

# Molecular architecture and activation of the insecticidal protein Vip3Aa from *Bacillus thuringiensis*

Rafael Núñez-Ramírez<sup>1</sup>, Juanjo Huesa<sup>2,3</sup>, Yolanda Bel <sup>3,4</sup>, Juan Ferré <sup>3,4</sup>, Patricia Casino <sup>2,3,5</sup> ✉ & Ernesto Arias-Palomo <sup>1</sup> ✉

*Bacillus thuringiensis* Vip3 (Vegetative Insecticidal Protein 3) toxins are widely used in biotech crops to control Lepidopteran pests. These proteins are produced as inactive protoxins that need to be activated by midgut proteases to trigger cell death. However, little is known about their three-dimensional organization and activation mechanism at the molecular level. Here, we have determined the structures of the protoxin and the protease-activated state of Vip3Aa at 2.9 Å using cryo-electron microscopy. The reconstructions show that the protoxin assembles into a pyramid-shaped tetramer with the C-terminal domains exposed to the solvent and the N-terminal region folded into a spring-loaded apex that, after protease activation, drastically remodels into an extended needle by a mechanism akin to that of influenza haemagglutinin. These results provide the molecular basis for Vip3 activation and function, and serves as a strong foundation for the development of more efficient insecticidal proteins.

<sup>1</sup>Centro de Investigaciones Biológicas Margarita Salas, CSIC, 28040 Madrid, Spain. <sup>2</sup>Departament de Bioquímica i Biologia Molecular, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain. <sup>3</sup>Estructura de Recerca Interdisciplinària en Biotecnologia i Biomedicina (ERI BIOTECMED), Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain. <sup>4</sup>Department of Genetics, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain. <sup>5</sup>CIBER de Enfermedades Raras (CIBERER-ISCI), Madrid, Spain. ✉email: [patricia.casino@uv.es](mailto:patricia.casino@uv.es); [earias@cib.csic.es](mailto:earias@cib.csic.es)

The bacterium *Bacillus thuringiensis* (Bt) synthesizes potent insecticidal proteins, such as the Cry toxins generated during sporulation, commonly used in agriculture<sup>1,2</sup>. In addition, Bt secretes to the medium numerous vegetative insecticidal proteins (Vip) that have been classified into three different families<sup>3</sup>. Notably, members of the Vip3 family have shown to have a broad-spectrum activity against Lepidoptera<sup>4</sup> and, since they target binding sites different to those of Cry proteins in the insect membranes<sup>5–11</sup>, they are considered a good complement to the action of these toxins. Crystal/spore mixtures of some Bt strains are indeed the active ingredient of bioinsecticides used in traditional and organic farming, and the genes coding for Cry and Vip3 proteins have been introduced in plants (Bt crops) to protect them from insect pests<sup>2,3</sup>.

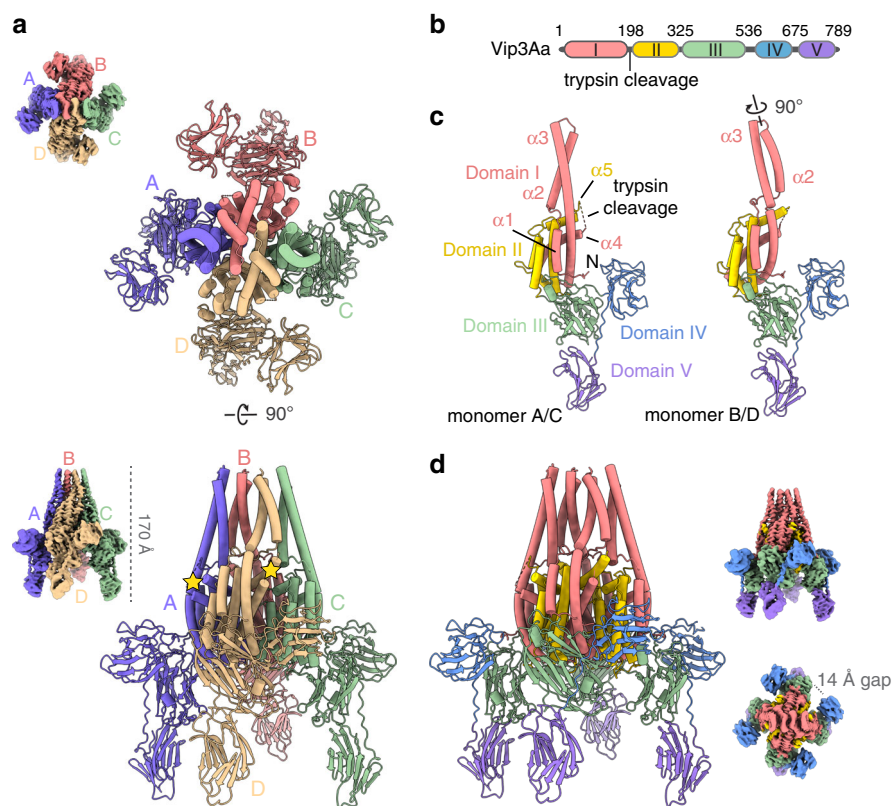
Vip3 proteins are synthesized in an inactive state, as protoxins, and they require to be activated by proteases present in the midgut tract of the insect. Trypsin-like enzymes cleave the molecule at residue 198, generating two fragments of around 19 and 65 kDa that remain strongly associated<sup>7,12–15</sup>. The activated protein then recognizes specific receptors in the brush border of columnar cells, whose identity is still under debate, and it has been proposed that binding to some of them can induce internalization<sup>16–18</sup>. Notably, although several cell lethality pathways have been proposed, numerous studies indicate that Vip3 proteins form pores in the membrane to trigger cell death<sup>6,8,13,19,20</sup>.

The first member of the Vip3 family was discovered more than 20 years ago. However, numerous questions persist regarding the three-dimensional organization, activation and mode of action of these proteins at the molecular level. To shed light on some of these questions, we have determined the near-atomic resolution

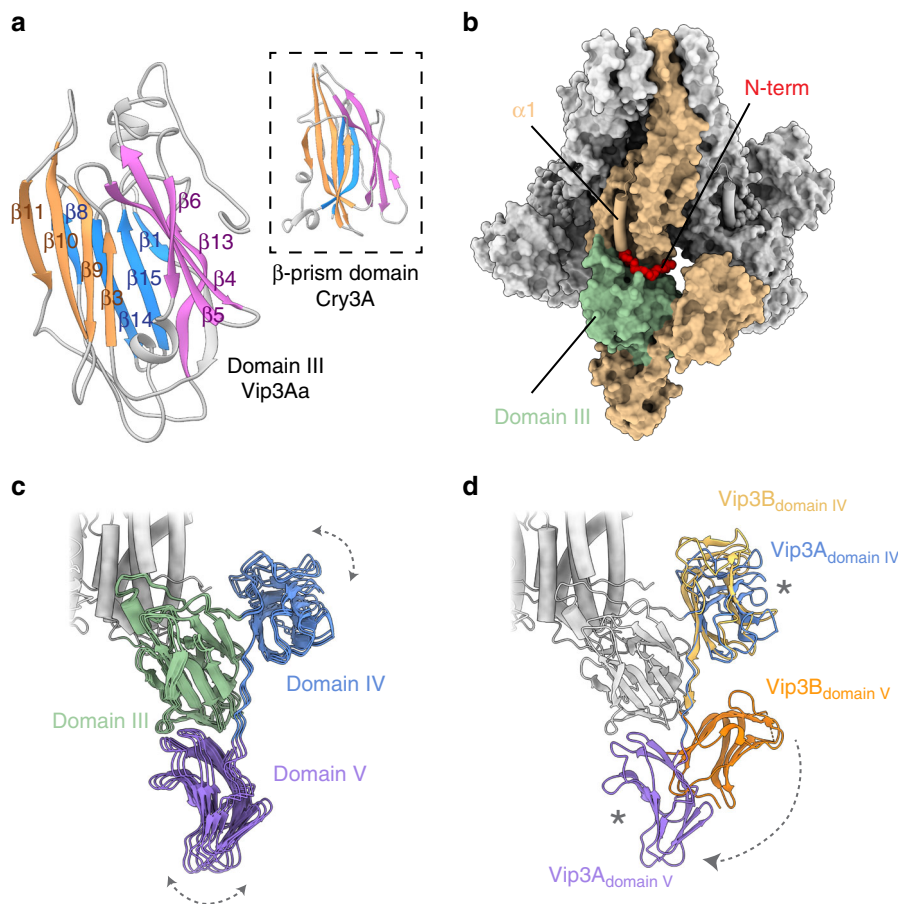
structures of Vip3Aa in the protoxin and activated states by cryo-electron microscopy (cryo-EM). The reconstruction of the uncleaved protein reveals that it assembles into a highly stable tetrameric dimer of dimers that contains three putative sugar-binding domains poised for receptor engagement. The trypsin-activated structure unravels the large conformational changes that occur upon protease digestion in the N-terminal region, and that leads to the formation of an extended four-helix coiled coil. Collectively, our work provides a high-resolution view of a Vip3 member in the protoxin and toxin conformation and allows us to propose an activation mechanism for this family of entomopathogenic proteins that both parallels and diverges from other toxins and viral fusion proteins.

## Results

**Molecular architecture of the Vip3Aa protoxin.** To better understand the three-dimensional organization and activation mechanism of these toxins, purified Vip3Aa was applied to electron microscopy grids coated with a thin carbon layer and analyzed using cryo-EM. The reference-free 2D averages of the motion-corrected micrographs show, in agreement with previous reports, the tetrameric organization of the protoxin<sup>12,15,21–24</sup> (Fig. 1, Supplementary Figs. S1, S2). Interestingly, the three-dimensional reconstruction, which converged at a global resolution of 2.9 Å (Fig. 1a, Supplementary Fig. S1 and Supplementary Video 1), revealed that the Vip3 tetramer assembles into a pyramid-shaped structure where the N-terminal region forms the core and apex, and the C-terminal domains are exposed to the solvent. Overall, the high-resolution structural features and the quality of the EM map have permitted unambiguous de novo



**Fig. 1** Prototoxin structure of Vip3Aa from *Bacillus thuringiensis*. **a** Two orthogonal views of the Vip3Aa protoxin. Each monomer has been colored differently. Yellow stars mark the position of the trypsin cleavage site in two of the monomers. Unsharpened EM map shown on the left. **b** Primary structure of Vip3Aa. **c** Vip3Aa is formed by two types of monomers, outer (left), and inner (right), that are mainly differentiated by a  $-90^\circ$  rotation of the  $\alpha 2$ – $\alpha 3$  helical bundle. **d** Vip3Aa tetramer colored according to the domain organization. Two orthogonal views of the EM reconstruction are shown on the right.



**Fig. 2** The flexible C-terminal domains contain motifs found in other toxins. **a** Domain III forms a  $\beta$ -prism fold similar to that found in the Cry toxins. **b** The N-terminal region of the four monomers (red) is inserted into a groove formed between domain III (green) and the core of the particle. **c** Superposition of domains III, IV and V, docked into the three most distant positions found in the focused 3D classification, shows the flexibility of these domains in the Vip3Aa reconstruction. **d** Although the comparison with the Vip3B protoxin (PDB ID 6V1V) reveals a similar overall architecture, the orientation of domain V is markedly different to that found in Vip3Aa. The putative sugar binding site at domain IV and V is labeled as an asterisk.

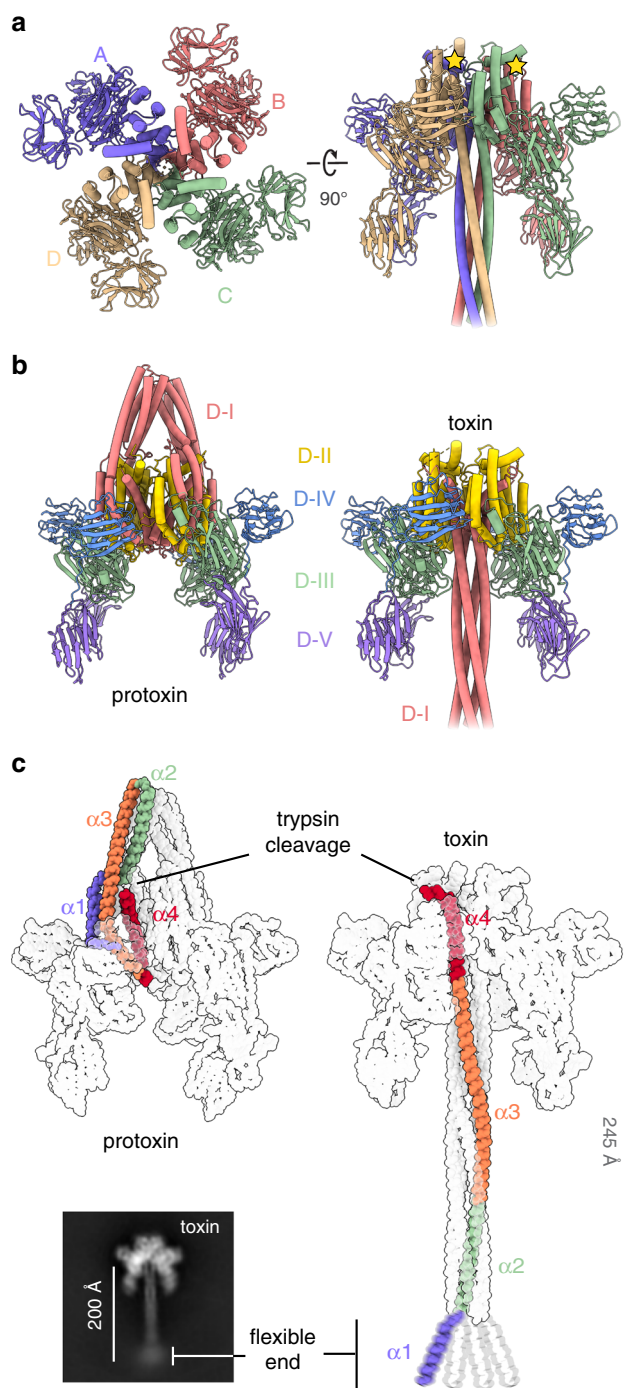
atomic modeling of the Vip3Aa protoxin (Fig. 1a and Supplementary Table S1).

The cryo-EM structure of the protoxin reveals key insights into the molecular architecture of the protein, and shows that Vip3 is composed of five distinct domains (Fig. 1b–d and Supplementary Fig. S3), confirming earlier predictions<sup>23</sup>. Domain I extends from the N-terminus to the primary protease cleavage site (K198)<sup>14,25</sup>, and contains four long, highly curved  $\alpha$ -helices ( $\alpha 1$ – $\alpha 4$ ). The first helix is markedly amphipathic (Supplementary Fig. S4), and the second and third helices form an antiparallel bundle that adopts a different configuration in the outer (chains A and C) and inner monomers (chains B and D) (Fig. 1c). In particular, the variable length of the  $\alpha 1$ – $\alpha 2$  loop allows the second and third helices of the internal monomers to rotate  $\sim 90^\circ$  compared with the external ones (Supplementary Video 2). This rotation further increases the curvature and tension of this element and leads to the formation of an eight-helix bundle that protrudes as a stalk from the tetrameric core formed by the subsequent domain (Fig. 1a, d). Domain II (residues 200–325) starts immediately after the cleavage site, which is located in a fully exposed loop that connects helix  $\alpha 4$  and  $\alpha 5$  (Fig. 1c). Although some of the most solvent accessible residues (T190–P202) were flexible—probably to facilitate trypsin cleavage—and were not modeled, there is clear density for the entire loop in the unsharpened EM map (Supplementary Fig. S5a, b). Notably, this second domain, composed of five  $\alpha$  helices, embodies the core of the tetramer (Fig. 1d) and plays a structural role mainly through two extended

loops (comprising residues 221–226 and 239–247, respectively) that project toward the adjacent subunit and stabilize the Vip3 oligomer (Supplementary Fig. S5c, d).

In sharp contrast, the C-terminal region of Vip3 is composed of three globular domains that are mostly solvent exposed and do not show any interactions between different monomers. Indeed, there is a clear gap of  $\sim 14$  Å (minimum distance) that separates these domains from the neighboring monomer (Fig. 1c, d). Domain III comprises residues 328–532 and contains three antiparallel  $\beta$ -sheets that form a  $\beta$ -prism fold strikingly similar to that found in members of the Cry insecticidal  $\delta$ -endotoxins<sup>26</sup> (Fig. 2a). Although in some Cry toxins this motif has been shown to recognize the sugar moieties that decorate the receptors of the target cell<sup>27,28</sup>, the exact role in Vip3A proteins has not yet been fully elucidated. This third domain, interestingly, is interacting and clamping the most N-terminal segment of the protein (residues 14–23), against the core of the tetramer (Fig. 2b).

The EM density for the last two domains is less well ordered. However, we were able to improve significantly the quality of the density in this region by performing a cycle of symmetry expansion, particle subtraction and focussed classification. The analysis confirmed the inherent flexibility of these domains—that make loose contacts with domain III—and allowed us to determine their 3D structure (Figs. 1, 2c and Supplementary Video 3). Notably, these two domains bear similar CBM (carbohydrate-binding module) folds that are connected by a long linker. Although the precise type of substrates that these domains could bind was not



**Fig. 3 Molecular architecture of the trypsin activated Vip3Aa toxin.**

**a** Two orthogonal views of the Vip3Aa toxin show that the four monomers are organized around a central pore. The different monomers have been colored as in Fig. 1a. Yellow stars indicate the position of the trypsin cleavage site in two of the monomers. **b** Similar views of the Vip3Aa protoxin and toxin colored by domain organization as in Fig. 1b. Domains are labeled as D-I through D-V. **c** The alpha helices that comprise domain I undergo a large reorganization that leads to the formation of an extended coiled coil. The four helical bundle is ordered starting from residue -45, and reference-free 2D averages show that the most N-terminal region appears to be flexible.

clear, a recent study indicates that they have a strong preference for chitosan and chitin<sup>29</sup>, and the DALI server revealed that they show similarity to glycoside hydrolases, which are often found in tandem to increase affinity for polysaccharides<sup>30,31</sup>.

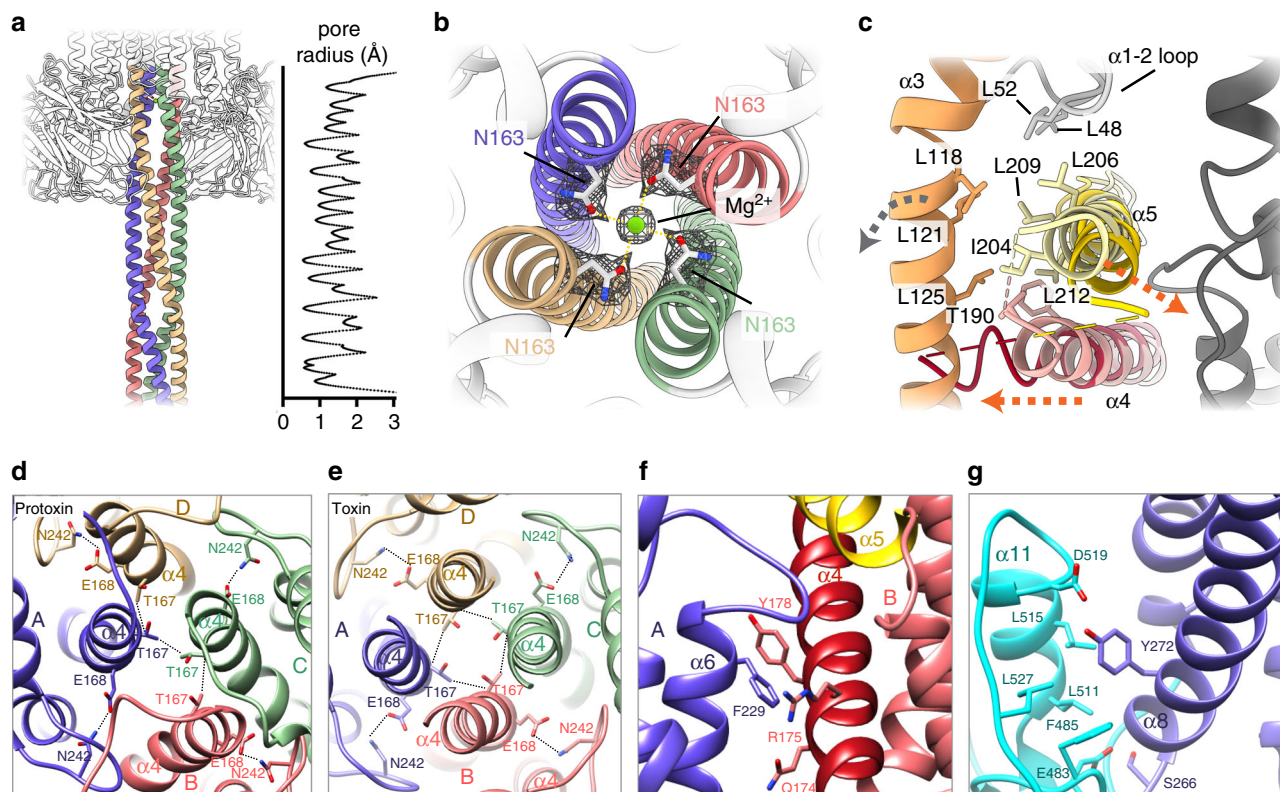
Interestingly, during the preparation of this manuscript the protoxin structure of Vip3B was published<sup>32</sup>, which allows us to compare the molecular architecture of two different inactive Vip3 variants. The structural analysis reveals that both proteins share a similar overall tetrameric organization. More precisely, domains I, II, and III show a remarkable structural similarity (RMSD value of  $\sim 1$  Å for 1932 residues), consistent with their elevated sequence homology ( $\sim 73\%$ ) (Supplementary Fig. S6). The isolated structures of the last two domains are also analogous. However, their position and orientation with respect to the core of the molecule differs slightly for domain IV and more pronouncedly for domain V, which rotates  $\sim 100^\circ$  (Fig. 2d and Supplementary Fig. S6). These results further confirm the flexible nature of the CBM domains that, importantly, maintain the putative glycan binding sites exposed to the solvent.

### Trypsin activation triggers a large molecular reorganization.

The analysis of the protoxin structure did not initially explain how protease digestion triggers the activation of Vip3 that leads to cell death. We therefore decided to solve the three-dimensional structure of Vip3Aa after trypsin treatment. The purified protein was first incubated with the protease using conditions that have been reported to produce full activation<sup>3,14</sup> (Supplementary Fig. S7a). The sample was then deposited onto open-hole grids coated with a thin layer of homemade carbon, flash-frozen, and examined using cryo-EM. Interestingly, although the efficiency of the digestion was elevated ( $>95\%$  according the PAGE gel), the 2D class averages showed that  $\sim 30\%$  of the molecules maintained a configuration remarkably similar to that of the undigested sample (Supplementary Fig. S7b). This observation probably reflects that, although protease cleavage is required, other factors might modulate the activation of Vip3. Notably, however, the majority ( $\sim 70\%$ ) of the particles adopted a syringe-like structure -toxin from now on- that after image processing converged into a 2.9 Å reconstruction (Fig. 3, Supplementary Figs. S7, S8).

Inspection of the structure of the protease-activated toxin reveals a tetramer where the four monomers adopt an identical conformation arranged around the fourfold symmetry axis of the particle, and confirms that the cleaved fragments remain tightly associated<sup>7,12,14,15</sup> (Fig. 3a, b and Supplementary Video 4). Activated Vip3Aa shows the same domain organization described in the protoxin state, and domains II-V maintain a similar global architecture (Fig. 3b and Supplementary Fig. S9). Domain I, however, suffers a large conformational change that remodels the apex of the pyramid into a  $\sim 200$  Å-long needle. Notably, the required set of movements can proceed unimpeded through the gap located between monomers. In particular, three antiparallel N-terminal helices ( $\alpha 2$ - $\alpha 4$ ) undergo a set of rotations, in a manner akin to a folding cane, to form a long continuous helix that assembles into a parallel four-helical coiled-coil in the digested Vip3Aa tetramer (Fig. 3c). The cryo-EM analysis shows that the coiled-coil is relatively ordered starting at the beginning of  $\alpha 2$  ( $\sim D45$ ), and that the flexible end is formed by the most N-terminal region of the protein, which contain the residues forming the first alpha helix in the protoxin state (Fig. 3c).

Close examination of the structure of the toxin revealed that the coiled-coil of Vip3Aa, which resembles other four-helix bundles such as the SNARE complex<sup>33,34</sup> (RMSD of 3.4 Å for 232 residues), extends  $\sim 160$  Å and that the average internal radius of the modeled section was 1.46 Å (Fig. 4a). Classic coiled-coils contain a series of heptad repeats (*abcdefg*) where the residues that line the lumen (*a* and *d*) are usually hydrophobic to maintain the distinctive knobs-into-holes packing of this motif<sup>35,36</sup>. In the case of Vip3Aa, although most residues found at positions *a* and *d* were hydrophobic ( $\sim 80\%$ ), a significant fraction were polar,



**Fig. 4 Dimensions and formation of the coiled coil.** **a** Side view of the coiled coil region modeled at atomic resolution (residues 95–167). The corresponding pore radius is plotted on the right. Source data are provided as a Source Data file. **b** A magnesium ion, coordinated by Asn 163, was modeled into the central pore of the cryo-EM structure. **c** A hydrophobic interaction network, formed by residues located in  $\alpha 3$ ,  $\alpha 4$ ,  $\alpha 5$ , and the  $\alpha 1$ – $\alpha 2$  loop (light orange, red, yellow and gray respectively), helps to maintain Vip3Aa in an inactive configuration. Helices  $\alpha 4$  and  $\alpha 5$  reposition upon trypsin cleavage (dark red and yellow helices, orange arrows indicate the direction of the movement), disrupting the hydrophobic patch and promoting the conformational change of the N-terminal region of the protein (indicated by a gray arrow). **d** Representation of inter-monomer interactions between residue T167 (domain I), as well as the intra-monomer interactions between E168 (domain I) and N242 (domain II), in the protoxin. **e** Same contacts shown in panel **d** in the toxin structure. **f** Inter-monomer interactions between F229 at domain II with residues located at the central part of  $\alpha 4$  in the adjacent monomer (dark red), highlighting  $\alpha 5$  (in yellow). **g** Intra-monomer interactions between domain II (dark blue) and domain III (light blue).

with asparagine being the most abundant one. Interestingly, comparative analysis of different structures, like that of the Sendai virus P protein, have shown that hydrophilic residues found in these positions—known as N@d layers—can bind ions and water molecules in the central cavity of the tetrameric coiled-coil<sup>37,38</sup>. Although the resolution of the cryo-EM reconstruction does not allow visualization of water molecules, there is a clear peak of density coordinated by four symmetry related asparagines (N163) that, based on the buffer composition (5 mM  $\text{MgCl}_2$ ), was assigned to a magnesium ion (Fig. 4b).

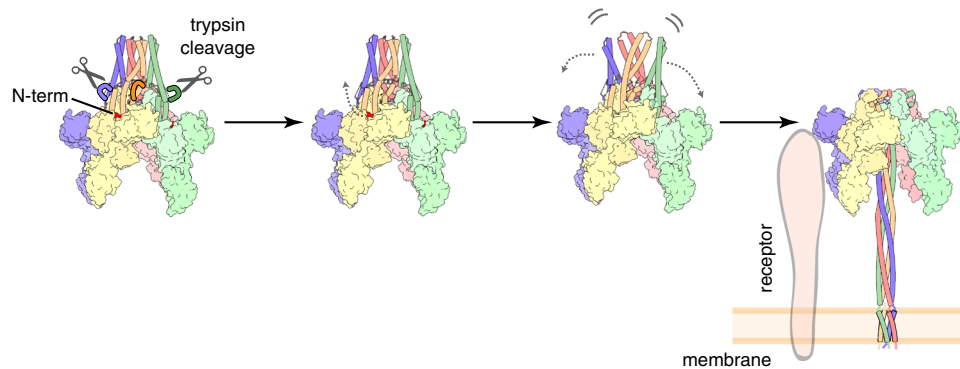
#### Remodeling and molecular basis for the activation of Vip3Aa.

Detailed comparison of the protoxin and toxin structures reveals the dramatic reorganization that the protein undergoes upon protease digestion, but what triggers the conformational change? Part of the answer is that the cleavage of the trypsin loop allows the re-configuration of the adjacent helices ( $\alpha 4$  and  $\alpha 5$ ) (Fig. 4c). The N-terminal region of  $\alpha 5$  moves toward the neighbor monomer and disrupts a hydrophobic interaction network, between helix  $\alpha 3$  and the loop between  $\alpha 1$  and  $\alpha 2$ , that had been previously shown to be critical for the protein<sup>39</sup>. The C-terminal part of  $\alpha 4$ , in turn, moves against  $\alpha 3$ , applying pressure to further facilitate the conformational change that leads to the formation of the N-terminal coiled-coil and the activation of the protein. Concomitantly, several residues in the N-terminal part of  $\alpha 4$  remodel to nucleate the formation of the coiled-coil. In this way,

T167, which is generating a network of inter-monomer interactions in the protoxin, is the first residue to realign and nucleate the apical section of the coiled-coil, with the side chains pointing toward its central axis (Fig. 4d, e). Similarly, N163 is also forming an inter-monomeric network of interactions in the protoxin but relocates right one helix turn below T167 at the coiled-coil to bind the divalent ion observed in the toxin configuration (Fig. 4b). Strikingly, although the N-terminal part of  $\alpha 4$  swivels during the activation of the protein, the C-terminal portion remains firmly attached to the domain II (Fig. 4f), which acts as an anchor point and provides a stable framework during the conformational changes that lead to the activation of the protein.

#### Discussion

In 1991 the first 3D structure of a Cry protein was reported<sup>26</sup>, providing key insights into their function and facilitating the improvement of their insecticidal properties. In contrast, structural studies of Vip3A have lagged behind limiting the understanding of their function. The results presented here solve many long-standing questions concerning the molecular organization and activation of the Vip3 family of toxins and help to explain and reconcile some difficult-to-interpret functional and biological data obtained during the last decades. The structures, for instance, shed light on the controversial role of the N-terminal 198 amino acids (domain D)<sup>3</sup> and the puzzling observation that after trypsinization this fragment remains associated to the rest of



**Fig. 5 Activation mechanism of Vip3Aa.** Protease digestion is required for Vip3 activation as it facilitates, possibly in conjunction with other factors, the remodeling of the spring-loaded helices that form the protoxin apex. This set of long-range conformational changes releases the N-terminus segment of the protein and leads to the formation of a four-helix coiled coil, which is sufficiently long to reach and permeate the lipid bilayer.

the protein. Domain I is indeed retained after protease digestion and stays tightly bound to the core of the protein through interactions with domain II, forming an essential part of the active toxin. Moreover, the structures explain why the exchange of this first domain between Vip3 variants can produce stable chimeras<sup>12,40</sup>, since the replacement of this highly conserved domain probably has a minor effect on the overall structure of the protein, and it also clarifies why attempts to produce Vip3 mutants in the trypsin loop or completely lacking domain I have been unsuccessful or have rendered an inactive or unstable proteins<sup>12,25,41</sup>.

In addition, the cryo-EM structures also reveal why some mutants have a dramatic effect on the stability and activity of the toxin. For instance, recent use of the alanine scanning technique has shown a cluster of residues around the interface of domain I and II (from residue 167 to 272) whose replacement by Ala highly decreased the insecticidal activity of Vip3Aa to *Spodoptera frugiperda*<sup>39</sup>. Residues E168A, F229A, Y272A, and E483 maintain fundamentally the same contacts in the protoxin and toxin structures and most likely contribute to the stabilization of the protein during the needle formation. In this way, residue E168 interacts with N242, found in the domain II of the same monomer, and contributes to stabilize both domains during remodeling (Fig. 4d, e). Similarly, F229, which is located at  $\alpha 6$  of domain II, interacts with residues in the central part of  $\alpha 4$  of the adjacent monomer generating a hydrophobic pocket (Fig. 4f) that seems to maintain  $\alpha 4$  bound to the core of the tetramer upon needle formation. Residue Y272, located at  $\alpha 8$  in domain II, is setting network contacts with residues from  $\alpha 11$  at the C-terminal end of domain III of the same monomer that are also maintained in the toxin structure (Fig. 4g). E483 reinforces this interaction engaging the base of  $\alpha 8$  to bridge the core of the tetramer and the  $\beta$ -prism domain (Fig. 4g). Taken together, our data show that numerous mutants known to decrease the insecticidal activity lay in conserved key residues that have a crucial role in maintaining the integrity of the tetramer.

Overall, the structural transformations observed during Vip3 activation allow us to propose a spring-loaded mechanism that shares numerous similarities with that of influenza haemagglutinin and other viral fusion proteins<sup>42,43</sup>. It is well established that haemagglutinin is primed by protease digestion and that, after receptor engagement and stimulation by other factors, a hidden N-terminal insertion peptide unlocks from the core of the protein that undergoes a large conformational change to form a three-helix coiled-coil<sup>42,44</sup>. In the case of Vip3, the protein is first secreted and spontaneously assembles into a tetrameric protoxin. Following ingestion, the protein is processed by the insect's midgut proteases and crosses the peritrophic membrane to interact with surrogate

receptors in the apical membrane of the epithelial cells. Proteolytic cleavage is required to trigger a set of events that lead to the activation of the toxin. The N-terminus of the protein detaches from a pocket formed between domain III and the main body of the tetramer, a process that could be further stimulated by other factors like the interaction of the membrane receptors with the C-terminal region of Vip3, which could also increase the concentration of the toxin at the membrane (Fig. 5). Protease digestion, in addition, would release the accumulated tension in  $\alpha 3$  and would allow the  $\alpha 2$ - $\alpha 3$  bundle to rotate and revert to the configuration found in the outer monomers (A and C), destabilizing the apex of the protoxin structure and promoting the formation of the tetrameric N-terminal coiled-coil.

Interestingly, it has been previously established that Vip3, similarly to the Cry toxins, forms pores in the insect cell and it is thought to act as an ion channel<sup>6,8,13</sup>. Notably, the coiled-coil observed upon trypsin digestion forms a 160 Å-long dipole that can accommodate small ions in the inner cavity and, therefore, the dimension of this four-helix needle would be sufficient to reach and pierce the lipid bilayer. It is not completely clear, however, whether the first  $\alpha$ -helix that forms the flexible end of the toxin could aid in the process. Its length, commensurable with the width of the membrane<sup>45</sup>, and highly amphipathic nature suggests that it could become ordered in contact with the lipid bilayer to form a four-helix bundle similar to the Rocker transporter or the influenza M2 channel, which are known to be responsible of transporting protons and divalent metals through the membrane<sup>46-49</sup>. Further studies will be required to establish a detailed mechanism for receptor recognition and cell permeation.

In summary, this work offers a high-resolution picture of the architecture and activation mechanism of Vip3. The structural information and precise definition of the domain boundaries can help to generate chimeras with modified target specificity and, overall, these studies provide a molecular framework that can be used to guide future studies to develop engineered proteins with increased stability and insecticidal activity.

## Methods

**Protein expression and purification.** The Vip3 protein used in this study was Vip3Aa16 (NCBI accession No. AAW65132), previously named Vip3LB<sup>50</sup>. The *vip3Aa16* gene (residues from 10 to 789) was cloned into the vector LIC 1.5 ampicillin resistant (pETNKI-StrepII3C-LIC-amp) containing an N-terminal StrepII-tag followed by a 3C PreScission protease cleavage site<sup>51</sup> (the primers used were Vip3\_start10+: CAGGGACCCGGTACAAGAGCCTTACCAAGTTTATTTG and Vip3-FL: CGAGGAGAAGCCCGTTACTTAATAGAGACATCGTAAAAATGTAC). The protein was expressed in *Escherichia coli* C43 (DE3) cells grown in Luria-Bertani broth adding 0.5 mM isopropylthiogalactoside, upon reaching exponential phase, followed by incubation at 37 °C for 3 h. Cells were harvested by centrifugation at 4000 g for 15 min and stored at -20 °C until use. For protein purification, thawed cells were resuspended in buffer A (50 mM Tris-HCl

pH 8.0, 500 mM NaCl, 50 mM MgCl<sub>2</sub> plus 1 mM PMSF and 1 mM TCEP) and sonicated for 5 min. Then, the cell lysate was centrifuged at 25000 g for 30 min and the clarified supernatant was loaded into a 5 ml StrepTrap column (GE, Healthcare). The protein was eluted in 2.5 mM *d*-Desthiobiotin dissolved in buffer A. The StrepII-tag was removed upon incubation with 3C PreScission protease, fused to an N-terminal GST-tag, at 4 °C overnight during a dialysis procedure to remove *d*-Desthiobiotin. Again, affinity chromatography was performed using a 5 ml StrepTrap column to remove cleaved protein from uncleaved protein, followed by a GSTrap column (GE Healthcare) to remove the 3C PreScission protease. The Vip3Aa protein was further purified by gel filtration chromatography with a ProSEC 16/60 6-600 HR SEC column (Generon, UK). SDS-PAGE was used to assess the purity of the fractions and the ones with the highest purity (>95%) were pooled and concentrated with a Vivaspin 20–30 kDa cut off filter (Sartorius, Germany) till 6 mg ml<sup>-1</sup>. Then the protein (protoxin) was aliquoted, frozen with N<sub>2</sub>(l) and stored. To prepare the activated toxin sample, 250 µl protoxin (at 6 mg ml<sup>-1</sup>) was incubated with 2.5 µl trypsin (at 6 mg ml<sup>-1</sup>), prepared in buffer 100 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 mM MgCl<sub>2</sub> and 2 mM DTT, for 2 h at 37 °C. Then the sample was frozen and stored as described above.

**Sample preparation for electron microscopy.** For cryo-EM studies Vip3Aa was diluted (to a final concentration of 67.3 µg ml<sup>-1</sup>) in 0.1 M Tris (pH 8.0), 0.5 M NaCl, 5 mM MgCl<sub>2</sub>, 2 mM DTT and 0.01% NP40. After dilution, 3 µl of sample were applied to Quantifoil R2/2 300 mesh grids coated with a second layer of homemade thin continuous carbon, previously glow discharged in a GlowCube (25 mA, 10 s). After 1 min incubation, grids were blotted for 2 s and blot force of ~20 using FEI Vitrobot Mark IV at room temperature and plunged into liquid ethane. Flash-frozen grids were subsequently stored in liquid nitrogen. Trypsin digested Vip3Aa was prepared similarly, using a slightly different protein concentration (75.2 µg ml<sup>-1</sup>) and grid type (quantifoil R2/1 300 mesh).

**Electron microscopy data acquisition.** Cryo-EM grids were pre-screened in a JEOL 1230 and in a FEI Talos Arctica microscopes equipped with a TemCam-F416 (TVIPS, Gauting, Germany) and a Falcon III (FEI) camera, respectively. High-resolution data were collected on a Titan Krios electron microscope operated at 300 kV, and images were acquired using a Gatan K2 Summit direct electron detector, operated in electron counting mode. For undigested Vip3Aa, EPU (Thermo) was used to record 3511 micrographs at a defocus range of -1.0 to -3.0 µm, with a magnification of 1.048 Å px<sup>-1</sup> at the specimen level, and a total dose of 56.6 e<sup>-</sup> Å<sup>-2</sup> accumulated over 12 s and fractionated across 40 frames. For the digested Vip3Aa, 3090 micrographs were collected at a defocus range of -1.0 to -2.9 µm, with a magnification of 1.055 Å px<sup>-1</sup> at the specimen level, and a total dose of 60 e<sup>-</sup> Å<sup>-2</sup> accumulated over 11 s and fractionated across 44 frames.

**Image processing.** For both datasets, frames were aligned using MOTION-COR2<sup>52</sup>, which also generated drift-corrected summed images with and without electron-dose weighting. Micrographs were manually inspected to remove pictures that contained crystalline ice or other forms of visible contamination, resulting in a final dataset of 3411 micrographs for the undigested and 3090 for the digested sample. CTF parameters were estimated with GCTF<sup>53</sup> using the non dose-weighted images. Micrographs were picked with GAUTOMATCH, binned by two, and subjected to 2D classification using RELION 3<sup>54</sup>. The resulting 2D averages were then used as templates to pick again the entire dataset with GAUTOMATCH to generate initial particle stacks (729,405 Vip3Aa and 697,393 digested Vip3Aa particles). After one round of 2D classification, only averages showing high-resolution features were retained, which resulted in 478,797 for the undigested sample (protoxin from now on). 2D classification of the trypsin treated sample revealed a mixture of two conformations; some molecules (93,875 particles) remained in a configuration similar to the protoxin, whereas the majority (193,153 particles) adopted a distinct extended structure (toxin from now on) reminiscent to some elongated negative stained particles reported in earlier studies<sup>13</sup>.

Initial models were calculated using the SGD algorithm<sup>55</sup> implemented in RELION 3. These reconstructions were low-pass filtered to 40 Å and used as starting model for 3D refinement with RELION 3 using different symmetries (C1, C2, and C4), revealing that the protoxin contained symmetry C2 whereas the toxin showed clear C4 symmetry. The unbinned particles were then re-centered and re-extracted, and subjected to a round of 3D refinement CTF and beam tilt refinement and particle polishing as implemented in RELION 3. After performing the Bayesian polishing, the particles were analyzed using 3D classification using six classes and without imposing symmetry. The protoxin did not show mayor signs of heterogeneity and all the particles (478,797) were included into a final 3D refinement round, using C2 symmetry and a spherical mask, that resulted in a 2.9 Å reconstruction by gold-standard FSC at 0.143<sup>56</sup> (resolution and B-factor were calculated using a mask generated from a low-pass filtered map that followed the contour of the protein). In the case of the toxin, only the particles assigned to the class showing high-resolution features were selected (92,303 particles), and subjected to a 3D refinement run, imposing symmetry C4, that yielded a 2.9 Å reconstruction.

The quality of the density for most parts of the reconstructions was excellent. However, the four appendices that extend from the core of the structure, and that resulted to be the C-terminal domains (III–V), appeared to be not as well resolved,

which hampered the subsequent model building efforts of the last two domains. We, therefore, decided to perform a round of symmetry expansion, density subtraction (to erase from the particles the density corresponding to domains I and II), and focused 3D classification without alignments ( $K = 6$ ,  $T = 400$ ). The result from the classification revealed that the last three domains move indeed as a rigid-body. The density showing better quality and best connectivity was used to build the atomic model of domains IV and V. This procedure gave nearly identical results for the protoxin and toxin datasets.

**Model building and refinement.** The atomic models of domain I, II, and III were built de novo into the globally sharpened EM maps, using COOT<sup>57</sup>, for both the toxin and protoxin structures. The models were then subjected to iterative rounds of model building and real-space refinement with COOT and PHENIX using NCS constrains, and Ramachandran, rotamer, geometry, and secondary structure restraints<sup>57,58</sup>. To build domains IV and V, the Robetta server<sup>59</sup> was used to generate an initial model (the recently determined structure of Vip3B, PDB 61V1, was automatically selected as a reference). The homology model of these two C-terminal domains was then docked into the best model obtained during the focused classification and subjected to several rounds of manual model building and real-space refinement. Next, domains IV and V were placed in all the subunits of the tetramers using the symmetry operators to generate the full-length structures. Finally, the complete tetramers were subjected to an ADP-only refinement cycle into the overall reconstructions (global sharpening values of -64 and -69 Å<sup>2</sup>) to allow the temperature factors of the domains IV and V to change slightly to better reflect the flexibility of this part of the molecule. In the case of the protoxin, the structure was modeled from residue 14 to 789, with the exception of the solvent-exposed trypsin cleavage loop (residues 190–202). For the toxin, the atomic model was built from residues 95 to 789, although 2D and 3D analysis showed that the coiled coil region extends until residue ~45 and only the most N-terminal amino acids appear to be disordered. The internal diameter of the coiled-coil region of toxin was estimated with HOLE<sup>60</sup>, and figures were generated with CHIMERA<sup>61</sup> and CHIMERAX<sup>62</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data supporting the findings of this manuscript are available from the corresponding authors upon reasonable request. A reporting summary for this article is available as a Supplementary information file. Source data are provided with this paper. The cryo-EM densities (including sharpened and unsharpened maps, masks, focused classification reconstructions and half-maps), and atomic coordinates for the protoxin and toxin Vip3Aa structures have been deposited in the EMDB: EMD-10492 and EMD-10493, and PDB 6TFJ [<https://doi.org/10.22101/pdb6TFJ/pdb>] and PDB 6TFK [<https://doi.org/10.22101/pdb6TFK/pdb>], respectively.

Received: 20 December 2019; Accepted: 8 July 2020;

Published online: 07 August 2020

## References

- Palma, L., Muñoz, D., Berry, C., Murillo, J. & Caballero, P. *Bacillus thuringiensis* toxins: an overview of their bioicidal activity. *Toxins (Basel)*. **6**, 3296–3325 (2014).
- Sanchis, V. From microbial sprays to insect-resistant transgenic plants: history of the biopesticide *Bacillus thuringiensis*. A review. *Agron. Sustain. Dev.* **31**, 217–231 (2011).
- Chakroun, M., Banyuls, N., Bel, Y., Escriche, B. & Ferré, J. Bacterial vegetative insecticidal proteins (Vip) from entomopathogenic bacteria. *Microbiol. Mol. Biol. Rev.* **80**, 329–350 (2016).
- Estruch, J. J. et al. Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *Proc. Natl Acad. Sci. U.S.A.* **93**, 5389–5394 (1996).
- Lee, M. K., Miles, P. & Chen, J.-S. Brush border membrane binding properties of *Bacillus thuringiensis* Vip3A toxin to *Heliothis virescens* and *Helicoverpa zea* midguts. *Biochem. Biophys. Res. Commun.* **339**, 1043–1047 (2006).
- Lee, M. K., Walters, F. S., Hart, H., Palekar, N. & Chen, J.-S. The mode of action of the *Bacillus thuringiensis* vegetative insecticidal protein Vip3A differs from that of Cry1Ab -endotoxin. *Appl. Environ. Microbiol.* **69**, 4648–4657 (2003).
- Chakroun, M. & Ferré, J. In vivo and in vitro binding of Vip3Aa to Spodoptera frugiperda midgut and characterization of binding sites by <sup>125</sup>I radiolabeling. *Appl. Environ. Microbiol.* **80**, 6258–6265 (2014).
- Liu, J.-G., Yang, A.-Z., Shen, X.-H., Hua, B.-G. & Shi, G.-L. Specific binding of activated Vip3Aa10 to *Helicoverpa armigera* brush border membrane vesicles results in pore formation. *J. Invertebr. Pathol.* **108**, 92–97 (2011).

9. Sena, J. A. D., Hernández-Rodríguez, C. S. & Ferré, J. Interaction of *Bacillus thuringiensis* Cry1 and Vip3A proteins with *Spodoptera frugiperda* midgut binding sites. *Appl. Environ. Microbiol.* **75**, 2236–2237 (2009).
10. Chen, W.-B. et al. Transgenic cotton co-expressing chimeric Vip3AcAa and Cry1Ac confers effective protection against Cry1Ac-resistant cotton bollworm. *Transgenic Res.* **26**, 763–774 (2017).
11. Kahn, T. W. et al. Efficacy and resistance management potential of a modified Vip3C protein for control of *Spodoptera frugiperda* in maize. *Sci. Rep.* **8**, 16204 (2018).
12. Zack, M. D. et al. Functional characterization of Vip3Ab1 and Vip3Bc1: two novel insecticidal proteins with differential activity against lepidopteran pests. *Sci. Rep.* **7**, 11112 (2017).
13. Kunthich, T. et al. pH regulates pore formation of a protease activated Vip3Aa from *Bacillus thuringiensis*. *Biochim. Biophys. Acta Biomembr.* **1859**, 2234–2241 (2017).
14. Bel, Y., Banyuls, N., Chakroun, M., Escriche, B. & Ferré, J. Insights into the structure of the Vip3Aa insecticidal protein by protease digestion analysis. *Toxins*. **9**, 131 (2017).
15. Banyuls, N., Hernández-Martínez, P., Quan, Y. & Ferré, J. Artefactual band patterns by SDS-PAGE of the Vip3Af protein in the presence of proteases mask the extremely high stability of this protein. *Int. J. Biol. Macromol.* **120**, 59–65 (2018).
16. Singh, G., Sachdev, B., Sharma, N., Seth, R. & Bhatnagar, R. K. Interaction of *Bacillus thuringiensis* vegetative insecticidal protein with ribosomal S2 protein triggers larvicidal activity in *Spodoptera frugiperda*. *Appl. Environ. Microbiol.* **76**, 7202–7209 (2010).
17. Jiang, K. et al. Scavenger receptor-C acts as a receptor for *Bacillus thuringiensis* vegetative insecticidal protein Vip3Aa and mediates the internalization of Vip3Aa via endocytosis. *PLoS Pathog.* **14**, e1007347 (2018).
18. Jiang, K. et al. Fibroblast growth factor receptor, a novel receptor for vegetative insecticidal protein Vip3Aa. *Toxins* **10**, 546 (2018).
19. Hernández-Martínez, P., Gomis-Cebolla, J., Ferré, J. & Escriche, B. Changes in gene expression and apoptotic response in *Spodoptera exigua* larvae exposed to sublethal concentrations of Vip3 insecticidal proteins. *Sci. Rep.* **7**, 16245 (2017).
20. Jiang, K. et al. Vip3Aa induces apoptosis in cultured *Spodoptera frugiperda* (Sf9) cells. *Toxicol.* **120**, 49–56 (2016).
21. Kunthich, T., Surya, W., Promdonkoy, B., Torres, J. & Boonserm, P. Conditions for homogeneous preparation of stable monomeric and oligomeric forms of activated Vip3A toxin from *Bacillus thuringiensis*. *Eur. Biophys. J.* **46**, 257–264 (2017).
22. Palma, L. et al. The Vip3Ag4 insecticidal protoxin from *Bacillus thuringiensis* adopts a tetrameric configuration that is maintained on proteolysis. *Toxins (Basel)*. **9**, 165 (2017).
23. Quan, Y. & Ferré, J. Structural domains of the *Bacillus thuringiensis* Vip3Af protein unraveled by tryptic digestion of alanine mutants. *Toxins* **11**, 368 (2019).
24. Şahin, B., Gomis-Cebolla, J., Güneş, H. & Ferré, J. Characterization of *Bacillus thuringiensis* isolates by their insecticidal activity and their production of Cry and Vip3 proteins. *PLoS One* **13**, e0206813 (2018).
25. Zhang, J. et al. Proteolytic activation of *Bacillus thuringiensis* Vip3Aa protein by *Spodoptera exigua* midgut protease. *Int. J. Biol. Macromol.* **107**, 1220–1226 (2018).
26. Li, J. D., Carroll, J. & Ellar, D. J. Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution. *Nature* **353**, 815–821 (1991).
27. Olson, R. & Gouaux, E. Crystal structure of the vibrio cholerae cytolysin (VCC) pro-toxin and its assembly into a heptameric transmembrane pore. *J. Mol. Biol.* **350**, 997–1016 (2005).
28. Levan, S., De, S. & Olson, R. *Vibrio cholerae* cytolysin recognizes the heptasaccharide core of complex N-glycans with nanomolar affinity. *J. Mol. Biol.* **425**, 944–957 (2013).
29. Jiang, K. et al. Structural insights into the insecticidal Vip3A toxin of *Bacillus thuringiensis*. Preprint at: <https://www.biorxiv.org/content/10.1101/2020.01.24.918433v1> (2020).
30. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480 (1995).
31. Borastom, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781 (2004).
32. Zheng, M., Evdokimov, A. G., Moshiri, F., Lowder, C. & Haas, J. Crystal structure of a Vip3B family insecticidal protein reveals a new fold and a unique tetrameric assembly. *Protein Sci.* **29**, 824–829 (2020).
33. Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* **395**, 347–353 (1998).
34. Zhou, Q. et al. Architecture of the synaptotagmin-SNARE machinery for neuronal exocytosis. *Nature* **525**, 62–67 (2015).
35. Crick, F. H. C. Is  $\alpha$ -keratin a coiled coil? *Nature* **170**, 882–883 (1952).
36. Crick, F. H. C. The packing of  $\alpha$ -helices: simple coiled-coils. *Acta Crystallogr.* **6**, 689–697 (1953).
37. Tarbouriech, N., Curran, J., Ruigrok, R. W. H. & Burmeister, W. P. Tetrameric coiled coil domain of Sendai virus phosphoprotein. *Nat. Struct. Biol.* **7**, 777–781 (2000).
38. Hartmann, M. D. et al. A coiled-coil motif that sequesters ions to the hydrophobic core. *Proc. Natl Acad. Sci. U.S.A.* **106**, 16950–16955 (2009).
39. Banyuls, N., Hernández-Rodríguez, C. S., Van Rie, J. & Ferré, J. Critical amino acids for the insecticidal activity of Vip3Af from *Bacillus thuringiensis*: Inference on structural aspects. *Sci. Rep.* **8**, 7539 (2018).
40. Fang, J. et al. Characterization of chimeric *Bacillus thuringiensis* Vip3 toxins. *Appl. Environ. Microbiol.* **73**, 956–961 (2007).
41. Li, C. et al. *Bacillus thuringiensis* Vip3 mutant proteins: insecticidal activity and trypsin sensitivity. *Biocontrol Sci. Technol.* **17**, 699–708 (2007).
42. Carr, C. M. & Kim, P. S. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**, 823–832 (1993).
43. Colman, P. M. & Lawrence, M. C. The structural biology of type I viral membrane fusion. *Nat. Rev. Mol. Cell Biol.* **4**, 309–319 (2003).
44. Blijleven, J. S., Boonstra, S., Onck, P. R., van der Giessen, E. & van Oijen, A. M. Mechanisms of influenza viral membrane fusion. *Semin. Cell Dev. Biol.* **60**, 78–88 (2016).
45. Hildebrand, P. W., Preissner, R. & Frömmel, C. Structural features of transmembrane helices. *FEBS Lett.* **559**, 145–151 (2004).
46. Hong, M. & DeGrado, W. F. Structural basis for proton conduction and inhibition by the influenza M2 protein. *Protein Sci.* **21**, 1620–1633 (2012).
47. Wang, J., Pielak, R. M., McClintock, M. A. & Chou, J. J. Solution structure and functional analysis of the influenza B proton channel. *Nat. Struct. Mol. Biol.* **16**, 1267–1271 (2009).
48. Cady, S. D. et al. Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* **463**, 689–692 (2010).
49. Joh, N. H. et al. De novo design of a transmembrane Zn<sup>2+</sup>-transporting four-helix bundle. *Science* **346**, 1520–1524 (2014).
50. Mesrati, L. A., Tounsi, S. & Jaoua, S. Characterization of a novel Vip3-type gene from *Bacillus thuringiensis* and evidence of its presence on a large plasmid. *FEMS Microbiol. Lett.* **244**, 353–358 (2005).
51. Luna-Vargas, M. P. A. et al. Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases. *J. Struct. Biol.* **175**, 113–119 (2011).
52. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
53. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
54. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
55. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
56. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
57. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
58. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
59. Kim, D. E., Chivian, D. & Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **32**, W526–W531 (2004).
60. Smart, O. S., Neduelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph* **14**, 354–360 (1996).
61. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
62. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

## Acknowledgements

The authors are grateful to the EM platforms of the CIB Margarita Salas and CNB (CSIC) for support in grid screening and preparation. We also acknowledge Diamond for access and support of the Cryo-EM facilities at the UK national electron bio-imaging center (eBIC proposals EM21062-2 and B124924-1), funded by the Wellcome Trust, MRC, and BBSRC. Access to eBIC was supported by iNEXT, grant numbers 4746 and 8736, funded by the Horizon 2020 program of the European Commission. This work was supported by the Spanish MINECO/AEI, the MCIU and the ERDF ESF (BFU2017-89143-P and RYC-2015-19059 to E.A.-P.; BFU2016-78606-P and RYC-2014-16490 to P.C.; RTI2018-095204-B-C21 to J.F.).



### Author contributions

J.F., P.C., and E.A.-P. designed this study. R.N.-R., J.H., P.C. and E.A.-P. designed the experiments. R.N.-R., J.H., Y.B., P.C., and E.A.-P. performed the experiments, data acquisition, analysis and interpretation. R.N.-R., J.F., P.C., and E.A.-P. wrote the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-17758-5>.

**Correspondence** and requests for materials should be addressed to P.C. or E.A.-P.

**Peer review information** *Nature Communication* thanks Charles Bayly-Jones, Neil Crickmore, and other, anonymous, reviewers for their contributions to the peer review of this work. Peer review reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020