

Structural Basis for Toxin Inhibition in the VapXD Toxin-Antitoxin System

Marie B. Bertelsen^{a, 1}, Meriem Senissar^{a, 1, 2}, Maja H. Nielsen^{a, 3}, Francesco Bisiak^a, Marta V. Cunha^a, Ashley L. Molinaro^b, Dayle A. Daines^b, and Ditlev E. Brodersen^{1, 4, *}

^aDepartment of Molecular Biology and Genetics, Aarhus University, Aarhus C, DK-8000, Denmark. ^bDepartment of Biological Sciences, Old Dominion University, Norfolk, VA 23529, USA

¹These authors contributed equally.

²Present address: Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, DK-2800, Denmark.

³Present address: Department of Forensic Medicine, Aarhus University, Aarhus N, DK-8200, Denmark.

⁴Lead contact

^{*}To whom correspondence should be addressed. Email address: deb@mbg.au.dk, Twitter @ditlevbrodersen.

SUMMARY

Bacterial type II toxin-antitoxin modules encode a toxic protein that down-regulates metabolism and a specific antitoxin that binds and inhibits the toxin during normal growth. In non-typeable *Haemophilus influenzae*, a common cause of infections in humans, the *vapXD* locus was found to constitute a functional TA module and contribute to pathogenicity, however, the mode of action of VapD and the mechanism of inhibition by the VapX antitoxin remain unknown. Here, we report the structure of the intact *H. influenzae* VapXD complex revealing an unusual 2:1 toxin-antitoxin molecular stoichiometry where a Cas2-like homodimer of VapD binds a single VapX antitoxin. VapX consists of a oligonucleotide/oligosaccharide-binding (OB) domain that docks into an asymmetrical cavity on the toxin dimer. Structures of isolated VapD further reveal how a symmetrical toxin homodimer adapts to interacting with an asymmetrical antitoxin and suggest how a primordial TA system evolved to become part of CRISPR-Cas immunity systems.

Keywords: *Haemophilus influenzae*, VapD, Cas2, OB fold, RNase, CRISPR-Cas

INTRODUCTION

Bacterial type II toxin-antitoxin (TA) modules encode two small protein components: a stable toxin capable of down-regulating cellular metabolism and an antitoxin that specifically binds and inhibits the toxin (Harms et al., 2018). Antitoxins are metabolically and structurally unstable and must thus be continuously replenished in order to maintain toxin inhibition. This principle forms the basis for plasmid maintenance via post-segregational killing, in which cells that do not inherit a TA-containing plasmid fail to survive because they no longer express the antitoxin, and abortive infection that increases bacterial survival during phage infection (Harms et al., 2018). Most type II antitoxins are composed of two distinct domains; an N-terminal DNA-binding domain that regulates transcription from the TA locus through direct interaction with the promoter, and an intrinsically unstructured C-terminal region responsible for toxin binding and inhibition (Harms et al., 2018). During nutritional or oxidative stress, the antitoxins are specifically degraded by cellular proteases such as Lon, thus providing a metabolic cue to toxin activation (Page and Peti, 2016). Once activated, type II toxins typically act as nucleases that modulate gene expression through degradation of specific, stable RNAs including tRNA, rRNA, and mRNA (Goeders and Van Melder, 2014). By targeting central cellular processes that affect bacterial growth and cell division, the toxins thus generate a fitness advantage for their hosts by allowing them to quickly adapt their growth rate to environmental changes (LeRoux et al., 2020; Page and Peti, 2016).

Non-typeable (non-encapsulated) *Haemophilus influenzae* (NTHi) is a human pathogen and common cause of infections such as recurrent middle ear infection (*otitis media*) and *sinusitis* in children, and severe lower respiratory tract infections (*pneumonia*) in patients with chronic obstructive pulmonary disease (COPD) (Murphy, 2003). The NTHi genome contains four toxin-antitoxin operons, *toxAvapA*, *vapBC-1*, *vapBC-2*, and *vapXD*, of which *vapBC-1* and

vapBC-2 are canonical and well-characterised *vapBC*-type TA modules containing a PIN-domain RNase toxin (VapC) (Daines et al., 2007; Harms et al., 2018; Molinaro et al., 2019). The *vapXD* locus encodes a non-canonical toxin-antitoxin system, which was linked to survival and virulence of NTHi in an animal model system and demonstrated to act as a *bona fide* TA module *in vivo* (Daines et al., 2004; Ren et al., 2012) (**Figure 1A**). The VapD toxin was shown to be active as a homodimer *in vivo* but very little is known about the structure and function of the expressed VapXD protein complex (Ren et al., 2012). A structure exists of a homologous VapD protein from *Helicobacter pylori*, which is remarkable because it is expressed from a locus that does not appear to encode an antitoxin and the protein is thus apparently is not part of a canonical TA system (Kwon et al., 2012). The *H. pylori* structure revealed a clear structural similarity to Cas2, which is required for spacer integration during CRISPR-Cas adaptation (Nunez et al., 2014). This observation as well as comparative sequence studies have fuelled speculations that the bacterial CRISPR-Cas immunity systems might have in fact evolved from a primordial *vapXD*-type TA system (Makarova et al., 2012).

Most characterised Cas2-type proteins display nonspecific endonuclease activity, with some preferring single-stranded RNA (Beloglazova et al., 2008) and others double or single-stranded DNA (Ka et al., 2014; Nam et al., 2012). The activity has been found to depend on divalent cations (Mg^{2+} or Mn^{2+}) as well as a strictly conserved aspartic acid residue which was proposed to be involved in coordination of the ion in the active site (Beloglazova et al., 2008; Ka et al., 2017; Ka et al., 2014; Nam et al., 2012). Surprisingly, however, the RNase activity of Cas2 does not appear to be required for its role in CRISPR-Cas adaptation, where the protein functions as a spacer of dimeric Cas1 subunits, which in turn catalyse the dsDNA cleavage reaction required for spacer acquisition (Nunez et al., 2014). *H. pylori* VapD appears to be active solely on single-stranded RNA, thus suggesting that VapD orthologues, when part of a

toxin-antitoxin system, function as RNases similarly to RelE, MazF, and VapC, but independently of cations (Kwon et al., 2012; Pedersen et al., 2003; Zhang et al., 2003). Despite this, mutation of the strictly conserved aspartic acid residue abolishes VapD activity, suggesting that the protein contains an active site similar to the Cas2 proteins but possibly uses a different catalytic mechanism (Kwon et al., 2012).

Understanding the mechanistic details underlying toxin inhibition and activity in the VapXD TA system is highly relevant to understand their influence of pathogenicity in *H. influenzae* and could prove useful for drug development against infectious diseases. Moreover, insights into the VapXD system could support the proposed evolutionary links between CRISPR-Cas and the TA modules. Here, we provide structural insights into the architecture of the intact VapXD toxin-antitoxin complex and toxin activation. The crystal structure of the NTHi VapXD complex confirms the similarity to Cas2 and reveals an unusual 2:1 toxin-antitoxin complex in which an otherwise symmetrical toxin homodimer adapts to binding of an asymmetrical antitoxin. Surprisingly, the VapX antitoxin contains neither an intrinsically disordered region nor a DNA-binding domain but consists of a single oligonucleotide/oligosaccharide-binding (OB) fold, which intercalates between monomers of the toxin dimer. Our results thus shed light on the mechanism of toxin inhibition in this unique TA system and suggest that transcriptional regulation is markedly different from canonical type II TA systems.

RESULTS

The structure of the VapXD complex reveals an unusual 2:1 stoichiometry

In order to understand the molecular basis for VapD inhibition by its cognate VapX antitoxin, we determined the structure of the NTHi VapXD protein complex by x-ray crystallography. Briefly, seleno-methionine (Se-Met) labelled VapXD complex was produced by

overexpression of the natural, full-length operon, modified to include a C-terminal His-tag on the VapD toxin, in the methionine auxotroph *Escherichia coli* strain B834 (DE3) growing in minimal medium. The labelled protein was purified to homogeneity, crystallised in the cubic space group $I4_132$, and the structure determined to 3.2 Å using Se SAD phasing, followed by manual model building and refinement to a final $R_{\text{work}} = 25\%$ and $R_{\text{free}} = 29\%$ (see STAR Methods and **Table 1** for details). The relatively high crystallographic R factors are likely due to the an unusually high solvent content (77.5%, $V_M = 5.46 \text{ Å}^3/\text{Da}$) as well as the presence pseudosymmetry in the cubic space group resulting in extended solvent channels with weak unexplained and repetitive density, pathologies that could not be accounted for using regular refinement protocols. Modelling of this spurious density did not improve refinement. Nevertheless, the electron density is generally well defined and the model is nearly complete, with a few terminal residues missing in each chain likely due to flexibility (**Figure S1A**). The structure consists of two VapXD₂ heterotrimers in the crystallographic asymmetric unit, each consisting of a VapD homodimer (2·92 aa) and a single VapX antitoxin (63 aa) molecule, which is located in a cleft formed between the two VapD subunits (**Figure 1B**). Although the VapD toxin is known to be active as a dimer, such a 2:1 molecular stoichiometry is highly unusual for type II TA systems, for which the toxins usually always associate 1:1 with their cognate antitoxin proteins (Page and Peti, 2016).

VapXD₂ heterotrimers associate into higher order structures

Symmetry expansion of the crystal structure reveals that three VapXD₂ heterotrimers come together to form a larger, ring-shaped structure, which is stabilised by the conserved charged residues Arg79 and Glu81 that form a circle of alternating positive and negative charges at the centre (**Figure 1C**, inset, and **Figure 1D**). PISA analysis (Krissinel and Henrick, 2007) suggests several possible higher order assemblies, including the VapXD₂ heterotrimer, VapX₃D₆ (the

ring) as well as VapX₆D₁₂ (two stacked rings), and even VapX₁₂D₂₄ (four stacked rings, **Figure S1B**) all with favourable calculated free energies of dissociation (**Table S1**), suggesting that they all potentially could be stable and in theory biologically relevant. To investigate which of these species is most prevalent in solution, we performed gel filtration chromatography, dynamic light scattering (DLS) and protein cross-linking. During gel filtration chromatography, the VapXD complex elutes with an apparent molecular weight of 150 kDa (**Figure S2A**), which is significantly larger than expected for a single VapXD₂ trimer (32.3 kDa) and closer to the expected mass of the ring structure ($3 \cdot 32.3 = 96.9$ kDa). The presence of higher order assemblies is further supported by glutaraldehyde cross-linking *in vitro*, which reveals several species with molecular masses up to around 90 kDa (**Figure S2B**). Finally, we measured the DLS from a solution containing purified VapXD, which provides an overall average size (not mass) estimate (**Figure S2C**). This revealed a largely monodisperse solution, with a single species having an average hydrodynamic diameter of 10.7 ± 0.1 nm. Calculation of theoretical hydrodynamic diameters of the various assemblies suggested by PISA using HYDROPRO (Ortega et al., 2011) shows that the measured size fits both one and two stacked rings, which have calculated hydrodynamic diameters of 8.1 and 9.8 nm, respectively. Together, these results support that the VapXD complex forms complexes of higher order than a single VapXD₂ heterotrimer in solution, and most likely in the form of trimers of trimers (rings), and possibly stacked as suggested by PISA and DLS. Moreover, we note that ring stacking is coaxial, which in theory would allow binding of longer macromolecules through the central hole (Figure S1B). Further studies will be necessary to determine the stoichiometry of these assemblies *in vivo* as well as their biological relevance.

The NTHi VapD toxin dimer is structurally similar to Cas2

The VapD toxin contains a modified ferredoxin fold ($\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$), where each of the two α -helices have been split into two shorter helices connected by short loops, resulting in a $\beta_1\alpha_1'\alpha_1\beta_2\beta_3\alpha_2\alpha_2'\beta_4$ topology (**Figure 2A**). In the VapXD complex, two VapD monomers form a butterfly-shaped dimer with a deep cleft between them. This dimerisation is primarily mediated by a fifth C-terminal β -strand (β_5) outside the core ferredoxin fold, via formation of an extended β -sheet with the corresponding region of the other VapD monomer (**Figure 2A** and **2B**). The structure of NTHi VapD is highly similar to *H. pylori* VapD both at the sequence and structure level (40% sequence identity, 68% sequence similarity, r.m.s.d. C α atoms 0.98 Å, **Figure 2C**) (Kwon et al., 2012). In NTHi VapD, β_2 and β_3 are very short (only two residues each), and β_2 is slightly distorted from an ideal β strand conformation. This contributes to the dimerisation interface by providing shape complementarity at the bottom of the cleft. Interestingly, the sequence of β_4 is well conserved among VapD homologues (**Figure 1D**) suggesting that a similar mode of dimerisation/cleft is likely to be found for these and that sidechains could be functionally important. Indeed, the conserved hydrophobic sidechains of Leu44, Phe78, and Ile80 appear to form a hydrophobic core that further contributes to the stability of the dimer. In addition, the conserved charged residues Arg79 and Glu81, which are involved in formation of the ring-shaped higher-order structure also interact across monomers at the bottom of the dimer (**Figure S3A**).

The VapD dimer is structurally similar to Cas2 as found within the Cas1-Cas2 complex required for spacer cleavage during the CRISPR-Cas adaptation stage in many microorganisms (Beloglazova et al., 2008; Kwon et al., 2012; Nunez et al., 2014). The NTHi VapD monomer superposes with *Sulfolobus solfataricus* Cas2 (PDB:2I8E) with an overall r.m.s.d. of C α atoms of 2.6 Å. Interestingly, alignment of complete dimers is significantly worse (5.5 Å), suggesting that mostly the monomer fold and less so dimer interaction, is conserved between the TA

systems and CRISPR-Cas. Comparison of the structures show that the charged Arg79/Glu81 interaction is unique to VapD, while the Cas2 dimer does have a cluster of hydrophobic residues between the monomers in a similar way as VapD. Compared to *E. coli* Cas2, which has been structurally characterised in complex with Cas1 and DNA, NTHi VapD has two short, additional α -helices (α_1' and α_2') at the top, but lacks a β -hairpin at the C-terminus, which is responsible for Cas1 interaction in the Cas1-Cas2 complex (**Figure 2C**) (Nunez et al., 2014). Moreover, the additional helices in VapD, which are involved in VapX antitoxin binding, would overlap with DNA in the Cas1-Cas2-DNA structure explaining their absence in Cas2 (Nunez et al., 2015).

Mutational studies of *S. solfataricus* Cas2 and *H. pylori* VapD have mapped the nuclease active site to the cleft between the two monomers (Beloglazova et al., 2008; Kwon et al., 2012). In particular, Asp10 was found to be crucial for activity of *S. solfataricus* Cas2 (Beloglazova et al., 2008). The two equivalent Asp10 residues are adjacent to each other in the cleft between the molecules of the corresponding Cas2 dimer, and have been proposed to coordinate a Mg^{2+} ion, which in turn could prepare a water molecule for a nucleophilic attack on the phosphate backbone of RNA and/or DNA (Beloglazova et al., 2008). This acidic residue is conserved among both Cas2 and VapD homologues and corresponds to Glu9 in *E. coli* Cas2 and Asp7 of NTHi VapD (**Figure 2C** and **1D**). In the structure of NTHi VapD bound to its antitoxin, the two Asp7 residues are located 11.9 Å apart (opposed to 6.5 Å in *S. solfataricus* Cas2), which is clearly too far for coordination of a single Mg^{2+} ion (**Figure S3B**). It thus seems more likely that if divalent cations are involved in catalysis that one is bound in each site in a symmetrical fashion.

VapD is structurally similar to the antitoxin of the type V TA system, GhoST

The GhoST system from *E. coli* is the only known type V TA system, which is defined by the fact that the antitoxin (GhoS) is an RNase that cleaves the toxin mRNA under normal conditions, thus preventing its expression and activity (Wang et al., 2012). Curiously, the structure of GhoS (**Figure 2D**) revealed a ferredoxin-like fold similar to *S. solfataricus* Cas2 including the conserved catalytic residues (Wang et al., 2012). The monomer of NTHi VapD superposes well with *E. coli* GhoS antitoxin with a root-mean-square deviation of the C α atoms of 2.4 Å (**Figure 2D**). When aligning the two structures, the conserved catalytic residue Asp7 in VapD superimposes with Asp17 in GhoS, which is oriented into the cleft of the dimer, suggesting the two residues could have the same function (**Figure 2D**, inset). However, GhoS does not contain the extra C-terminal β -strand and does not form a dimer (Wang et al., 2012). Instead, GhoS has a small β -hairpin near its C-terminus, which precludes the dimer interaction (**Figure 2D**). Interestingly, the GhoS antitoxin is specific for its substrate, the mRNA of its cognate toxin (Wang et al., 2012). Mutational studies of five conserved residues revealed that Arg30 and Phe57 are crucial for the activity of GhoS *in vitro*, and furthermore that Arg30 is important for the activity *in vivo* (Wang et al., 2012), but none of these residues are conserved in VapD. In summary, it is likely that VapD and GhoS derived from the same, original ferredoxin-like fold, possibly with similar reaction mechanisms involving an Asp, and later diverged in terms of oligomeric state and substrate specificity.

The VapX antitoxin contains an OB fold with swapped architecture

VapX is a small, monomeric and single-domain antitoxin with a compact $\alpha(\beta)_4\alpha$ fold consisting of an N-terminal region composed of a loop and an α -helix linked to a core fold of four highly curved and antiparallel β -sheets connected by loops and finally a C-terminal region also composed by an α -helix and a loop (**Figure 3A**). VapX structurally resembles the Oligonucleotide/oligosaccharide-Binding (OB) domain proteins, of which translation

elongation factor P is an example (EF-P, Z score 6.1, r.m.s.d. C α atoms 1.8 Å, **Figure 3B**). VapX superimposes well with EF-P, however, closer inspection reveals that its topology is different in that it contains a swapped OB fold in which the β 4-strand, which runs alongside the β 1 in the canonical OB fold and forms a classical Greek key motif, is replaced by an N-terminal β -strand (β 1', **Figure 3A**). The OB domain is found ubiquitously in all kingdoms of life despite a general lack of sequence similarity (Murzin, 1993), and is usually involved in binding either RNA, ssDNA, oligosaccharides, or other proteins (Arcus, 2002). Regardless of binding target, OB fold proteins tend to use the same face for interactions, a region which is bounded by several loop regions (**Figure 3B**). Specifically, binding of single-stranded nucleic acids is accomplished through aromatic stacking, hydrogen bonding, and polar interactions by side chains located between β -strands 1 and 2 (L12), and between β -strands 4 and 5 (L45). Together, these loops generate a cleft that runs across the surface of the domain, perpendicularly to the β -strands, with a fixed polarity in nucleic acid binding running 5' to 3' from β 4 and β 5 towards β 2 (Theobald et al., 2003). Interestingly, the OB fold binding face is conserved in VapX and located on the flat side of the VapXD complex. However, due to the swapped OB fold topology, the L45 loop is structurally replaced by the loop between the extra N-terminal β -strand (β 1') and β 1 (L1'1), while L12 is in the canonical location (**Figure 3A**). Consistently, a sequence comparison does not show any pattern of sequence similarity between VapX and the canonical OB fold proteins (**Figure 3C**). In summary, we conclude that VapX is a highly unusual antitoxin in that it neither contains a well-known DNA binding domain nor an intrinsically disordered region used for toxin inhibition.

Binding of VapX introduces asymmetry in the VapD dimer

Probably the most unusual feature of the VapXD TA system is that a single antitoxin (VapX) binds between two monomers of a toxin (VapD) homodimer. This interaction is largely driven

by charged and polar interactions with VapX being mostly negatively charged and VapD being positively charged (**Figure S3C**). Moreover, the 2:1 molecular interaction between toxin and antitoxin proteins allows us to understand how an asymmetrical domain interacts with a symmetrical dimer. Interestingly, the β -strands and loops of the VapX OB fold take a pseudo two-fold symmetrical conformation, which allows them to bind in similar ways into the two VapD subunits. In this interaction, the two loops connecting $\beta 1'$ and $\beta 1$ (L1'1) and that connecting $\beta 2$ and $\beta 3$ (L23) in VapX form two legs that are able to interact with each of the two VapD monomers (**Figure 4**). The two loops interact with some of the same residues (Arg39, Tyr25, Tyr45, and Gln22) in the two VapD monomers, but the details of the interactions differ to accommodate the lack of symmetry in VapX. Also, there is no sign up an internal pseudo-symmetry such as a palindromic protein sequence in VapX (**Figure 3C**). In VapD chain A, the L1'1 loop of VapX thus inserts into a small cleft between $\alpha 1$ and the conserved loop that connects $\beta 2$ and $\beta 3$ of the toxin. Here, Arg39 of VapD forms several hydrogen bonds to the backbone carbonyl groups of Phe18 and Gly19 in VapX (**Figure 4A**). In the other VapD monomer (chain B), the L23 loop of VapX instead binds on top of the conserved $\beta 2$ - $\beta 3$ loop of VapD, further away from $\alpha 1$ (**Figure 4B**). Here, both Arg39 and Ser43 form hydrogen bonds to the carbonyl group of VapX Gln44. Moreover, there is an interaction from VapD Tyr25, via VapX the Gln44 side chain, to VapD Asp7, which could directly affect activity.

Whereas VapX interacts differently with the conserved $\beta 2$ - $\beta 3$ loops of the two VapD molecules, its interactions with the N-terminal ends of $\alpha 1$ mirror each other and only induce slight differences in the conformation of the two VapD monomers in this region (**Figure 4A and 4B**). In one VapD monomer, the side chain of Gln22 forms a hydrogen bond with Asp59 of VapX, while in the other monomer it interacts with Thr41. The nearby hydrophobic residues, Val21 and Tyr25 are buried by hydrophobic residues of VapX for both monomers of VapD, further

stabilising the interaction. Finally, there is interaction between the dedicated $\alpha 1'$ helix of VapD chain A and VapX, which is tilted and comes closer to VapD at the top of the trimer. Several hydrogen bonds bind the helix to $\beta 3$ of VapX, mainly using backbone atoms as well as the side chains of Arg73 and Gln14 on VapD (**Figure 4C**).

In summary, the analysis of the VapD-VapX interaction demonstrates an interesting example of how a symmetrical homodimer adapts to binding of a folded, asymmetrical partner protein, in this case primarily facilitated via the $\beta 2$ - $\beta 3$ loop in VapD. The structure also sheds new light on the diversity of the classic OB fold as a protein-protein interaction domain. The two different modes of interaction induce different conformations in the conserved $\beta 2$ - $\beta 3$ loop of the two VapD monomers, which can be seen upon superposition of the two VapD monomers, including both orientations of VapX (**Figure 4D**). Interestingly, this loop was also found in two different conformations and with increased B-factors compared to the rest of the model in the *H. pylori* VapD dimer (Kwon et al., 2012). In both cases, however, and in line with what is observed for most type II toxin-antitoxin complexes, the putative active site region around Asp7 of VapD is sterically blocked by interaction of the antitoxin.

Isolated VapD has a closed active site cleft

In order to investigate the conformation of the activated VapD toxin as well as any changes induced by VapX binding, we determined crystal structures of isolated VapD as well as a D7N active site mutant. Surprisingly, expression of the *wt* toxin did not affect cell growth despite earlier results showing that VapD expression in *E. coli* causes growth arrest, suggesting that the effects of the toxin in this heterologous host is strain dependent or too subtle to observe during large-scale growth (Daines et al., 2004). Isolated VapD elutes with an apparent MW of 50 kDa during gel filtration chromatography indicating that it also forms a higher-order structure,

possibly a tetramer, in the absence of VapX (data not shown). The wt protein was crystallised in the hexagonal space group, P₆₃22 and the structure was refined to 2.5 Å, while the D7N mutant crystals could be best modelled in P321 and diffracted to 2.2 Å and (See STAR Methods and **Table 1**). In both cases, a VapD monomer was used to determine the structure by molecular replacement. The structures were refined to R=22.0% / R_{free}=26.4% (wt) and R=21.4% / R_{free}=26.7% (D7N) and are complete except for a few residues in the C-terminus and in the case of the D7N mutant, the extended region 10-23. Symmetry expansion of the structures reveals that the VapD dimer as well as the ring-shaped trimer (of dimers), are present in both cases in the absence of VapX (**Figure S3D** and **Figure S4A**). Moreover, since the structure of the isolated *wt* VapD toxin only contains a single molecule in the crystallographic asymmetric unit, it is perfectly two-fold symmetrical and thus that VapD forms a completely symmetrical homodimer (or trimer of dimers) in the absence of VapX. The D7N mutant structure contains a pseudo-symmetrical dimer in the asymmetric unit, with very few differences between the monomers.

Structural alignment of isolated VapD with the VapX-bound form reveals differences at the top of the cleft between VapD monomers (Leu8-Thr26 of the $\beta 1\alpha 1'\alpha 1$ motif), where the long $\alpha 1$ -helix has straightened and become extended by one turn while the shorter $\alpha 1'$ has unfolded completely into a loop that points upwards and towards the other VapD monomer (**Figure 5A**). The overall effect of this conformational change is that the binding cleft between the monomers of the dimer is closed by contacts between the loops on either side, with Tyr16 and Pro18 from one monomer coming into close contact with the corresponding residues from the other monomer (**Figure 5B**). This hydrophobic interaction is stabilised by a hydrogen bond from His17 to Asp12 of the same monomer on each side. Furthermore, the $\beta 2$ - $\beta 3$ loop that forms the bottom of the cleft, takes a slightly different turn compared to the VapX-bound form, with the

side chains pointing more inward packing more closely together. This also affects the putative catalytic residue, Asp7, which points more down into the cleft in the isolated structure (**Figure 5A**). However, the distance between these residues is even longer than in the VapX-bound structure (14.5 Å), supporting that they do not need to come close during catalysis and perhaps are part of separate ion binding sites. Interestingly, the isolated VapD dimer has a highly electropositive channel running through the closed cleft, which is compatible with binding of negatively charged nucleic acids (**Figure S4B**). This potential is generated by the combined positive charged from several basic arginine residues, Arg39, Arg73, and Arg76 likely in combination with backbone amino groups. Together, this observation is thus compatible with a potential for nucleic acid binding at the cleft region. This area is also the most conserved part of the structure, again supporting that it constitutes an important site (**Figure S4C**).

In the VapD D7N mutant, the loop connecting $\beta 1$ and the first α -helix ($\alpha 1$, Val10-Gln23) is not visible in the electron density. Since the missing part exactly covers the region that differs between the VapX-bound and isolated form of VapD, the structure is therefore compatible with both the open and closed conformation (**Figure 5C**). This suggests that this region of the structure is highly flexible in absence of VapX and thus might adjust to accommodate substrate during binding. Interestingly, the unit cell dimensions and optimal space group differ between the wt and D7N mutant structures, suggesting a slightly different packing (**Figure S4B**). The sidechain of the mutated residue (D7N) points in the same direction as the wt but slightly more into the cleft than the corresponding residue of wt VapD, now forming a hydrogen bond to the carbonyl atom of Thr40 in the other VapD monomer, which is absent in the wt structure due to the more retracted position of Asp7 (**Figure 5C**, inset).

DISCUSSION

In this paper, we present the structure of a toxin-antitoxin complex of the VapXD type, revealing a highly unusual stoichiometry of interaction, in which one antitoxin protein binds and inhibits a toxin dimer. Binding of the antitoxin directly to the active site of the toxin is a common feature of type II TA complexes and by far the most common molecular stoichiometry is 1:1 where each toxin molecule is inhibited by a cognate antitoxin (Page and Peti, 2016). There are examples of 2:1 toxin:antitoxin interactions among the VapBC TA modules, where extended C-terminal tails of the VapB antitoxins adopt a linear structure and use pseudo-palindromic protein sequences to inhibit both active sites of a nearby VapC toxin dimer (Bendtsen et al., 2017). However, in these cases, there is always a second VapB antitoxin present, thus preserving the overall 1:1 molecular stoichiometry of the complex, even though one antitoxin is not used to inhibit a toxin directly. The interaction observed between VapD and VapX is more similar to the gyrase poison toxin-antitoxin system, CcdBA, where an extended structure of the antitoxin CcdA interacts with both molecules of a symmetrical CcdB toxin dimer (De Jonge et al., 2009). However, common to nearly all TA interactions is that the toxin-interacting part of the antitoxin forms an extended (and intrinsically disordered) structure. VapXD is thus unique in that a single, folded domain antitoxin is able to bind both molecules of a toxin dimer. The compact structure of VapX thus raises the question of how toxin activation takes place in the VapXD system. For many TA systems, a stress-induced cellular protease like Lon or Clp is believed to be involved in targeted degradation of the antitoxin via the disordered region (Muthuramalingam et al., 2016). Given that VapX consists of a compact single-domain OB fold, this suggests that perhaps another, unknown, mechanism is involved in activation.

Mapping of the electrostatic surface potential reveals that VapX has an overall electronegative surface that binds to a positively charged cleft of VapD. The OB fold of the VapX antitoxin is a well-known nucleic acid-binding fold, however, the overall charge of VapX suggests that this

might not be the case. It thus appears that the VapX OB fold evolved specifically to bind a protein instead of nucleic acids (Arcus, 2002). OB domains involved in protein-protein interaction are seen found in some bacterial enterotoxins, such as the superantigen SEC-B from *Staphylococcus aureus*, which uses an OB fold to bind the T-cell receptor: major histocompatibility complex (TCR-MHC) on the surface of immune cells (Rodstrom et al., 2014). Despite an overall negative surface charge, the binding region of VapX does contain several residues that potentially could be used for nucleic acid interaction, including acidic (Asp33, Asp35, Glu 47, and Asp49), basic (Arg52), and aromatic (Phe18 and Tyr55) residues. Several of these residues are conserved and especially those in the loop between $\beta 1$ and $\beta 2$ (L12) appear not to interact with VapD and could thus have other functions. Future studies will hopefully reveal if VapX has a separate role that involves DNA/RNA binding aside from its interaction with VapD. If VapX does not bind DNA, this is contrary to what is observed for most type II TA modules, where the antitoxin functions as a repressor of transcription from the TA locus. This also raises the question of how the transcription of the *vapXD* locus is controlled; it is possible that it is regulated by a global trans-activator protein (such as Fis), as has been shown for the *vapBC-1* module (Cline et al., 2012). However, this would also have to be confirmed experimentally.

VapD is structurally similar to Cas2, which we know is active as a nuclease. Some RNase activity of VapD has been shown *in vitro*, but we do not know the specificity of the reaction (Ren et al., 2012). We know that a conserved residue (Asp 7 in NTHi VapD) is important, but its role in catalysis remains enigmatic. Moreover, reports are conflicting when it comes to the requirement for divalent cations such as Mg^{2+} and Mn^{2+} , which are required for the function of many nucleases. The consensus is that while Cas2 requires the presence of such ions (Beloglazova et al., 2008; Ka et al., 2017; Ka et al., 2014; Nam et al., 2012), VapD-type proteins

do not (Kwon et al., 2012; Pedersen et al., 2003; Zhang et al., 2003), which is surprising given their obvious structural homology. For Cas2, it has been suggested that there is hinge motion involving the loop region between $\beta 4$ and $\beta 5$, which would reduce distance between the aspartates to become compatible with Mg^{2+} binding (Ka et al., 2017; Ka et al., 2014; Nam et al., 2012). However, these conclusions are based on separate experiments, and careful parallel experiments using identical conditions would be needed to confirm if this is really the case. Finally, the concept of a toxin dimer acting as a nuclease involves the question of whether there are one or two active sites, a question which is also pertinent for understanding the VapC toxins and still not resolved. To better understand the catalytic mechanism of VapD, a structure with RNA substrate bound is needed. Such a structure would be highly valuable and would also give insights into Cas2 RNase activity, as there currently is no structure of Cas2 with RNA substrate bound.

The VapXD and VapD structures also shed light on the (yet) speculative evolutionary connections between CRISPR-Cas and the TA systems (Makarova et al., 2012). According to the proposed model, a primordial VapXD-type TA system became associated with an ancestral transposable DNA element containing *cas* genes (a "casposon"), which was later incorporated next to a *cas* cassette to allow for adaptive immunity via incorporation of novel genetic elements in the nearby CRISPR arrays (Koonin and Krupovic, 2015). According to this hypothesis, the VapD fold and nuclease evolved to become the Cas2 protein, which is known today for its role in the Cas1-Cas2 universal CRISPR-Cas adaptation module (Koonin and Krupovic, 2015). Curiously, Cas2 appears to have kept its nuclease activity despite no apparent need for this during cleavage of protospacers during CRISPR-Cas adaptation (Nunez et al., 2014; Nuñez et al., 2015). Studying the VapD-type nucleases thus provides a unique glimpse into a world that predates CRISPR-Cas and can perhaps offer explanations to why this is the case. Comparison

of VapD and Cas2, however, offers some insight into how the transformation from a TA toxin to a CRISPR Cas protein took place. We know that Cas2 lacks the small $\alpha 1'$ helix, which sits on top of the VapD dimer and is involved in recognition of VapX. This functionality is not required in Cas2, which has likely meant that the helix was lost over time. Moreover, Cas2 contains an additional β -hairpin at its C-terminus, which we know are involved in CRISPR-specific functions (Cas1 binding) (Nunez et al., 2014). This element is not present in VapD and could thus represent a later addition to the protein as it evolved to become part of the Cas1-Cas2 complex. In the light of this, it is intriguing that natural variations of VapD have been found in some strains of NTHi that lack residues 30-77 (Daines et al., 2004). This would remove a large part of the core Cas2 fold and connect the long $\alpha 1$ with $\beta 4$, thus making it difficult, from a structural perspective, to understand how the protein can remain folded (Supplementary Figure 4D). Nevertheless, all natural deletions are all in-frame and the resulting proteins have been shown to interact with full length VapD *in vivo*. It is thus possible that this represents an extra twist on the VapXD toxin-antitoxin system and its relevance in the pathogenicity of *H. influenzae*.

ACKNOWLEDGEMENTS

This work was funded by the Danish National Research Foundation's Centre for Bacterial Stress response and Persistence (BASP, DNRF 120), the Novo Nordisk Foundation, and the U.S. National Institute on Deafness and Other Communication Disorders grant DC014756. The authors are indebted to beamline personnel at P13, EMBL-Hamburg, Germany and ID29 at ESRF, France, for help and advice during data collection and Lan Bich Van for invaluable help with protein purification and crystallisation.

AUTHOR CONTRIBUTIONS

Conceptualization, D. A. D. and D. E. B.; Resources, D. A. D. and D. E. B.; Investigation, A. L. M., M. C. V., M. B. B., and M. S.; Formal Analysis, M. B. B., M. S., F. B., M. H. N., and D. E. B.; Data Curation, M. S., M. H. N., and D. E. B.; Writing - Original Draft, M. S., M. B. B., and M. H. N.; Writing - Reviewing & Editing, M. B. B., M. H. N., D. A. D., and D. E. B.; Funding Acquisition, D. A. D. and D. E. B.; Supervision, M. S., D. A. D., and D. E. B.

DECLARATIONS OF INTERESTS

The authors declare no competing interests.

MAIN FIGURE TITLES AND LEGENDS

Figure 1. The structure of *H. influenzae* VapXD. **A.** The genetic organisation of the *H. influenzae* *vapXD* operon. The *vapX* gene (192 nt/63 aa) precedes the *vapD* gene (276 nt/92 aa) spaced by 8 nt. **B.** Overall structure of the VapXD₂ heterotrimer composed of a toxin of two VapD molecules (sea green/light blue) and one VapX antitoxin molecule (red). **C.** Symmetry expansion of the crystal structure showing the ring-shaped higher-order structure composed of three VapXD₂ heterotrimers. The inset shows the conserved residues, Arg79 and Glu81, that stabilise the assembly of the ring structure through electrostatic interactions. **D.** Structure-based sequence alignment of four VapD homologues and three Cas2 homologues, *H. influenzae* VapD (PDB:6ZN8, UniProt P71351, this work), *H. pylori* VapD (HP0315, PDB:3UI3, UniProt 19909) (Kwon et al., 2012), *Neisseria gonorrhoeae* VapD (UniProt Q51003), *Moraxella nonliquefaciens* VapD (UniProt A0A1B8QQI3), *Xanthomonas albilineans* Cas2 (PDB:5H1O, UniProt D2UG58) (Ka et al., 2017), *Streptococcus pyogenes* Cas2 (PDB:4QR0, UniProt Q99YS8) (Ka et al., 2014), and *S. solfataricus* Cas2 (PDB:2I8E, UniProt Q97YC2) (Beloglazova et al., 2008). Strictly conserved residues are shown on a blue background, and similar residues are in blue letters and boxed. Secondary structure elements as observed for the

NTHi VapD structure are indicated on top of the alignment and labelled as in **Figure 2A**. The conserved Asp7 involved in catalysis is marked with an orange star. Arg39, Ser43, and Tyr45 located near the active site are marked with blue spheres, hydrophobic residues Leu44, Phe78, and Ile80 involved in dimerisation are marked with green crosses, and the charged residues Arg79 and Glu81 involved in dimerisation and formation of the ring shape are marked with red crosses. The sequences were aligned using PROMALS 3D (Pei and Grishin, 2014) and printed using ESPript 3.0 (Robert and Gouet, 2014). See also Figures S1, S2 and Table S1.

Figure 2. The VapD toxin is similar to Cas2 and GhoS. A. The NTHi VapD monomer with secondary structure elements coloured individually and labelled. **B.** Dimerisation interface of the VapD homodimer with monomers in different colours and backbone of residues 75-85 shown as sticks to illustrate the hydrogen bonds stabilising the dimer. **C.** Structural alignment of *H. pylori* VapD (PDB 3UI3, red, r.m.s.d. C α atoms 0.98 Å) and *E. coli* Cas2 (PDB 5DS4, green, r.m.s.d. C α atoms 3.9 Å) on the NTHi VapD dimer (blue) (Kwon et al., 2012; Nuñez et al., 2015). In both cases, the alignment was based on the complete dimers and using the VapD dimer from the VapXD complex. Differences are shown with bright green (Cas2) and dark blue (VapD), respectively. The conserved acidic residues (Asp7 in NTHi VapD, Asp65 in *H. pylori* VapD, and Glu9 in Cas2) are shown with sticks. **D.** Left, structural alignment of the type V antitoxin, GhoS, from the *E. coli* GhoST TA system (orange and red, PDB 2LLZ) (Wang et al., 2012) with an NTHi VapD monomer (blue) with structural differences shown in red (GhoS) and dark blue (VapD); Right, close-up of the active site region with relevant amino acids marked. See also Figure S3.

Figure 3. VapX has an OB fold. A. The structure of NTHi VapX (left) and topology (right). The coloured parts of the diagrams represent the canonical elements of the OB fold. **B.** Structure

of the *Acinetobacter baumannii* EF-P OB fold with conserved secondary structure in gold (PDB 5J3B, left, J. Abendroth *et al.*, *unpublished*) and corresponding topology (right). EF-P has a classical Greek key topology consisting of a three-stranded, anti-parallel sheet followed by a helix and $\beta 4$ hydrogen bonding to $\beta 1$. The conserved binding face of the OB fold is indicated in grey. **C.** Sequence alignment of NTHi VapX with homologues from other bacteria as well as *A. baumannii* EF-P (bottom). The alignment shows the secondary structure on top of the sequences, with the sequence of EF-P wrapped around to show the structural homology between $\beta 1'$ in VapX and $\beta 4$ in EF-P. Conserved residues are shown in red and the conserved glycine residues at the beginning of some of the β -strands in black boxes.

Figure 4. The VapD homodimer adapts to binding the asymmetrical VapX. **A.** Interactions between VapD (chain A) and L23 of VapX with relevant residues labelled. **B.** Interactions between VapD (chain B) and L1'1 of VapX with relevant residues labelled. **C.** Interactions between VapD helix $\alpha 1'$ (chain A) and VapX. These interactions are not found in chain B due to the asymmetry of VapX. **D.** Structural alignment of VapD chain B (in grey) onto VapD chain A (blue), including its interaction with VapX to show how the symmetrical dimer adapts to binding of an asymmetrical molecule.

Figure 5. The structure of the isolated wt and D7N VapD toxin. **A.** Overview of the structure of the isolated VapD dimer (green) and comparison to the VapX-bound conformation (grey). The conserved Asp7 is shown and the arrows indicate the closure of the lobes surrounding the active site cleft upon loss of the VapX antitoxin. **B.** Details of the interactions at the top of the isolated VapD dimer where the lobes meet. The hydrophobic interaction area involving Tyr16 and Pro18 is shown in grey. **C.** Structural alignment of the isolated VapD D7N structure (in two shades of blue) on isolated wt VapD (grey). The inset shows the conformation of the

mutated residue (D7) as well as nearby residues in the β 2- β 3 loop (in blue, D7N only). In VapD D7N, there is a hydrogen bond between Asn7 and the carbonyl oxygen of Thr41 of the opposing VapD monomer (dashed line). See also Figure S4.

MAIN TABLES AND LEGENDS

PDB entry	VapXD (Se-Met) 6ZN8	VapD (wt) 6ZI0	VapD (D7N) 6ZI1
Data Collection			
Wavelength (Å)	0.979	0.98021	0.98021
Resolution range	46.88 - 3.21 (3.33 - 3.21)*	85.88 - 2.50 (2.56 - 2.50)	42.19 - 2.20 (2.26 - 2.20)
Space group	<i>I4₁32</i>	<i>P6₃22</i>	<i>P321</i>
Unit cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	256.75, 256.75, 256.75	99.168, 99.168, 44.486	84.380, 84.380, 49.476
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90, 90, 90	90, 90, 120	90, 90, 120
Total no. of reflections	172,895 (17,802)	49,715 (3673)	34,356 (2279)
Unique reflections	23,771 (2337)	4800 (344)	10,306 (736)
Multiplicity	7.3 (7.6)	10.3 (10.7)	3.3 (3.1)
Completeness (%)	99.7 (99.7)	99.9 (100.0)	97.2 (97.9)
R _{merge}	0.169 (1.37)	0.212 (1.581)	0.060 (0.928)
I/σ(I)	9.2 (1.3)	9.5 (1.3)	12.1 (1.3)
CC _{1/2}	0.99 (0.53)	0.996 (0.55)	0.999 (0.61)
Refinement			
Average B-factor (Å ²)	100.7	76.2	58.9
No. of reflections	23,774 (2337)	4560 (326)	9282 (668)
No. of reflections (free)	2351 (209)	241 (18)	1028 (68)
R-work (%)	25.8 (34.9)	22.0 (28.4)	21.4 (38.6)
R-free (%)	29.9 (39.6)	26.4 (28.9)	26.7 (40.9)
No. of atoms			
protein	3939	722	1333
solvent	0	55	88
ligand	0	0	0
RMSD bonds (Å)	0.010	0.008	0.007
RMSD angles (°)	1.11	1.74	1.48
Ramachandran statistics			
favoured (%)	93.7	91.6	84.8
allowed (%)	6.3	7.2	14.6
generously allowed (%)	-	1.2	0.7
outliers (%)	0.0	0.0	0.0

Table 1. Crystallographic data collection and refinement statistics. *Numbers in parentheses indicate values for the outermost resolution shell.

STAR Methods

RESOURCE AVAILABILITY

Lead Contact

Lead contact: Ditlev Egeskov Brodersen (deb@mbg.au.dk).

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact.

Materials Availability

There is no restriction on materials generated for this study and first reported here. They are available from the lead author upon reasonable request.

Data and Code Availability

Atomic coordinates and structure factors have been deposited to the protein database (PDB:6ZN8, VapXD; 6ZI0, VapD(wt); 6ZI1, VapD(D7N)). All other data are available from the lead author upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbes

E. coli B834 DE3 cells were cultured at 20°C to 37°C in Se-Met Complete media. *E. coli* BL21 cells were cultured at 20°C to 37°C in LB media.

METHOD DETAILS

Protein purification. The plasmid pDD791 expressing the *wt* VapXD complex from NTHi strain 86-028NP C-terminally His-tagged on VapD, was transformed into the Met auxotroph *E. coli* strain B834 (DE3) (Merck) for seleno-methionine (Se-Met) incorporation using Se-Met

Complete medium (Molecular Dimensions). Cells were grown at 37°C to OD₆₀₀ = 0.6-0.8, followed by induction with 0.1 mM IPTG overnight at 20°C. Cell pellets were resuspended in a lysis buffer (50 mM Tris-HCl, pH=8, 300 mM NaCl) followed by sonication and clarification by centrifugation at 14,000 rpm for 45 min. Cell extract containing the VapX:VapD-His₆ complex was loaded onto a 1 ml His-Trap column (GE Healthcare). After binding, a multi-step wash/elution was performed using a buffer identical to the lysis buffer but including 10, 20, 40, 80, 150, and 300 mM imidazole, respectively. Fractions containing most protein complex were further purified on a 1 ml Source 15Q column (GE Healthcare) running in 50 mM Tris-HCl, pH=8.5, 5 mM β-mercaptoethanol (BME) and eluted with a gradient into 1 M NaCl. Final separation was achieved on a Superdex 200 10/300 GL column (GE Healthcare), equilibrated in 20 mM Tris-HCl, pH 7.5, 300 mM NaCl, and 5 mM BME. Peak fractions were concentrated to 12 mg/ml prior to crystallisation.

For expression of isolated VapD, the *wt vapD* gene from NTHi strain 86-028NP was inserted into pET24b encoding a C-terminal Leu-Glu-His₆ affinity tag. *E. coli* BL21 (DE3) transformed with this plasmid were grown to OD₆₀₀ = 0.6-0.8 in Luria-Bertani (LB) medium including kanamycin (0.05 mg/mL) at 37°C in a shaking incubator. Expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and growth continued overnight at 20°C before harvesting cells at 6000 rpm. Cells were resuspended in 50 mM Tris-HCl, pH=8.0, 300 mM NaCl supplemented with 0.2 mg/mL lysozyme, incubated 30 minutes on ice and were lysed by sonication. After removal of cell debris, the supernatant was applied to a pre-packed 1 mL HisTrap FF (GE Healthcare) column, washed in 20-30 column volumes (CV) of lysis buffer + 10 mM imidazole before elution using lysis buffer + 300 mM imidazole. Peak fractions were further purified on a 1 mL Source 15Q column (GE Healthcare Life Sciences) equilibrated in 50 mM Tris-HCl, pH=8.5, 50 mM NaCl, 5 mM BME by gradient salt elution up to 700 mM

NaCl. Finally, the pooled and concentrated fractions were applied to a Superdex200 increase 10/300 column (GE Healthcare) pre-equilibrated in 20 mM Tris-HCl, pH=8.0, 150 mM NaCl, 5 mM BME running at 0.4 mL/min. The peak fractions were analysed on a 4-20 % SDS-PAGE and pooled for crystallisation trials and biochemical experiments. The column was calibrated with standard proteins of known mass. The VapD D7N site-directed mutant was synthesised by Integrated DNA Technologies (Coralville, IA, USA) and inserted into pET24b encoding a C-terminal Leu-Glu-His₆ affinity tag (pDD943) and expressed and purified in a similar way, except pH=7.5 was used in the final gel filtration buffer. Both protein samples were concentrated to 6-7 mg/mL using a Vivaspin 500 mini spin column (Sartorius) before crystallisation.

Structure determination. The NTHi VapXD complex was crystallised using the sitting-drop vapor-diffusion method at 19°C. 96-well crystallisation plates (Hampton research) were dispensed using a Mosquito automated crystallisation robot (TTP labtech). Protein crystals belonging to the cubic space group I4132 appeared in 1.6 M ammonium sulphate, 0.1 M MES pH=6.5, and 10 % v/v 1,4-dioxane. Individual crystals were transferred to a drop of the corresponding mother liquor supplemented with 25% (v/v) glycerol and immediately flash-cooled in liquid nitrogen. A complete Se-Met SAD data set was collected at the ID29 beamline of ESRF and phases obtained using SHELXD/hkl2map pipeline (de Sanctis et al., 2012; Sheldrick, 2010). The structure was built manually in Coot (Emsley et al., 2010) based on the initial model output by hkl2map and iteratively refined using Phenix.refine (Adams et al., 2010). R factors were significantly improved by optimising bulk solvent mask parameters. Ramachandran outliers and clashes were corrected by hand in Coot, and Ramachandran restraints were used in the final rounds of refinement to avoid some Ramachandran outliers from reappearing.

For isolated VapD (wt), crystals were obtained in 1 M LiCl, 0.1 M citrate, pH=4.0, 20% PEG 6000 while crystals of the VapD (D7N) mutant appeared in 1.8 M ammonium sulphate, 50 mM Na cacodylate, pH=6.0. In both cases, the crystals grew at 19°C and were frozen in mother liquor supplemented with 20% PEG 400 before data collection (360 degrees at 0.1° per frame) at 100 K at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany) (Cianci et al., 2017). Data were processed in XDS through XDSGUI (Kabsch, 2010) and the structures were determined by molecular replacement in Phenix.phaser (McCoy et al., 2007) using a monomer of VapD from the VapXD as a search model. Initial models were built using Phenix.autobuild (Terwilliger et al., 2008) and the structures were iteratively refined in Phenix.refine (Adams et al., 2010) with manual rebuilding in Coot (Emsley et al., 2010). Both crystal forms suffer from various crystallographic irregularities and pseudo-symmetries likely due to the higher order symmetry of the VapD dimer and ring structure. In both cases, automated software initially suggested the high-symmetry space group P6₃22, which was also used for the VapD (wt) structure. For the higher resolution structure of VapD (D7N), the map in this space group revealed overlaps between symmetry mates and the structure was therefore expanded to P321 to accommodate slightly different structures of some of the side chains of the two monomers in the crystallographic asymmetric unit. This structure thus has a non-crystallographic symmetry (NCS) two-fold axis parallel to the two-fold axis in P6₃22 but is modelled as an NCS.

Analytical gel filtration and cross-linking. Peak fractions from a Source 15Q column were pooled and concentrated to app. 6 mg/mL using a Vivaspin 6 spin filter (Sartorius) with a 5 kDa molecular mass cut-off. During concentration, the buffer was exchanged to 20 mM Hepes, pH=7.5, 150 mM NaCl, 5 mM BME. 400 µL of the protein sample was loaded on a Superdex

200 increase 10/300 GL column (GE Healthcare), which had been pre-equilibrated in the same buffer. The protein was eluted at 0.4 mL/min while measuring OD280. The elution volume (V_e) from the Superdex 200 increase 10/300 GL column was used to calculate the apparent molecular mass of the sample based on a calibration run with the Gel Filtration Calibration Kit HMW (GE Healthcare). Peak fractions from the analytical SEC run were concentrated to 4 mg/ml, and a reaction mixture containing 63 μ L of VapXD complex and 7 μ L of 0.5% or 1% glutaraldehyde was prepared. At the time points 5, 10, 20, 40, 60, and 120 min, 11 μ L aliquots were removed, and the reaction quenched with 4 μ L of 2 M Tris, pH=8.0. Samples were analysed on precast protein gels (Bio Rad).

Dynamic light scattering (DLS). The hydrodynamic diameter of the VapXD complex in solution was measured by DLA using a Zetasizer μ V (Malvern) cuvette setup. For this, the concentration of purified VapXD complex was first adjusted to 1.5 mg/mL, and DLS was measured at 5°C. After an equilibration time of 120 s, three consecutive measurements were made, each consisting of 13 runs.

Calculation of theoretical hydrodynamic diameter. The program HYDROPRO version 10.0 (Ortega et al., 2011) was used to predict the translational diffusion coefficient D , of the different suggested assemblies of VapXD. The viscosity of the solvent (water), was set to 0.0015 N s/m² and the temperature was set to 278 K, to mimic the conditions in the DLS measurements. The predicted translational diffusion coefficients (D) were then used to calculate the hydrodynamic diameters (D_{hyd}) by the Stokes-Einstein equation, $D_{hyd} = kT/3\pi\eta D$, where k is Boltzmann's constant, T is the absolute temperature, and η is the viscosity. The same values for temperature and viscosity as used in the HYDROPRO program were used for this calculation.

QUANTIFICATION AND STATISTICAL ANALYSIS

XDS/XDSGUI (Kabsch, 2010) were used for reduction of diffraction data, and the SHELXD/hkl2map pipeline for phase calculation for the structure of Se-Met VapXD (de Sanctis et al., 2012; Sheldrick, 2010). For VapD(wt) and VapD(D7N), Phenix.phaser (McCoy et al., 2007) we used for molecular replacement. Structure models were built using Phenix.autobuild (Terwilliger et al., 2008) and Coot (Emsley et al., 2010) and refined in Phenix.refine (Adams et al., 2010) with manual rebuilding in Coot (Emsley et al., 2010). All crystallographic data statistics are available in Table 1.

The program HYDROPRO version 10.0 (Ortega et al., 2011) was used to predict the translational diffusion coefficient D for VapXD.

ADDITIONAL RESOURCES

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* *66*, 213-221.
- Arcus, V. (2002). OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.* *12*, 794-801.
- Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* *44*, W344-350.
- Beloglazova, N., Brown, G., Zimmerman, M.D., Proudfoot, M., Makarova, K.S., Kudritska, M., Kochinyan, S., Wang, S., Chruszcz, M., Minor, W., *et al.* (2008). A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J. Biol. Chem.* *283*, 20361-20371.
- Bendtsen, K.L., Xu, K., Luckmann, M., Winther, K.S., Shah, S.A., Pedersen, C.N.S., and Brodersen, D.E. (2017). Toxin inhibition in *C. crescentus* VapBC1 is mediated by a flexible pseudo-palindromic protein motif and modulated by DNA binding. *Nucleic Acids Res.* *45*, 2875-2886.
- Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I., Roessle, M., Cipriani, F., Fiedler, S., and Schneider, T.R. (2017). P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J Synchrotron Radiat* *24*, 323-332.
- Cline, S.D., Saleem, S., and Daines, D.A. (2012). Regulation of the vapBC-1 toxin-antitoxin locus in nontypeable *Haemophilus influenzae*. *PLoS One* *7*, e32199.
- Daines, D.A., Jarisch, J., and Smith, A.L. (2004). Identification and characterization of a nontypeable *Haemophilus influenzae* putative toxin-antitoxin locus. *BMC Microbiol.* *4*, 30.
- Daines, D.A., Wu, M.H., and Yuan, S.Y. (2007). VapC-1 of nontypeable *Haemophilus influenzae* is a ribonuclease. *J. Bacteriol.* *189*, 5041-5048.

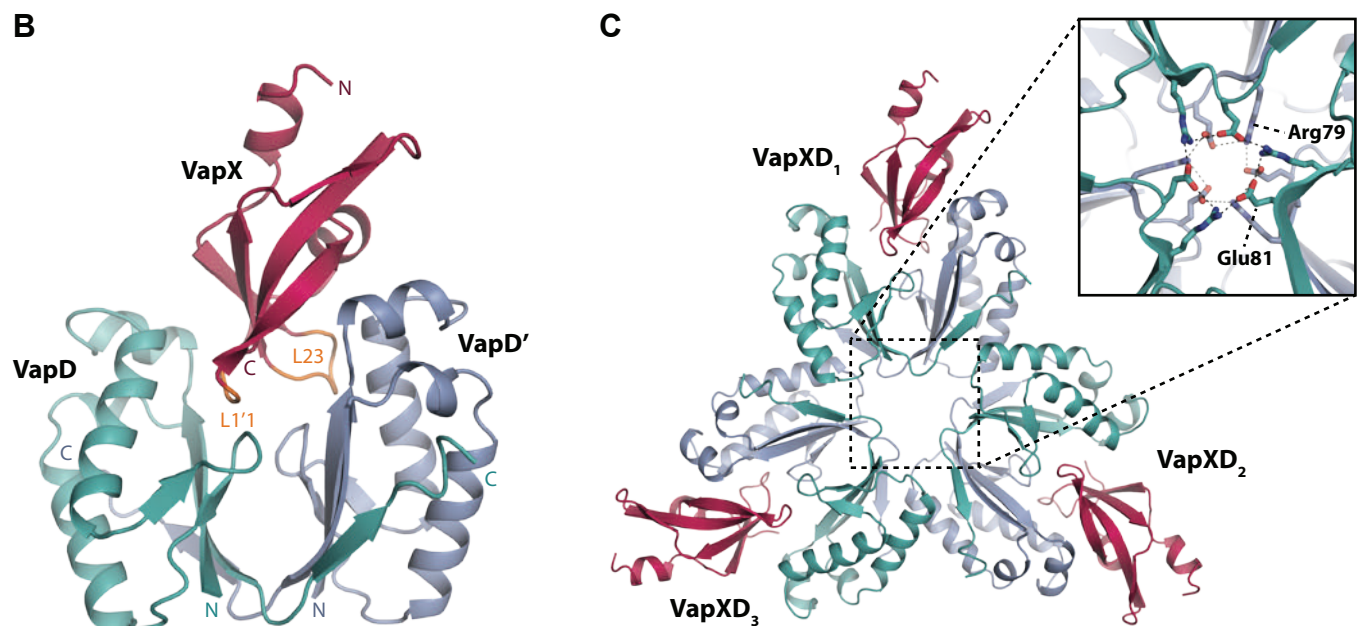
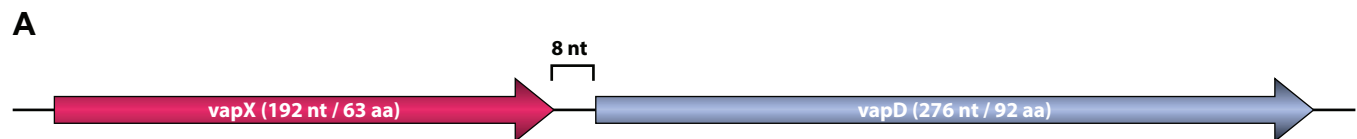
- De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., Wyns, L., De Greve, H., and Loris, R. (2009). Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Mol. Cell* *35*, 154-163.
- de Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., *et al.* (2012). ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *Journal of Synchrotron Radiation* *19*, 455-461.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* *66*, 486-501.
- Goeders, N., and Van Melderren, L. (2014). Toxin-antitoxin systems as multilevel interaction systems. *Toxins (Basel)* *6*, 304-324.
- Harms, A., Brodersen, D.E., Mitarai, N., and Gerdes, K. (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol. Cell* *70*, 768-784.
- Ka, D., Hong, S., Jeong, U., Jeong, M., Suh, N., Suh, J.Y., and Bae, E. (2017). Structural and dynamic insights into the role of conformational switching in the nuclease activity of the *Xanthomonas albilineans* Cas2 in CRISPR-mediated adaptive immunity. *Struct Dyn* *4*, 054701.
- Ka, D., Kim, D., Baek, G., and Bae, E. (2014). Structural and functional characterization of *Streptococcus pyogenes* Cas2 protein under different pH conditions. *Biochem. Biophys. Res. Commun.* *451*, 152-157.
- Kabsch, W. (2010). Xds. *Acta Crystallogr D Biol Crystallogr* *66*, 125-132.
- Koonin, E.V., and Krupovic, M. (2015). Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nat. Rev. Genet.* *16*, 184-192.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* *372*, 774-797.
- Kwon, A.R., Kim, J.H., Park, S.J., Lee, K.Y., Min, Y.H., Im, H., Lee, I., Lee, K.Y., and Lee, B.J. (2012). Structural and biochemical characterization of HP0315 from *Helicobacter pylori* as a VapD protein with an endoribonuclease activity. *Nucleic Acids Res.* *40*, 4216-4228.

- LeRoux, M., Culviner, P.H., Liu, Y.J., Littlehale, M.L., and Laub, M.T. (2020). Stress Can Induce Transcription of Toxin-Antitoxin Systems without Activating Toxin. *Mol. Cell*.
- Makarova, K.S., Anantharaman, V., Aravind, L., and Koonin, E.V. (2012). Live virus-free or die: coupling of antiviral immunity and programmed suicide or dormancy in prokaryotes. *Biology direct* 7, 40.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.
- Molinaro, A.L., Kashipathy, M.M., Lovell, S., Battaile, K.P., Coussens, N.P., Shen, M., and Daines, D.A. (2019). Crystal Structure of VapBC-1 from Nontypeable *Haemophilus influenzae* and the Effect of PIN Domain Mutations on Survival during Infection. *J. Bacteriol.* 201.
- Murphy, T.F. (2003). Respiratory infections caused by non-typeable *Haemophilus influenzae*. *Curr. Opin. Infect. Dis.* 16, 129-134.
- Murzin, A.G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* 12, 861-867.
- Muthuramalingam, M., White, J.C., and Bourne, C.R. (2016). Toxin-Antitoxin Modules Are Pliable Switches Activated by Multiple Protease Pathways. *Toxins (Basel)* 8.
- Nam, K.H., Ding, F., Haitjema, C., Huang, Q., DeLisa, M.P., and Ke, A. (2012). Double-stranded endonuclease activity in *Bacillus halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. *J. Biol. Chem.* 287, 35943-35952.
- Nunez, J.K., Harrington, L.B., Kranzusch, P.J., Engelman, A.N., and Doudna, J.A. (2015). Foreign DNA capture during CRISPR-Cas adaptive immunity. *Nature* 527, 535-538.
- Nunez, J.K., Kranzusch, P.J., Noeske, J., Wright, A.V., Davies, C.W., and Doudna, J.A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528-534.
- Nuñez, J.K., Lee, A.S.Y., Engelman, A., and Doudna, J.A. (2015). Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature* 519, 193-198.

- Ortega, A., Amoros, D., and Garcia de la Torre, J. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* *101*, 892-898.
- Page, R., and Peti, W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat. Chem. Biol.* *12*, 208-214.
- Pedersen, K., Zavialov, A.V., Pavlov, M.Y., Elf, J., Gerdes, K., and Ehrenberg, M. (2003). The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell* *112*, 131-140.
- Pei, J., and Grishin, N.V. (2014). PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol Biol* *1079*, 263-271.
- Ren, D., Walker, A.N., and Daines, D.A. (2012). Toxin-antitoxin loci vapBC-1 and vapXD contribute to survival and virulence in nontypeable *Haemophilus influenzae*. *BMC Microbiol.* *12*, 263.
- Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* *42*, W320-324.
- Rodstrom, K.E., Elbing, K., and Lindkvist-Petersson, K. (2014). Structure of the superantigen staphylococcal enterotoxin B in complex with TCR and peptide-MHC demonstrates absence of TCR-peptide contacts. *J. Immunol.* *193*, 1998-2004.
- Sheldrick, G.M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* *66*, 479-485.
- Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* *64*, 61-69.
- Theobald, D.L., Mitton-Fry, R.M., and Wuttke, D.S. (2003). Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.* *32*, 115-133.

Wang, X., Lord, D.M., Cheng, H.Y., Osbourne, D.O., Hong, S.H., Sanchez-Torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W., *et al.* (2012). A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nat. Chem. Biol.* 8, 855-861.

Zhang, Y., Zhang, J., Hoeflich, K.P., Ikura, M., Qing, G., and Inouye, M. (2003). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Mol. Cell* 12, 913-923.



D

Sequence alignment of VapD and Cas2 proteins from various species, highlighting conserved regions and mutations. The alignment is divided into two main sections, each with a diagram above showing domain structure (β1, α1', α1, β2, β3, α2, α2', β4, β5) and residue positions (1, 10, 20, 30, 40, 50, 60, 70, 80, 90).

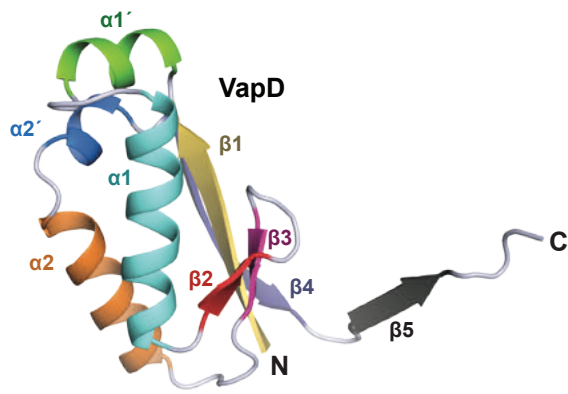
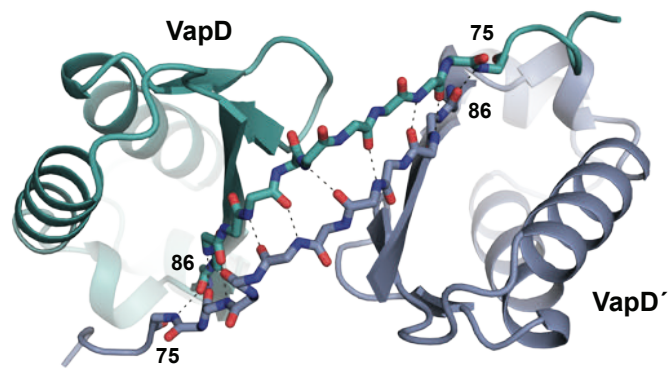
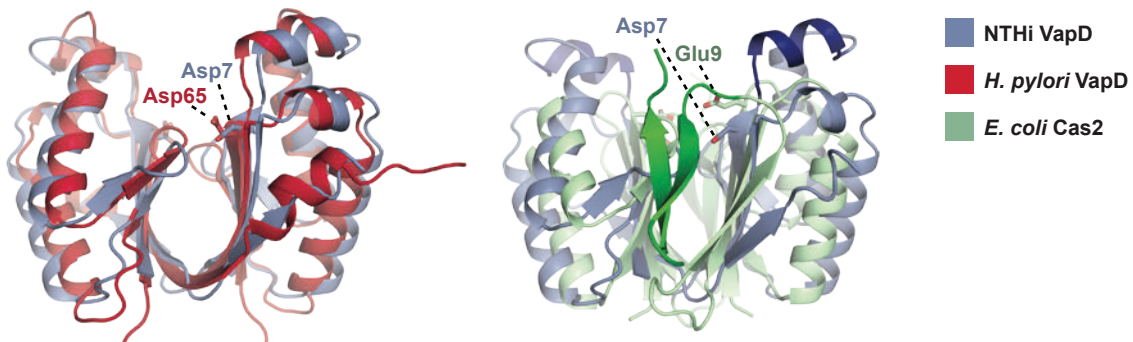
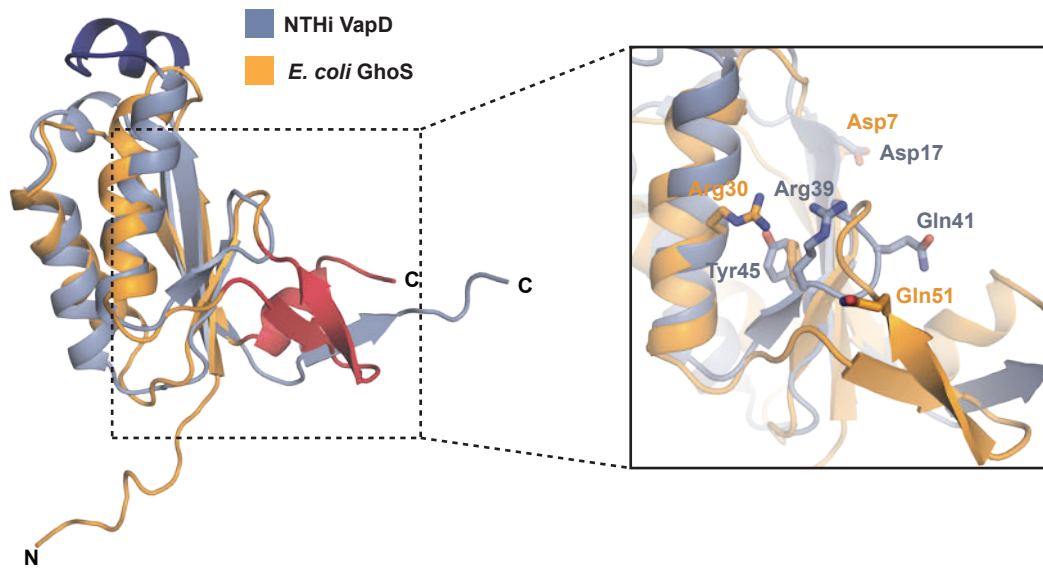
Top Section (Residues 1-50):

- VapD:**
 - H. influenzae* VapD: ..MYAI..AFDLVVKDTDYHPKGVQEAAYTDIGAVLA..KFGFV**RT**OG..SLYTNMNE..DM
 - H. pylori* VapD: ..MYAL..AFDLKIEILKKEYGEPYNKAYDDLRQEL..LLGFETWTOG..SVYVNYSKENTL
 - N. gonorrhoeae* VapD: ..MYAI..SFDLVVADTAQNHPKGISQAYADIGYTLR..QFGFT**RV**OG..GLYTCQNE..DM
 - M. nonliquefaciens* VapD: ..MYAI..GFDLLVKEVEKHHPKGVRQAYSDEIILN..KFSEFN**RI**OG..SLYVTE**NE**..NM
- Cas2:**
 - X. albilineans* Cas2: .MMVLV..SYDVSTSS.....PGGD**KRLRK**VAKACR..DLGQ**RV**OF..SVFEIEV...DP
 - S. pyogenes* Cas2: .MMVLV..TYDVNTET.....PAGR**KRLRH**VAKLCV..DYGQ**RV**ON..SVFECSV...TP
 - S. solfataricus* Cas2: MLYLIFFYDIT.....DD**NLRNR**VAEFLKKGLLD**RI**OYS**SVFM**GDL...NS

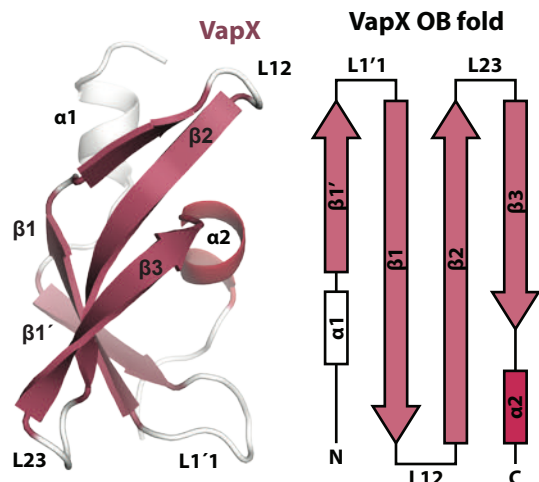
Bottom Section (Residues 53-90):

- VapD:**
 - H. influenzae* VapD: 53ANLFOAMN**ALKQL**..AWISQ....SVRD**IRAF**..RTEQ..WS..DFTDFI.....
 - H. pylori* VapD: 53AQVYKAIN**KLSQI**..EWFKK....SVRD**IRAF**..KVED..FS..DFTEIV.KSL...
 - N. gonorrhoeae* VapD: 53ANLFSAIN**ELKAL**..PWFPS....SVRD**IRAF**..RIEQ..WS..DFTS...LVKS..
 - M. nonliquefaciens* VapD: 53ELVTKAMV**SLSR**L..SWFSA....SVRD**IRAF**..KVEQ..WS..NFNS...FFKEEF
- Cas2:**
 - X. albilineans* Cas2: 46AOWTALRQ**RLCD**LI..DPD....IDS**LRFY**..HLGAKWEA..RVEH.VGAKP...
 - S. pyogenes* Cas2: 46AEFVDIKH**RLTQ**II..DEK....TDS**IRFY**..LLGKNWQR..RVETLG.....
 - S. solfataricus* Cas2: 47SRLKDVEAG**LKI**IGNRKKKLQEDE**RFF**ILIVVPITENQFR**REERIV**..IGYS.....

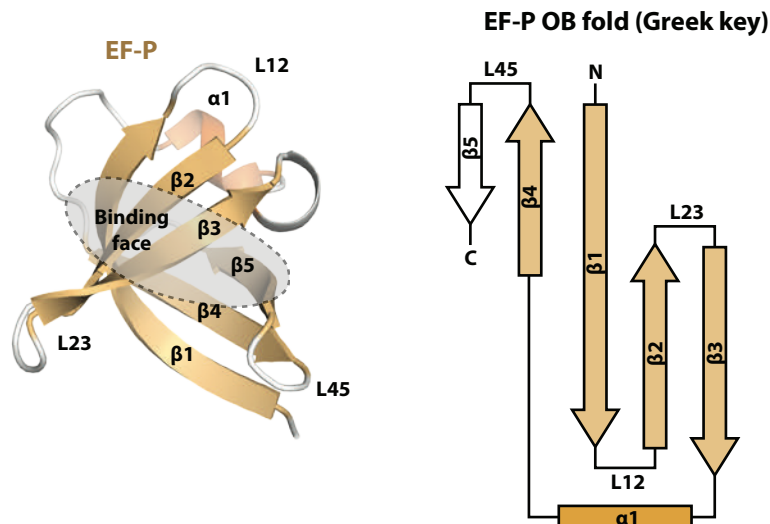
Conserved regions are indicated by blue boxes. Mutations are indicated by red boxes and red 'x' marks. The alignment shows high conservation between VapD and Cas2 proteins, particularly in the α1' and α1 regions.

A**B****C****D**

A

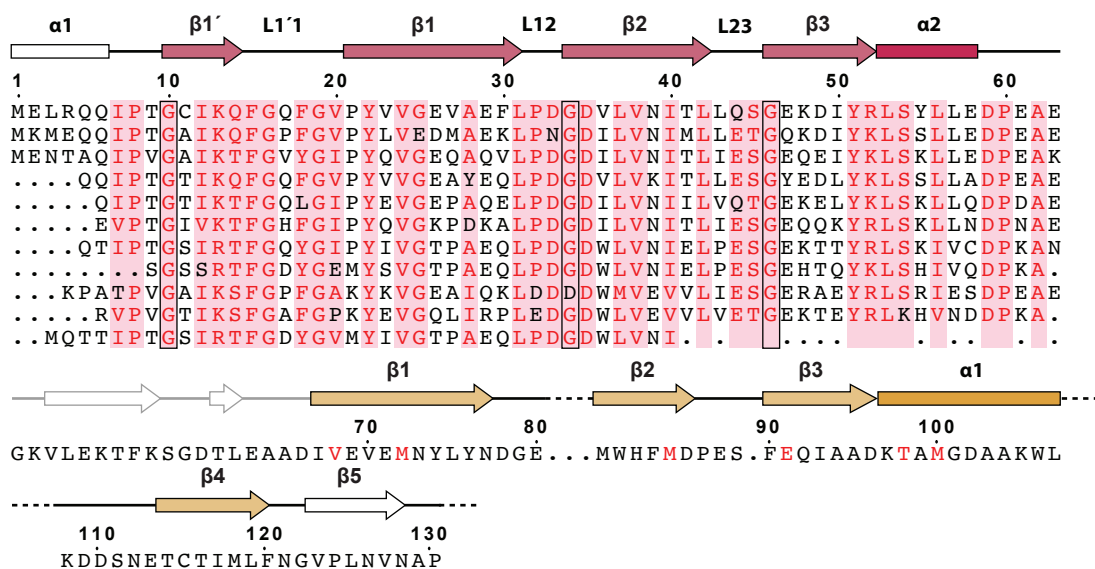


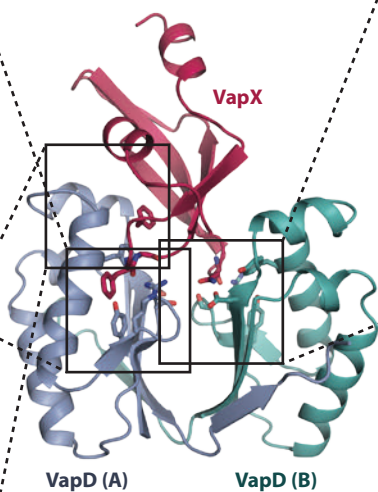
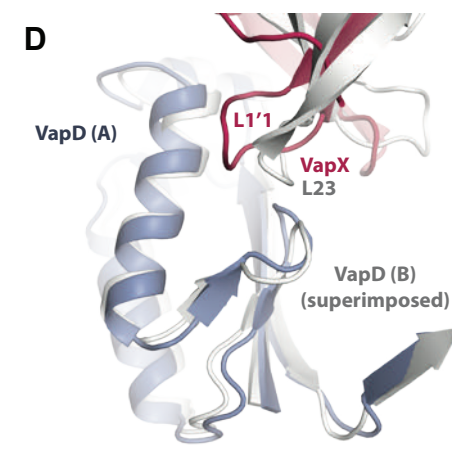
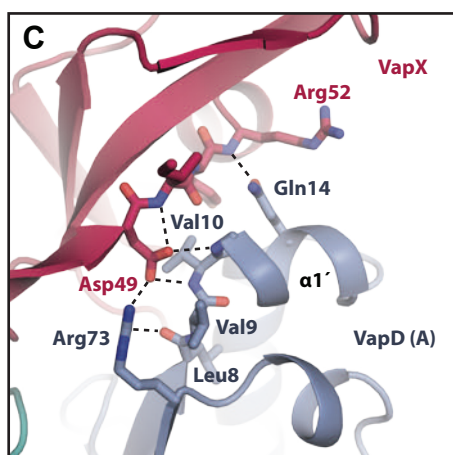
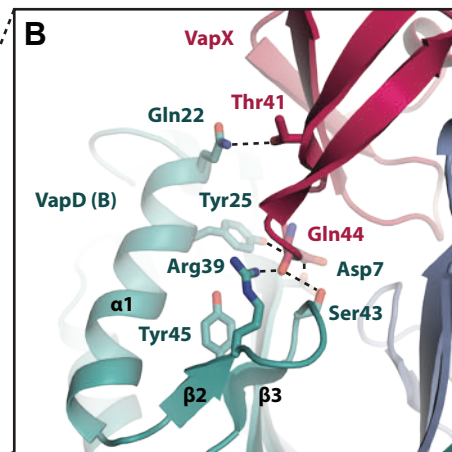
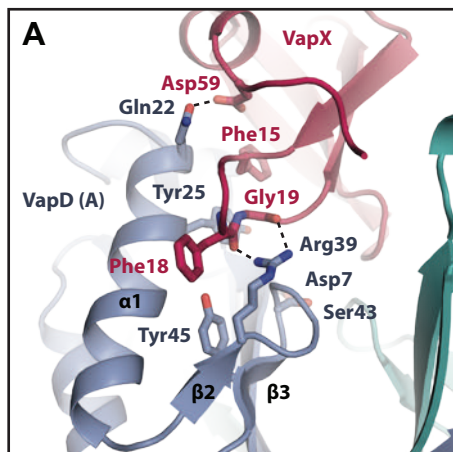
B

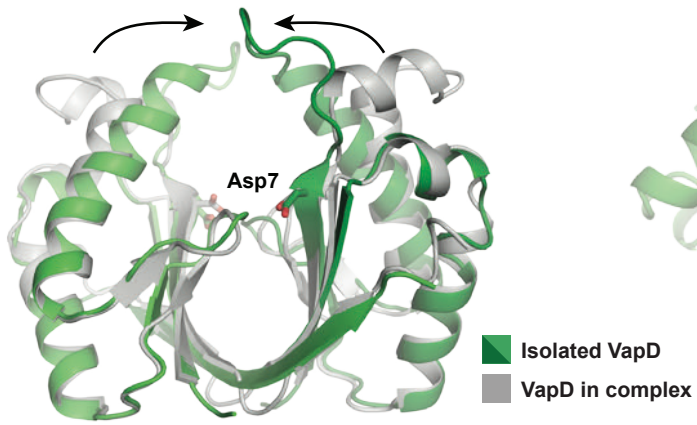
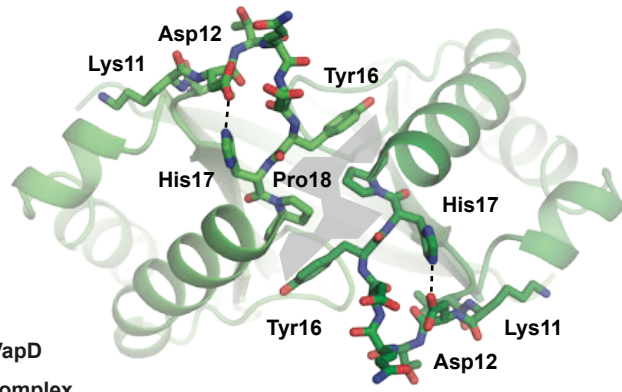
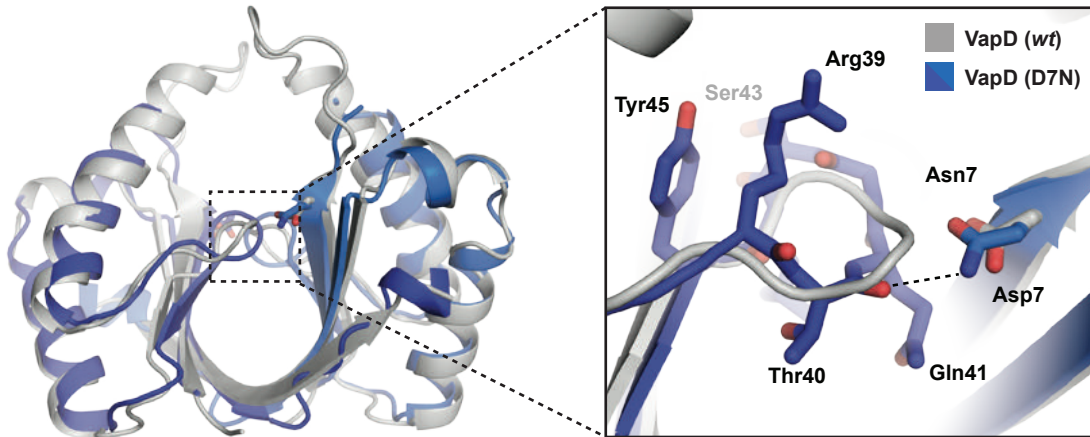


C

H. influenzae VapX
A. pleuropneumoniae
T. aquaticus
H. paracuniculus
E. coli
P. multocida
N. chenwenguii
S. enterica
Synechocystis spp.
P. luminescens
N. gonorrhoeae





A**B****C**

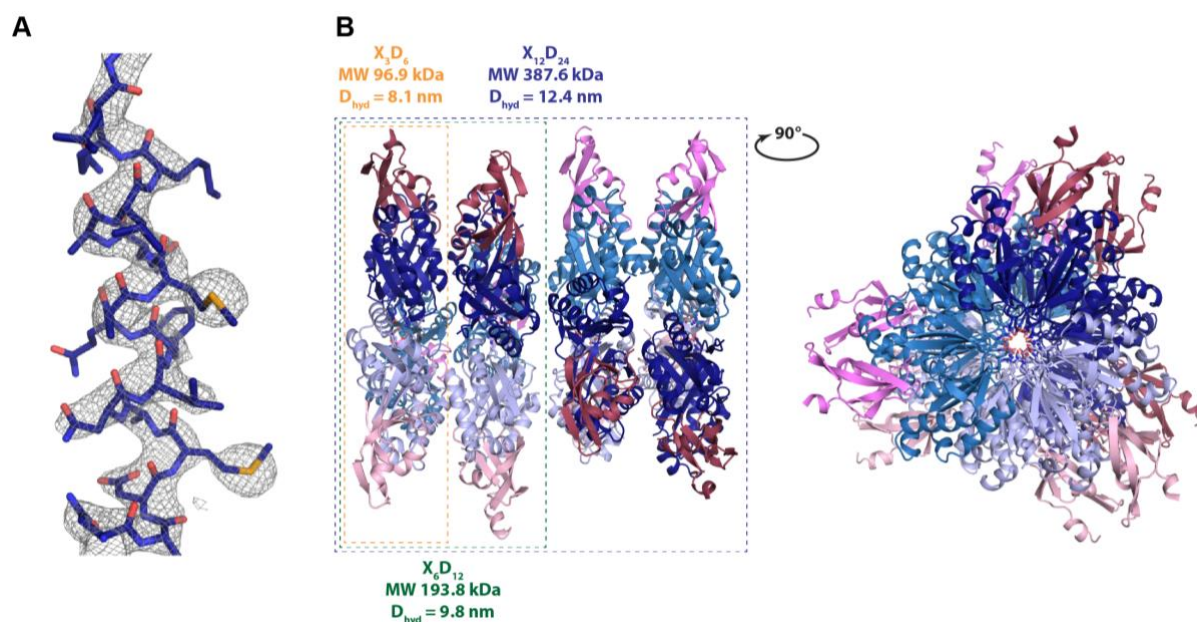


Figure S1. VapXD density and crystal packing, Related to Figure 1. **A.** Representative, final $2mF_o-DF_c$ density obtained during refinement of the structure of NTHi VapXD, contoured at 2.0σ . **B.** The ring-shaped trimer of VapXD complexes stacks in sets of four throughout the crystal as shown in the side (left) and end (right) views of the crystal packing. The orange, green, and blue boxes show the assemblies consisting of one, two, or four rings respectively, all proposed by PISA to be stable in solution (Krissinel & Henrick, 2007). The theoretical molecular masses in kDa (MW) and calculated hydrodynamic diameters (D_{hyd}) in nm for each of the possible assemblies are shown above. In the end view of four stacked rings, the conserved Arg79 and Glu81 that line the channel, are shown as sticks.

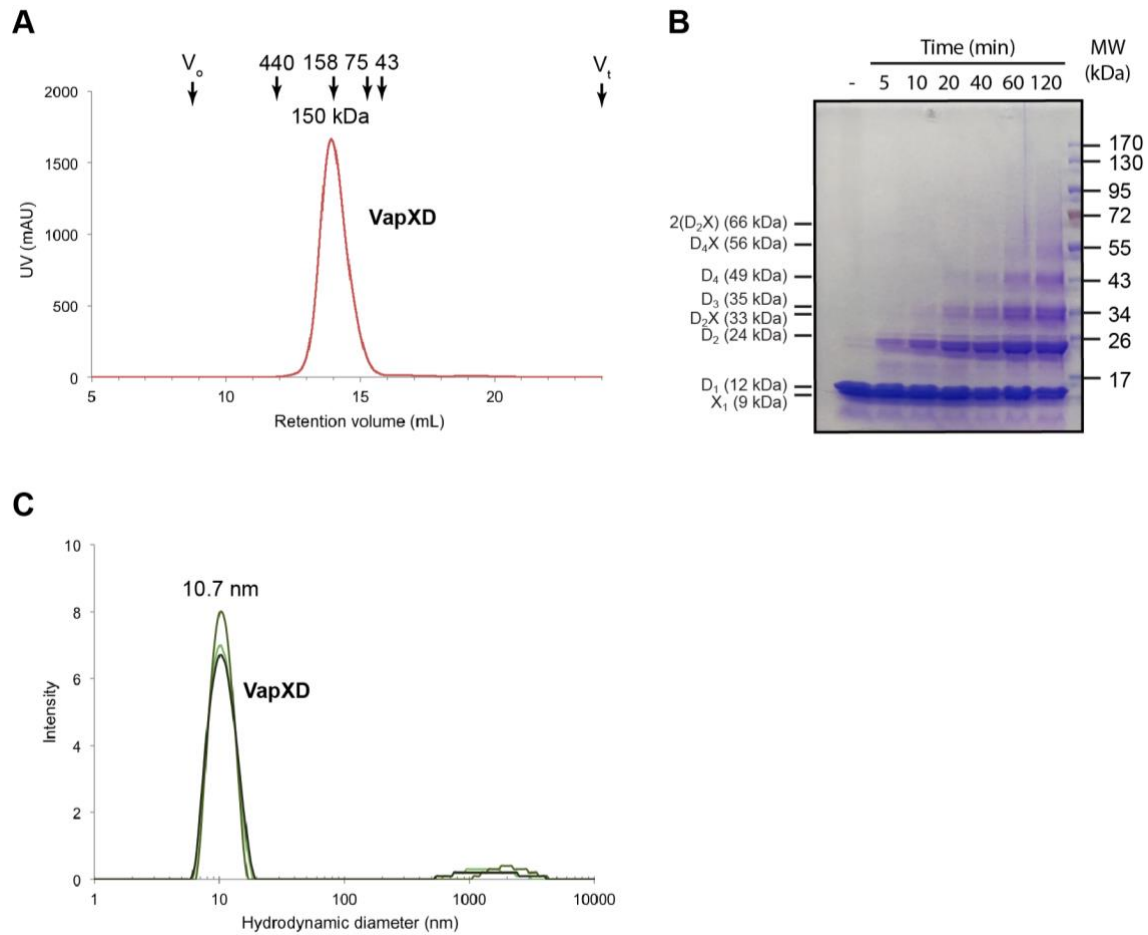


Figure S2. Biophysical analysis of VapXD higher-order structure, Related to Figure 1. A. Analytical gel filtration chromatogram of the VapXD complex using a 24 mL Superdex 200 Increase 10/300 GL column. The VapXD complex elutes with an apparent molecular mass of approximately 150 kDa. The void volume ($V_o \sim 8$ mL) and total volume ($V_t = 24$ mL) of the column are indicated along with the elution volume of standard proteins. **B.** Chemical cross-linking of the VapXD complex in solution using 0.5% glutaraldehyde. Possible stoichiometries of cross-linked species are shown on the left and a standard molecular weight marker is on the right. **C.** Dynamic light scattering measurements of VapXD in solution. Three consecutive DLS measurements are shown overlaid and reveal one major peak with a mean hydrodynamic diameter of 10.7 ± 0.1 nm.

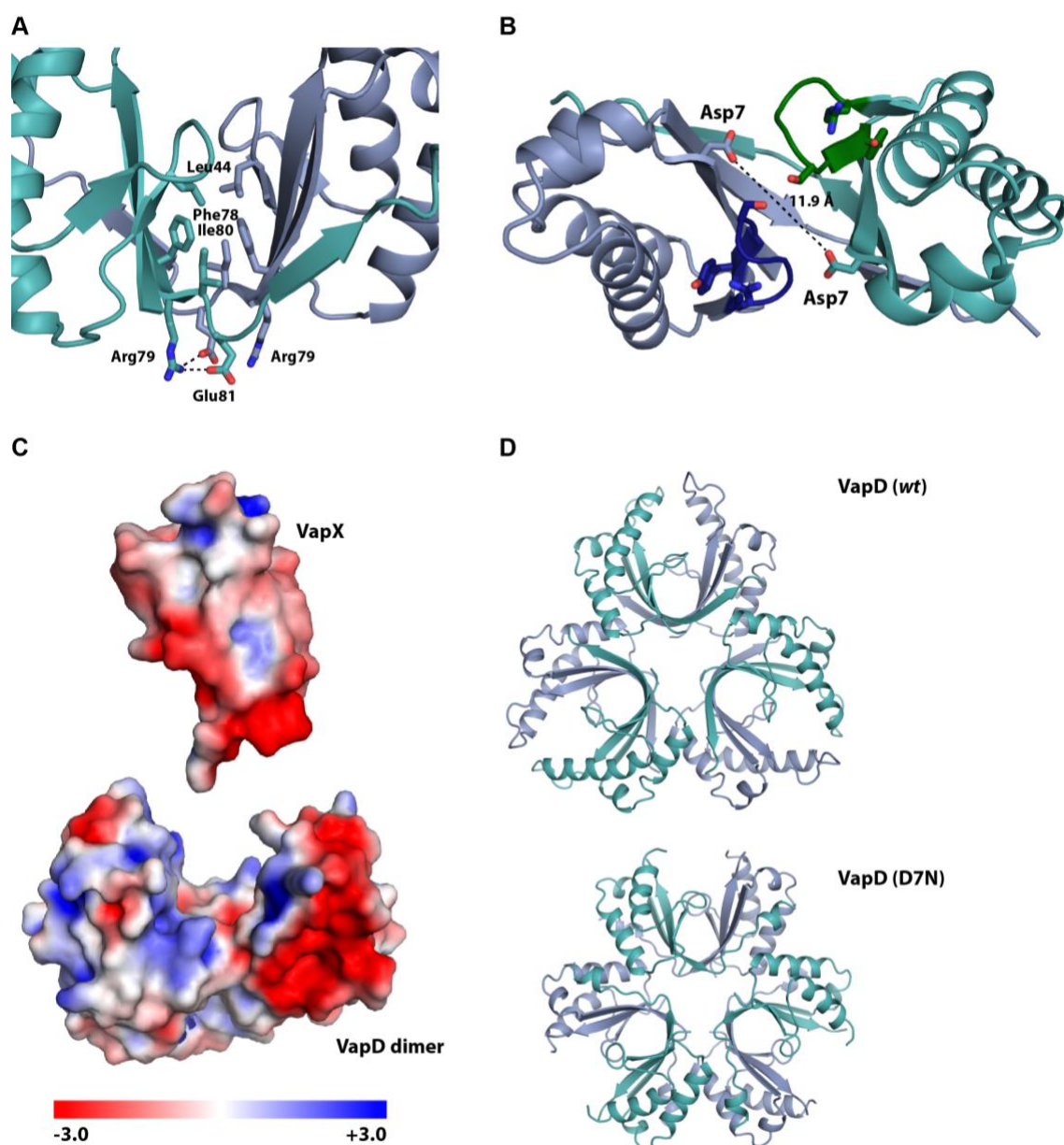


Figure S3. VapXD interactions and higher-order structure of isolated VapD, Related to Figure 2. **A.** Details of the interactions at the interface between VapD monomers. The cluster of hydrophobic residues (Leu44, Phe78, and Ile80) are shown along with the charged residues (Arg79 and Glu81) that also stabilise the ring structure (see text for details). **B.** Cartoon representation of the VapD dimer as observed in the VapXD complex with residues putatively involved in catalysis shown as sticks. The measured distance between the two Asp7 residues (11.9 Å) is shown. The loop connecting $\beta 2$ and $\beta 3$, harbouring conserved residues previously found to be important for activity is coloured dark blue/green in the two monomers. **C.** The electrostatic surface potential of the VapD dimer and VapX, which is lifted out of the VapD dimer for clarity. The scale shows the average surface charge from -3.0 to +3.0. Electropositive patches suitable for RNA binding is seen on the sides of the cleft in the VapD dimer. VapX has electronegative regions on the end that binds the VapD dimer, which nicely complements the electropositive patches in the cleft of the VapD dimer. **D.** Symmetry expansion of the structures of isolated VapD wt (top) and D7N mutant (bottom) structures reveal similar ring-shaped assemblies as the VapXD complex.

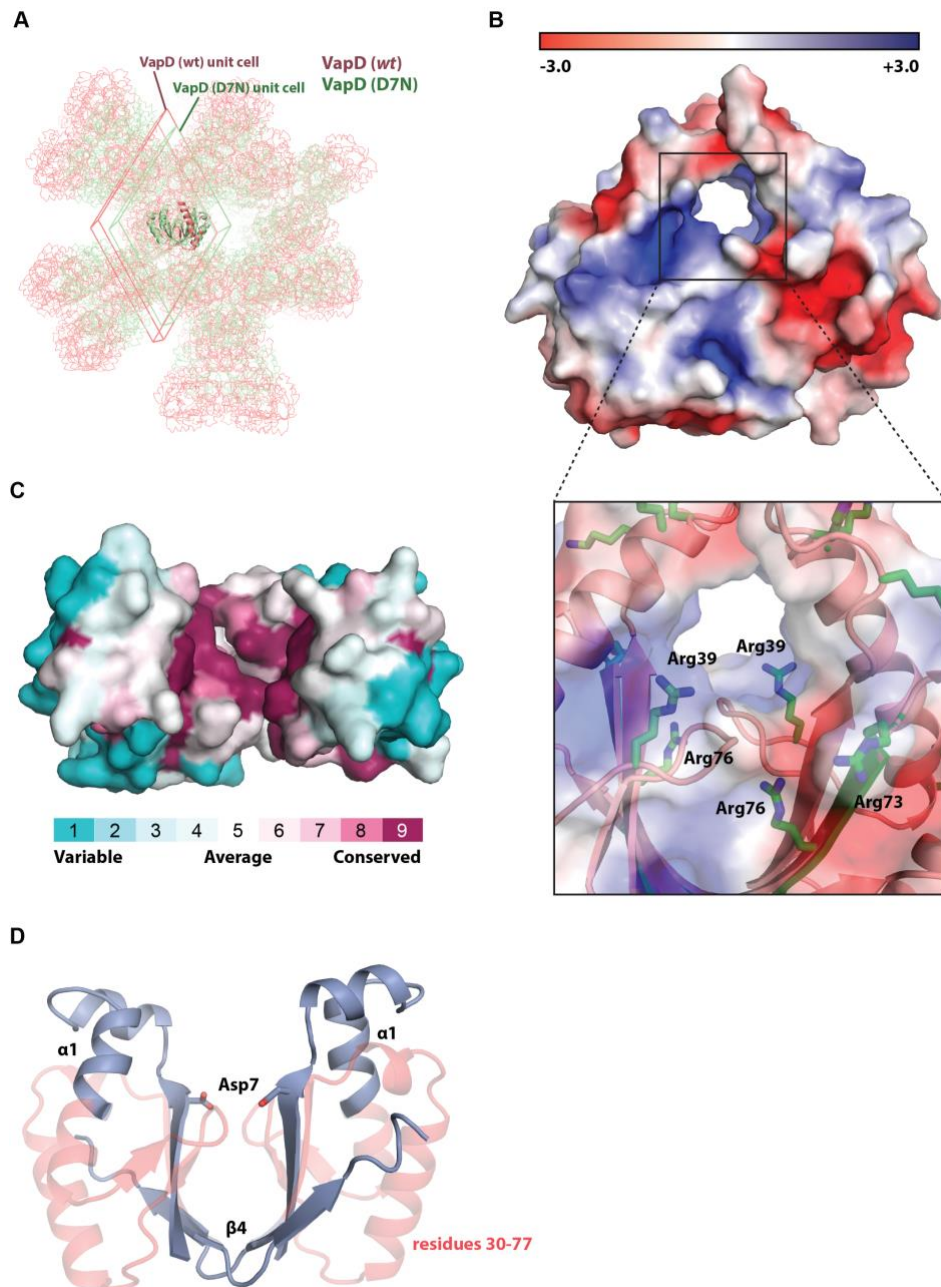


Figure S4. Structural features of isolated VapD, Related to Figure 5. **A.** Crystal packing of the isolated VapD wt (P6₃22) and D7N structures (P321), with unit cells indicated. **B.** The electrostatic surface potential of the isolated VapD dimer (top) showing the highly electropositive channel (in blue) through the dimer. The scale shows the average surface charge from -3.0 to +3.0. The inset (bottom) shows basic residues involved in generating the positive potential. **C.** Surface representation of the NTHi VapD dimer coloured according to conservation of the residues as shown in the bar below. The proposed active site in the cleft between the two monomers of the VapD dimer is highly conserved. The colouring scheme was prepared using the ConSurf Server (Ashkenazy et al., 2016) and is based on 150 VapD sequences with a sequence identity to NTHi VapD ranging from 35-95%. **D.** Natural variants of VapD exist for which residues 30-77 of the protein are missing (shown in faint red), thus connecting $\alpha 1$ to $\beta 4$. This would leave the catalytic Asp7 (indicated) as well as the β strands required for dimerisation in place.

Formula	Composition (chains)	Surface area (Å ²)	Buried area (Å ²)	ΔG^{int} (kcal/mol)	ΔG^{diss} (kcal/mol)
VapX ₆ VapD ₁₂	A ₆ B ₆ C ₆	60,760	42,830	-195.5	2.0
VapX ₃ VapD ₆	A ₃ B ₃ C ₃	31,230	20,570	-89.0	12.4
VapX VapD ₂	ABC	12,150	5110	-21.3	4.9
VapD ₁₂	A ₆ B ₆	48,570	28,710	-147.3	3.6
VapD ₆	A ₃ B ₃	25,130	13,510	-64.9	14.2
VapD ₂	AB	10,120	2760	-13.3	8.9

Table S1. Possible VapXD higher-order structures, Related to Figure 1. Molecular composition of proposed, stable molecular species, their surface and buried area, and calculated interaction energies from PISA based on the NTHi VapXD₂ crystal structure, for which VapD occupies chains A and B, and VapX chain C (Krissinel and Henrick, 2007).