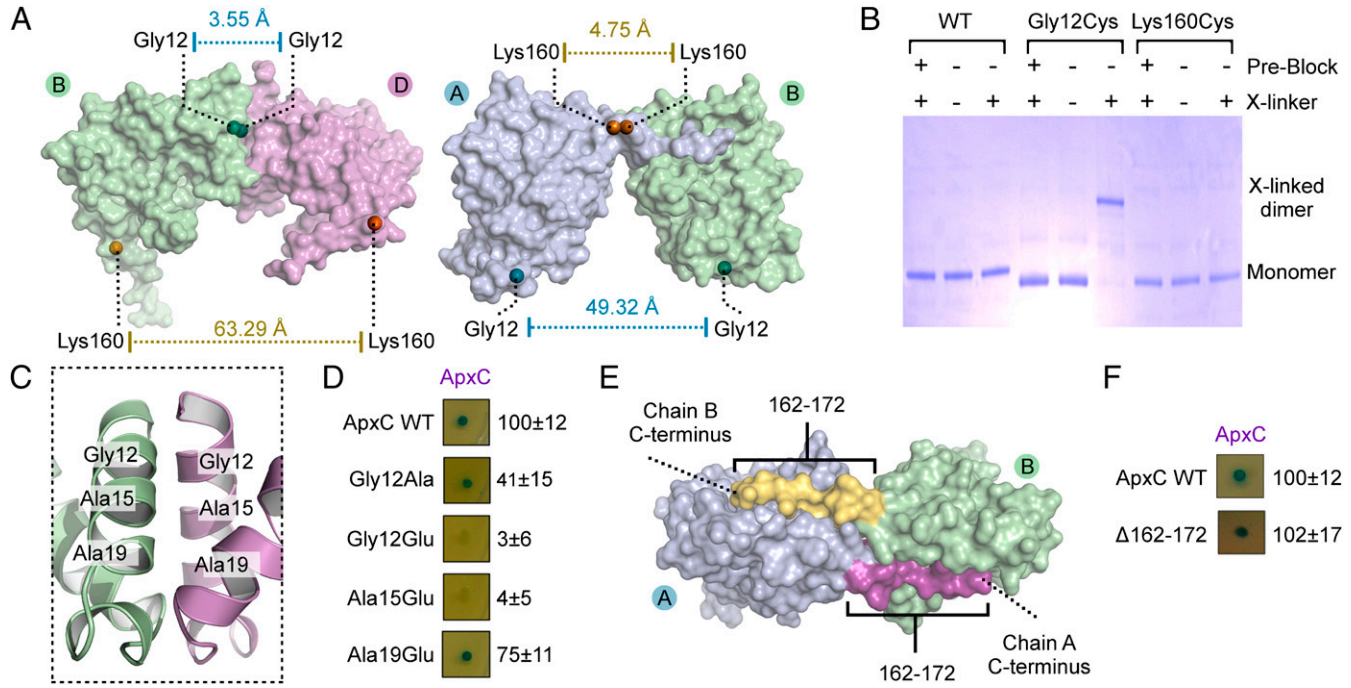

Fig. 4. Experimental determination of the ApxC solution dimer. (*A*) Locations of engineered cysteine residues (marked with spheres Gly12, blue; Lys160, orange) with respect to two dimeric arrangements (dimer III, *Left*; dimer I, *Right*). (*B*) Cysteine cross-linking of purified wild-type (WT) ApxC or cysteine variant. Protein (50 μg) was treated with or without 100 μM $CuCl_2$ (X-linker) for 30 min at 25 °C to catalyze cross-linking of proximal cysteines before blocking free cysteines with 5 mM *N*-ethylmaleimide (NEM). Preblock reactions were treated with NEM and EDTA before addition of cross-linker. (*C*) Locations of Gly12, Ala15, and Ala19 within the N-terminal α-helix at the dimer III interface. (*D*) Bacterial two-hybrid assay of homodimerization of WT ApxC or indicated variants. Interaction was assessed qualitatively by blue coloration and quantitatively by β-galactosidase activity (*Right*), expressed as a percentage of the wild type ± SD. (*E*) Location of an engineered C-terminal deletion (Δ162–172) in context of dimer I. (*F*) Bacterial two-hybrid assay assessing homodimerization of ApxC (Δ162–172), details as for *D*.

alanine variants still interacted with the target protoxin, mutation of Arg121 uniquely abrogated binding of ACP. In other unrelated ACP-binding enzymes, positively charged residues mediate direct interactions with phosphates of the ACP phosphopantetheine group (37) or the negatively charged residues of ACP (38), and so it seems likely that Arg121 plays a similar role in the TAAT:ACP interaction. For Asn35, the chemistry of the asparagine side chain precludes a role in acid-base catalysis and amide group is unlikely to be a nucleophilic, but it may have a role in orientating substrates or other essential residues e.g., Asp93 to which it is hydrogen bonded.

His24 and Asp93 are both candidates to be true catalytic residues and plausible roles within the context of the two most likely acyl transfer schemes are outlined in Fig. S6. In the first scheme, a ternary complex is formed between the enzyme and both substrates, facilitating direct nucleophilic attack of the protoxin target lysine on the acyl-ACP thioester with subsequent expulsion of ACP (a "direct attack" mechanism; Fig. S6*A*). In the second, the reaction proceeds in two steps through formation of a covalent acyl-enzyme intermediate that is resolved by attack of the toxin receptor lysine ("covalent catalysis"; Fig. S6*B*). In either mechanistic scheme, analogy to other GNATs, particularly histone acetyltransferases such as Gcn5 and Hat1, suggests Asp93 acts as a general base to deprotonate the receptor lysine of the protoxin, and in the specific case of covalent catalysis, it could additionally deprotonate the catalytic nucleophile. His24 would then serve as a general acid protonating the ACP thiolate leaving group in the direct attack scheme, or else as the enzymatic nucleophile in the covalent catalysis scheme.

That TAATs represent a subfamily of GNATs is an important finding as kinetically characterized GNATs appear to use a ternary complex to facilitate transfer of acyl groups from one substrate to the other, giving strong support for a ternary complex (direct attack)-based mechanism (27, 30, 39, 40). Formation of a ternary complex is consistent with early kinetic experiments on HlyC (41), but at variance with other results favoring sequential binary interactions between enzyme and substrate (42), and observations of acylenzyme formation inferred from experiments with radiolabeled fatty acids (23, 24). Incubation of TAAT HlyC with acyl-ACP yields acylated enzyme and depends on integrity of the active site histidine (equivalent to ApxC His24) (23). However, direct histidine modification has not been demonstrated, and in other GNATs, internal enzyme acylation is suggested to be a self-regulatory feature (36, 43), so it is not clear whether the acylated TAAT is an acyl transfer

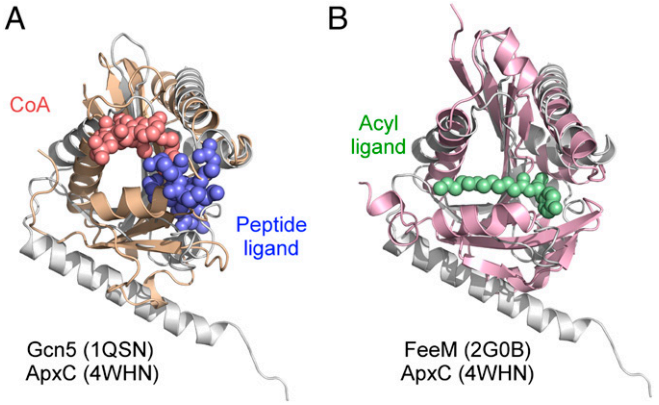

Fig. 5. Superposition of the TAAT ApxC with Gcn5 and FeeM. (*A*) Superposition of ApxC (white) with Gcn5 (wheat). The two Gcn5 substrates are CoA (red) and histone H3 peptide (blue). (*B*) Superposition of ApxC (white) with FeeM (pink). The FeeM *N*-lauroylayrosine ligand is shown in green.

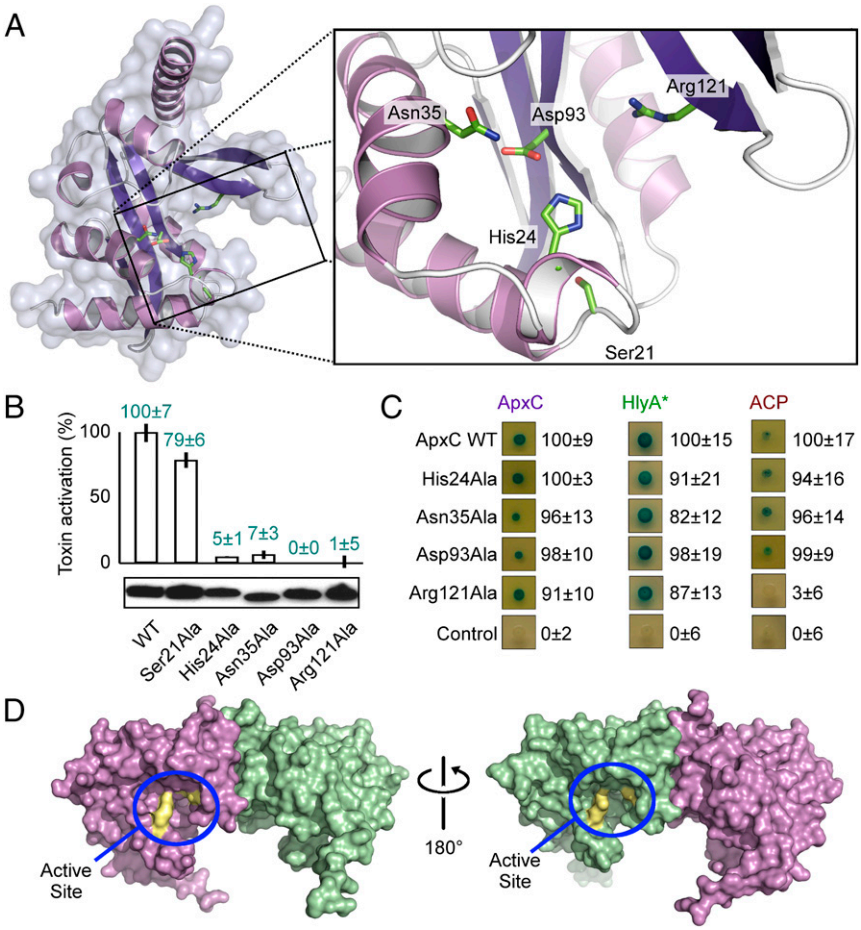

Fig. 6. Definition of the TAAT active site. (*A*) Location of the postulated substrate-binding groove and close-up of the active site (boxed). Side chains of key active site residues are shown in stick representation. (*B*) TAAT activity of active site variants assayed by in vivo toxin activation (erythrocyte lysis, percentage of wild-type activity ± SD with immunoblot showing expression underneath. (*C*) Bacterial two-hybrid experiments testing ApxC dimerization (*Left*), interaction with TAAT-binding protoxin peptide (residues 679–736 - HlyA*; *Center*), and ACP binding (*Right*) for wild-type (WT) ApxC and mutant variants. Interaction was assessed qualitatively by blue coloration and numerically by β-galactosidase activity (*Right*), expressed as a percentage of the wild type (*Top*) ± SD after subtraction of the negative control. (*D*) Solid surface representations of ApxC showing the location of the active site of each monomer in context of the soluble dimer (blue lasso), with important active site residues in yellow.

intermediate, regulatory state, or artifact of acyl donor interaction in the absence of the acyl acceptor.

In support of ternary complex formation, the TAAT surface groove is of comparable size to that of Gcn5, and so probably big enough to accommodate both the acylated phosphopantetheine arm of acyl-ACP and the protoxin lysine, if not also a substantial part of unfolded protoxin. Based on the superposition of TAATs with other GNATs (Fig. 5) and our identification of the TAAT catalytic residues (Fig. 6), we illustrate how TAATs could bring together the acyl-ACP and toxin substrates to effect toxin activation in Fig. 7. The central location of the active site residues within the groove (Fig. 7*B*, yellow) hints that each substrate could be bound on opposing faces of the enzyme with acyl transfer chemistry occurring at the crux of the split β-sheet. In support of this proposed arrangement, the ACP-interacting arginine of TAATs (Arg121 in ApxC) is located on the enzyme face equivalent to the part of the active site groove in FeeM that accommodates a lauryl acyl chain (Fig. 7*A*, green), suggesting the likely TAAT ACP binding site (Fig. 7*B*, green). The opposite TAAT face corresponds to where the histone tail is located in the Gcn5 ternary complex (Fig. 7*A*, blue) suggesting the TAAT protoxin-binding site (Fig. 7*B*, blue). If TAATs use a ternary complex, it seems likely they would also use the direct attack mechanism typical of other GNATs. Otherwise, TAATs will be of significant interest as a truly unique subset of the GNAT superfamily.

Elucidation of the TAAT protein structure, solution oligomer, and active site represents a significant step forward toward understanding how bacterial pore-forming toxins are activated. The structural and possibly functional distinction between TAATs and all known GNAT families also identifies these enzymes as possible targets for developing new antivirulence compounds.

Methods

Cloning and Site-Directed Mutagenesis. The *apxIC* gene was PCR-amplified from *A. pleuropneumoniae* (ATCC27088) genomic DNA using primers ApxC QE80 F and ApxC QE80 R (all primer sequences are given in Table S3) engineered to encode restriction sites in the 5′ ends. The PCR product was digested with BamHI and HindIII and ligated into pQE80 (Qiagen). Site-specific mutations were introduced by using QuikChange (Stratagene). Clones from PCR-amplified DNA were sequenced (Source Bioscience). TAAT genes from uropathogenic *E. coli,* enterohemorrhagic *E. coli*, *Bordetella pertussis, Mannheimia hemolytica*, *Aggregatibacter actinomycetemcomitans,* and *Aeromonas hydrophila* were cloned with a similar strategy. For bacterial two-hybrid ApxC or variant was amplified with primers ApxC HindIII F/ApxC SacI REV (ApxC HindIII F/ApxC 161 R for the Δ162–172 construct), digested with HindIII and SacI, and cloned into pUT18 and pKNT25, resulting in the adenylate cyclase fragment being placed at the C terminus of ApxC. Similarly *E. coli* ACP and a fragment corresponding to HlyA residues 679–736 were amplified with primer pairs ACP F/ACP R and HlyA

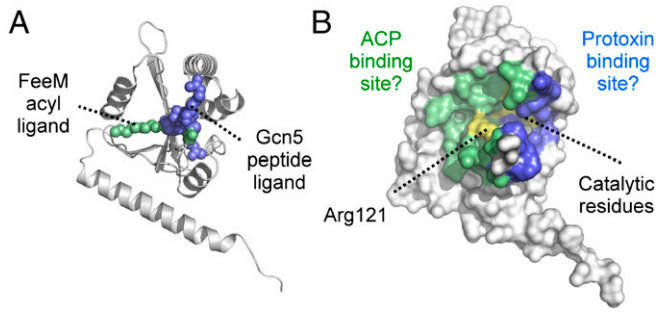

Fig. 7. A model for ternary complex TAAT function. (*A*) Location of the acyltyrosine ligand (green) from the FeeM:acyltyrosine costructure and histone H3 peptide (blue) from the Gcn5:CoA:H3 peptide ternary complex after superposition with ApxC (white). (*B*) Proposed locations of the acylACP and protoxin binding sites on opposite faces of the TAAT enzyme (ACP-binding site green; protoxin binding site blue). The substrates are predicted to meet at the crux of the split β-sheet where the catalytic residues (yellow) are located. Arg121 (essential for ACP binding) is annotated.

679 F/HlyA 736 R, respectively, digested with HindIII-SacI, and cloned into pUT18 digested with the same enzymes.

ApxC Protein Expression and Purification. N-terminally His-tagged ApxC was produced in *E. coli* C41 cells (44) bearing plasmid pQE80/ApxC, grown in 2TY medium at 30 °C until A_{600} 0.6, then the temperature was reduced to 18 °C and expression induced with 0.1 mM IPTG for 16 h. Cells were harvested by centrifugation (10,000 × *g*) and resuspended in 50 mM Hepes pH 7.5, 400 mM NaCl, 5% (vol/vol) glycerol supplemented with Complete EDTA-free protease inhibitor mixture (Roche), and broken by two passages through a cell disruptor (30,000 psi). Unbroken cells were removed by ultracentrifugation at 150,000 × *g* for 1 h at 4 °C. The supernatant was supplemented with 4 mM Imidazole and added to Profinity IMAC resin (Bio-Rad) for 1 h at 4 °C. Protein-bound resin was washed with 25 mM Hepes pH 7.5, 500 mM NaCl, 8 mM Imidazole, 0.1% Triton X-100, then with the same buffer without Triton X-100, and eluted with 25 mM Hepes pH 7.5, 200 mM NaCl, 250 mM Imidazole. Protein fractions were pooled, concentrated, and loaded onto a Superdex S75 column equilibrated in 25 mM Hepes pH 7.5, 200 mM NaCl. Peak fractions were concentrated to 5 mg/mL in an Amicon 10-kDa centrifugal filter device (Millipore). For phase determination, SeMet-labeled ApxC was produced in *E. coli* B834 transformed with pQE80/ApxC. Cells grown in M9 minimal medium supplemented with 50 mg/mL methionine to A_{600} 0.8 were resuspended in minimal media lacking methionine for 20 min and selenomethionine was added to 50 mg/mL final concentration. After another 20 min, expression was induced by IPTG as before, and SeMet-labeled protein purified as unlabeled protein except that 1 mM TCEP (Merck) was added to buffers. Protein identity and SeMet incorporation were verified by mass spectrometry (PNAC; Cambridge Department of Biochemistry).

Crystallization and Structure Determination. ApxC crystals were obtained via the sitting-drop vapor diffusion method by using crystallization reagent composed of 32% (wt/vol) PEG 300, 100 mM Phosphate/Citrate Buffer pH 4.4, and a protein solution of 5 mg/mL ApxC in 150 mM NaCl, 20 mM Hepes pH 7.5. Sitting drops were formed by mixing 2 μL of protein solution with 1 μL of crystallization reagent and equilibration over a 500-μL volume of the crystallization reagent at 15 °C in 24-well plates. Crystals appeared after 3 d and were harvested a few days later. For freezing, crystals were transferred into a cryoprotectant solution composed of 20 mM Phosphate/Citrate buffer, 50 mM NaCl, 7.5% (wt/vol) PEG 300, and 10% (vol/vol) glycerol, then to the same solution containing 20% (vol/vol) glycerol before harvesting with a cryoloop and flash-freezing in liquid nitrogen. Crystals of the selenomethionine derivative were prepared similarly, although they were typically much smaller than those of the native protein. X-ray diffraction data were collected on beamlines I02 and I04 of Diamond Light Source, UK. Structure determination used the Se-SAD method of phase determination aided by programs from both the PHENIX (45) and CCP4 software suites (46). X-ray data were indexed and integrated by using MOSFLM (47) and scaled by using SCALA. A partial substructure of Selenium sites was identified by using HYSS (48) and experimental phases calculated with PHASER (49). Density modification using PARROT (50) benefitted from identification of a twofold noncrystallographic symmetry relating a subset of the heavy atom sites. Manual inspection of the density-modified maps identified further selenium sites and derivation of the likely fourfold NCS operators relating each of the four monomers in the asymmetric unit. After further density modification by using fourfold NCS averaging, an initial selenoApxC model was built with BUCCANEER (51) and completed with iterative rounds of manual model-building with COOT (52) and refinement with REFMAC (53) using NCS restraints. Refinement was switched to an isomorphous, high-resolution, native dataset, maintaining the same "free" reflection set, with NCS restraints on the monomers relaxed and solvent components such as ordered water molecules and crystallization reagents added to complete the crystallographic model. Further model validation used PROCHECK (54) and RAMPAGE (55). The final model comprises 4 molecules of ApxC, 219 waters, and a citrate molecule. Atomic coordinates and structure factors are deposited in the Protein Data Bank, ID code 4WHN.

Structure Analysis, Sequence Alignments and Homology Modeling. Buried surface areas were calculated by using AREAIMOL (56) with a solvent probe radius of 1.2 Å. Predicted dimer stability was assessed with PISA (34). Root mean square deviations (rmsds) were calculated over all C_α atoms after alignment by using superpose in the CCP4 suite (57). Figures were prepared in Pymol (58). For homology modeling, a single monomer of ApxC was used as a template on which to impose the sequence of other TAATs by using the Phyre2 server in sequence threading mode (59). For multiple sequence alignments, amino acid sequences were processed by using Clustal Omega (60) and conserved residues mapped with ESPript (61). Further editing to add the secondary structure elements and essential residues was performed manually by using the GNU Image Manipulation Program.

SAXS Analysis. SAXS was conducted at the PETRA P12 SAXS beamline in Hamburg, Germany (Deutsches Elektronen-Synchrotron) using purified ApxC protein in buffer 20 mM Hepes pH 7.5, 150 mM NaCl. Data used in final analysis is the average of four separate samples (each the average of 10–20 frames) collected by using 56 μM ApxC. Computational analysis of the data used the ATSAS package of programs (62). Inspection and characterization of the data used PRIMUS (63) and GNOM (64), whereas ab initio shape reconstruction used DAMMIN (65). The final particle was constructed by averaging 20 independent models with DAMAVER (66) before filtering with DAMFILT (66). The final model also used enforced P2 symmetry, but runs without symmetry restraints produced similar results. CRYSOL (67) was used to compare predicted scattering behavior of dimers/monomers in our crystal structure with scattering data obtained in solution.

Hemolysis Assay of Toxin Activation. *E. coli* MC1061 (TolC -positive) cells bearing plasmids expressing the export system HlyBD (LG575; ref. 68) and HlyA (LG583; ref. 68) were transformed with pQE80 vector alone or expressing wild-type or mutant ApxC. Bacteria were grown overnight on blood agar plates supplemented with 20 μg/mL chloramphenicol and 50 μg/mL carbenicillin. LB cultures supplemented with antibiotics to maintain plasmid selection were grown to A_{600} 0.8 and cells were removed. Dilutions of supernatant containing secreted HlyA were added to 2% (vol/vol) washed horse erythrocytes, 20 mM $CaCl_2$, 150 mM NaCl, and incubated 30 min at 42 °C. Hemolytic activity was assayed as hemoglobin release at A_{543} after subtracting the absorbance of the vector control. Assays were performed in triplicate, with each absorbance measurement itself the mean of triplicates.

Bacterial Two-Hybrid Assays. *E. coli* BTH101 was cotransformed with recombinant pUT18 and pKNT25 plasmids expressing the protein of interest fused to either the T18 or T25 domain of *B. pertussis* CyaA and grown at 30 °C. Overnight culture (LB with 100 μg/mL carbenicillin and 50 μg/mL kanamycin) grown at 30 °C was spotted (2 μL) on LB agar containing 0.5 mM IPTG and 40 μg/mL X-Gal. After overnight incubation at 30 °C, plates were photographed under white light. β-galactosidase assays were performed as described (69).

Cross-Linking. Cysteine variants of ApxC were purified as for the wild-type protein except that 2 mM TCEP was added to all buffers. For cross-linking reactions, TCEP was removed by desalting column and 100 μM $CuCl_2$ was added to 50 μg of protein in 25 mM Hepes pH 7.5, 150 mM NaCl and incubated at 25 °C for 30 min before addition of 10 mM EDTA (to chelate copper) and 5 mM NEM (to block free cysteines). Reactions were analyzed by SDS/PAGE. For the control reaction (preblock), samples were pretreated with EDTA and NEM before addition of $CuCl_2$.

Western Blot Analysis. Samples were resolved on SDS/PAGE, transferred to nitrocellulose membrane, and immunoblotted with anti-His (Qiagen) or anti-cyaA (Santa Cruz) antisera as appropriate. Immunoblots were revealed with a chemiluminescence kit (GE Healthcare ECL) or fluorescence imaging (Odyssey Licor), respectively.

ACKNOWLEDGMENTS. We thank staff at Diamond Light Source (UK) and European Molecular Biology Laboratory Hamburg for beam line provision, Dr. Phil Hinchliffe for assisting data collection, Dr. Len Packman for mass spectrometry services, and Gouzel Karimova (Institut Pasteur, Paris) for the generous gift of plasmids pKNT25 and pUT18. This work was supported by UK Medical Research Council and the Wellcome Trust Grants (to C.H. and V.K.).

1. Ulett GC, et al. (2013) Uropathogenic Escherichia coli virulence and innate immune responses during urinary tract infection. *Curr Opin Microbiol* 16(1):100–107.
2. Gur C, et al. (2013) Natural killer cell-mediated host defense against uropathogenic E. coli is counteracted by bacterial hemolysinA-dependent killing of NK cells. *Cell Host Microbe* 14(6):664–674.
3. Uhlén P, et al. (2000) Alpha-haemolysin of uropathogenic E. coli induces Ca^{2+} oscillations in renal epithelial cells. *Nature* 405(6787):694–697.
4. Stanley P, Koronakis V, Hughes C (1998) Acylation of Escherichia coli hemolysin: A unique protein lipidation mechanism underlying toxin function. *Microbiol Mol Biol Rev* 62(2):309–333.
5. Wiles TJ, Dhakal BK, Eto DS, Mulvey MA (2008) Inactivation of host Akt/protein kinase B signaling by bacterial pore-forming toxins. *Mol Biol Cell* 19(4):1427–1438.
6. Smith YC, Rasmussen SB, Grande KK, Conran RM, O'Brien AD (2008) Hemolysin of uropathogenic Escherichia coli evokes extensive shedding of the uroepithelium and hemorrhage in bladder tissue within the first 24 hours after intraurethral inoculation of mice. *Infect Immun* 76(7):2978–2990.
7. Dhakal BK, Mulvey MA (2012) The UPEC pore-forming toxin α-hemolysin triggers proteolysis of host proteins to disrupt cell adhesion, inflammatory, and survival pathways. *Cell Host Microbe* 11(1):58–69.
8. Nagamatsu K, et al. (2015) Dysregulation of Escherichia coli α-hemolysin expression alters the course of acute and persistent urinary tract infection. *Proc Natl Acad Sci USA* 112(8):E871–E880.
9. Koronakis V, Eswaran J, Hughes C (2004) Structure and function of TolC: The bacterial exit duct for proteins and drugs. *Annu Rev Biochem* 73:467–489.
10. Koronakis V, Koronakis E, Hughes C (1989) Isolation and analysis of the C-terminal signal directing export of Escherichia coli hemolysin protein across both bacterial membranes. *EMBO J* 8(2):595–605.
11. Thanabalu T, Koronakis E, Hughes C, Koronakis V (1998) Substrate-induced assembly of a contiguous channel for protein export from E.coli: Reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *EMBO J* 17(22): 6487–6496.
12. Balakrishnan L, Hughes C, Koronakis V (2001) Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by Escherichia coli. *J Mol Biol* 313(3):501–510.
13. Issartel JP, Koronakis V, Hughes C (1991) Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation. *Nature* 351(6329):759–761.
14. Stanley P, Packman LC, Koronakis V, Hughes C (1994) Fatty acylation of two internal lysine residues required for the toxic activity of Escherichia coli hemolysin. *Science* 266(5193):1992–1996.
15. Stanley P, Koronakis V, Hardie K, Hughes C (1996) Independent interaction of the acyltransferase HlyC with two maturation domains of the Escherichia coli toxin HlyA. *Mol Microbiol* 20(4):813–822.
16. Ludwig A, et al. (1996) Analysis of the in vivo activation of hemolysin (HlyA) from Escherichia coli. *J Bacteriol* 178(18):5422–5430.
17. Lim KB, et al. (2000) Escherichia coli alpha-hemolysin (HlyA) is heterogeneously acylated in vivo with 14-, 15-, and 17-carbon fatty acids. *J Biol Chem* 275(47): 36698–36702.
18. Hardie KR, Issartel JP, Koronakis E, Hughes C, Koronakis V (1991) In vitro activation of Escherichia coli prohaemolysin to the mature membrane-targeted toxin requires HlyC and a low molecular-weight cytosolic polypeptide. *Mol Microbiol* 5(7):1669–1679.
19. Hackett M, Guo L, Shabanowitz J, Hunt DF, Hewlett EL (1994) Internal lysine palmitoylation in adenylate cyclase toxin from Bordetella pertussis. *Science* 266(5184): 433–435.
20. Forestier C, Welch RA (1990) Nonreciprocal complementation of the hlyC and lktC genes of the Escherichia coli hemolysin and Pasteurella haemolytica leukotoxin determinants. *Infect Immun* 58(3):828–832.
21. Gygi D, et al. (1990) Isolation of the Actinobacillus pleuropneumoniae haemolysin gene and the activation and secretion of the prohaemolysin by the HlyC, HlyB and HlyD proteins of Escherichia coli. *Mol Microbiol* 4(1):123–128.
22. Koronakis V, Cross M, Senior B, Koronakis E, Hughes C (1987) The secreted hemolysins of Proteus mirabilis, Proteus vulgaris, and Morganella morganii are genetically related to each other and to the alpha-hemolysin of Escherichia coli. *J Bacteriol* 169(4): 1509–1515.
23. Trent MS, Worsham LM, Ernst-Fonberg ML (1999) HlyC, the internal protein acyltransferase that activates hemolysin toxin: Roles of various conserved residues in enzymatic activity as probed by site-directed mutagenesis. *Biochemistry* 38(29): 9541–9548.
24. Trent MS, Worsham LM, Ernst-Fonberg ML (1999) HlyC, the internal protein acyltransferase that activates hemolysin toxin: Role of conserved histidine, serine, and cysteine residues in enzymatic activity as probed by chemical modification and site-directed mutagenesis. *Biochemistry* 38(11):3433–3439.
25. Trent MS, Worsham LM, Ernst-Fonberg ML (1999) HlyC, the internal protein acyltransferase that activates hemolysin toxin: The role of conserved tyrosine and arginine residues in enzymatic activity as probed by chemical modification and site-directed mutagenesis. *Biochemistry* 38(27):8831–8838.
26. Holm L, Rosenström P (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res* 38(Web Server issue):W545–W549.
27. Dyda F, Klein DC, Hickman AB (2000) GCN5-related N-acetyltransferases: A structural overview. *Annu Rev Biophys Biomol Struct* 29:81–103.
28. Vetting MW, et al. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys* 433(1):212–226.
29. Gu L, et al. (2007) GNAT-like strategy for polyketide chain initiation. *Science* 318(5852):970–974.
30. Rojas JR, et al. (1999) Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* 401(6748):93–98.
31. Clements A, et al. (1999) Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A. *EMBO J* 18(13): 3521–3532.
32. Yan Y, Harper S, Speicher DW, Marmorstein R (2002) The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nat Struct Biol* 9(11):862–869.
33. Van Wagoner RM, Clardy J (2006) FeeM, an N-acyl amino acid synthase from an uncultured soil microbe: Structure, mechanism, and acyl carrier protein binding. *Structure* 14(9):1425–1435.
34. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.
35. Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* 95(10): 5752–5756.
36. Sun B, et al. (2011) Regulation of the histone acetyltransferase activity of hMOF via autoacetylation of Lys274. *Cell Res* 21(8):1262–1266.
37. Masoudi A, Raetz CRH, Zhou P, Pemble CW, 4th (2014) Chasing acyl carrier protein through a catalytic cycle of lipid A production. *Nature* 505(7483):422–426.
38. Nguyen C, et al. (2014) Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* 505(7483):427–431.
39. Lau OD, et al. (2000) p300/CBP-associated factor histone acetyltransferase processing of a peptide substrate. Kinetic analysis of the catalytic mechanism. *J Biol Chem* 275(29):21953–21959.
40. Berndsen CE, Albaugh BN, Tan S, Denu JM (2007) Catalytic mechanism of a MYST family histone acetyltransferase. *Biochemistry* 46(3):623–629.
41. Stanley P, Hyland C, Koronakis V, Hughes C (1999) An ordered reaction mechanism for bacterial toxin acylation by the specialized acyltransferase HlyC: Formation of a ternary complex with acylACP and protoxin substrates. *Mol Microbiol* 34(5):887–901.
42. Worsham LM, Trent MS, Earls L, Jolly C, Ernst-Fonberg ML (2001) Insights into the catalytic mechanism of HlyC, the internal protein acyltransferase that activates Escherichia coli hemolysin toxin. *Biochemistry* 40(45):13607–13616.
43. Yang C, Wu J, Sinha SH, Neveu JM, Zheng YG (2012) Autoacetylation of the MYST lysine acetyltransferase MOF protein. *J Biol Chem* 287(42):34917–34926.
44. Miroux B, Walker JE (1996) Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 260(3):289–298.
45. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
46. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.
47. Read RJ (2007) *Evolving Methods for Macromolecular Crystallography*, ed Sussman JL (Springer Netherlands, Dordrecht, The Netherlands).
48. Grosse-Kunstleve RW, Adams PD (2003) Substructure search procedures for macromolecular structures. *Acta Crystallogr D Biol Crystallogr* 59(Pt 11):1966–1973.
49. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
50. Cowtan K (2010) Recent developments in classical density modification. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):470–478.
51. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 62(Pt 9):1002–1011.
52. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
53. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3): 240–255.
54. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 26(2):283–291.
55. Lovell SC, et al. (2003) Structure validation by Calpha geometry: Phi,psi and Cbeta deviation. *Proteins* 50(3):437–450.
56. Lee B, Richards FM (1971) The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* 55(3):379–400.
57. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2256–2268.
58. Schrodinger L (2010) *The PyMOL Molecular Graphics System,* Version~1.3r1.
59. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the Web: A case study using the Phyre server. *Nat Protoc* 4(3):363–371.
60. Goujon M, *et al*. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38 (Web Server issue):W695–W699.
61. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42 (Web Server issue):W320–W324.

62. Petoukhov MV, et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Cryst* 45(Pt 2):342–350.
63. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst* 36(5):1277–1282.
64. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* 25(4):495–503.
65. Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76(6):2879–2886.
66. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Cryst* 36(3):860–864.
67. Svergun D, Barberato C, Koch MHJ (1995) CRYSOL – a program to evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Cryst* 28(6):768–773.
68. Mackman N, Nicaud JM, Gray L, Holland IB (1985) Genetical and functional organisation of the Escherichia coli haemolysin determinant 2001. *Mol Gen Genet* 201(2):282–288.
69. Battesti A, Bouveret E (2012) The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli. *Methods* 58(4):325–334.