ARTICLE

De novo determination of mosquitocidal Cry11Aa and Cry11Ba structures from naturally-occurring nanocrystals

Guillaume Tetreau[1,18], Michael R. Sawaya[2,18], Elke De Zitter[1,18], Elena A. Andreeva[1,3], Anne-Sophie Banneville[1], Natalie A. Schibrowsky[2,4], Nicolas Coquelle[5], Aaron S. Brewster[6], Marie Luise Grünbein[3], Gabriela Nass Kovacs[3], Mark S. Hunter[7], Marco Kloos[3,8], Raymond G. Sierra[7], Giorgio Schiro[1], Pei Qiao[9], Myriam Stricker[3], Dennis Bideshi[10,11], Iris D. Young[6], Ninon Zala[1], Sylvain Engilberge[1], Alexander Gorel[3], Luca Signor[1], Jean-Marie Teulon[1], Mario Hilpert[3], Lutz Foucar[3], Johan Bielecki[8], Richard Bean[8], Raphael de Wijn[8], Tokushi Sato[8], Henry Kirkwood[8], Romain Letrun[8], Alexander Batyuk[7], Irina Snigireva[12], Daphna Fenel[1], Robin Schubert[8], Ethan J. Canfield[13], Mario M. Alba[14], Frédéric Laporte[15], Laurence Després[15], Maria Bacia[1], Amandine Roux[16], Christian Chapelle[17], François Riobé[16], Olivier Maury[16], Wai Li Ling[1], Sébastien Boutet[7], Adrian Mancuso[8], Irina Gutsche[1], Eric Girard[1], Thomas R. M. Barends[3], Jean-Luc Pellequer[1], Hyun-Woo Park[10,11], Arthur D. Laganowsky[9], Jose Rodriguez[2,4], Manfred Burghammer[12], Robert L. Shoeman[3], R. Bruce Doak[3], Martin Weik[1], Nicholas K. Sauter[6], Brian Federici[10], Duilio Cascio[2], Ilme Schlichting[3] & Jacques-Philippe Colletier[1✉]

Cry11Aa and Cry11Ba are the two most potent toxins produced by mosquitocidal *Bacillus thuringiensis* subsp. *israelensis* and *jegathesan*, respectively. The toxins naturally crystallize within the host; however, the crystals are too small for structure determination at synchrotron sources. Therefore, we applied serial femtosecond crystallography at X-ray free electron lasers to in vivo-grown nanocrystals of these toxins. The structure of Cry11Aa was determined *de novo* using the single-wavelength anomalous dispersion method, which in turn enabled the determination of the Cry11Ba structure by molecular replacement. The two structures reveal a new pattern for in vivo crystallization of Cry toxins, whereby each of their three domains packs with a symmetrically identical domain, and a cleavable crystal packing motif is located within the protoxin rather than at the termini. The diversity of in vivo crystallization patterns suggests explanations for their varied levels of toxicity and rational approaches to improve these toxins for mosquito control.

The most commonly used biological insecticides for controlling mosquito and black fly vector populations are produced by the bacterium *Bacillus thuringiensis* subsp. *israelensis* (*Bti*), discovered in Israel in 1976[1]. These products target the larval stage of a wide variety of vectors, and due to their high efficacy and environmental safety, have replaced broad spectrum synthetic chemical insecticides in many vector control programs. These include Anopheles gambiae and related species that transmit malaria, as well as numerous Culex and Aedes species that spread viruses such as those that cause West Nile Encephalitis and Yellow Fever. *Bti* products are also used in Africa to regulate black fly species responsible for vectoring the filarial worms that cause River Blindness. Aside from vector populations, they are used to control nuisance mosquitoes in the Rhine Valley in Germany, in the Camargue in southern France, and throughout the U.S., Asia, and Latin and South America, with thousands of tons applied over the past 30 years.

The highly potent mosquitocidal activity of *Bti* is due to three nanocrystalline forms of four protoxins, viz. Cyt1Aa, Cry11Aa, and co-crystallized Cry4Aa and Cry4Ba. These are produced during sporulation and are remarkably stable in a variety of conditions, but dissolve after ingestion under the high alkaline pH levels characteristic of the larval mosquito midgut[2]. Solubilized protoxins are activated by insect gut proteases enabling binding to gut cell membranes, subsequent oligomerization, and ultimately gut cell lysis leading to larval death[2]. *Bti* toxins are environmentally safe because they are much more specific for target mosquitoes than broad-spectrum chemical larvicides.

The most potent of the four *Bti* toxins is Cry11Aa, but its activation and mechanism of toxicity are poorly understood, in large part because unlike Cry4Aa, Cry4Ba, and Cyt1Aa, its structure is unknown. A related toxin produced by *Bt* subsp. *jegathesan* (*Btj*) is Cry11Ba, which is seven to thirty-seven times more toxic than Cry11Aa against major mosquito vector species belonging to the genera *Aedes*, *Anopheles*, and *Culex*[3], and in some bacterial hosts appears to form slightly larger crystals. Cry11Ba's structure is also unknown, although it has been used as a replacement for Cry11Aa in recombinant strains of *Bti* to improve mosquitocidal activity significantly[3,4]. Thus, our goal was to determine the structures of Cry11Aa and Cry11Ba protoxins to help understand how they achieve formation of robust crystals labile only at alkaline pH, and to obtain structural insights for increasing the efficacy of these proteins for mosquito control.

Structure determination of Cry11Aa and Cry11Ba protoxins from natural nanocrystals requires cutting-edge technology. Conventional crystallography is limited to projects in which crystals are sufficiently large to mount and oscillate individually in a synchrotron X-ray beam. In the past, crystals of activated Cry4Aa[5], Cry4Ba[6] and Cyt1Aa[7] attained sufficient size by growing these in vitro from toxins dissolved from natural nanocrystals and activating the toxins enzymatically. However, Cry11Aa and Cry11Ba do not recrystallize in vitro from dissolved nanocrystals[8]. Moreover, enzymatic activation is unwanted since our goal is to understand the pH-controlled mechanism of natural crystal dissolution. To observe the protoxin state in natural nanocrystals produced in bacterial cells, we applied serial femtosecond crystallography (SFX) at X-ray free electron lasers (XFEL)[9–11]. In the SFX experiment, high brilliance XFEL beam pulses, each lasting only ~10–50 fs, intercept a series of nanocrystals, one pulse-per-crystal, eliciting the strongest possible diffraction signal from each tiny crystal before it vaporizes, and producing a series of diffraction snapshots, later assembled into a full data set. Feasibility of this strategy had been demonstrated by the recent elucidation of the full bioactivation cascade of Cyt1Aa[12].

Our success in determining the structures of Cry11Aa and Cry11Ba protoxins highlights the capability of XFEL sources to overcome limits of small crystal size. We relied on de novo phasing of the native SFX data because all attempts at molecular replacement (MR) failed despite detectable sequence similarity with thirteen structurally-determined members of the three-domain Cry δ-endotoxin family (Fig. 1)[13–15]. We opted to derivatize our Cry11Aa nanocrystals with a recently-introduced phasing-agent, a caged-terbium compound, Tb-Xo4[16,17]. The phases obtained from single-wavelength anomalous dispersion (SAD) were sufficient to reveal the Cry11Aa protoxin structure at 2.6 Å resolution and subsequently enable phasing of the Cry11Ba protoxin structure at 2.4 Å resolution by molecular replacement. In hindsight, we attribute the failure of early MR attempts to three extra β-strands in domain II which alter the relative orientation of the three domains in Cry11 toxins.

Our studies of Cry11Aa and Cry11Ba crystals reveal a new paradigm of molecular packing among Cry δ-endotoxins reported thus far. In particular, the cleavable peptides that constitute important crystal contacts are located near the middle of the toxin sequence, rather than at the termini. Molecules pack in tetramer units, exhibiting D2 symmetry; these tetramers in turn pack in a body centered pattern (like a 3-dimensional brick-wall in which successive rows are offset by half a brick). To achieve this pattern, each of the three domains in a Cry11 molecule packs with an identical domain from a symmetry related molecule: domain I packs with domain I, II with II, and III with III. Thus, each Cry11 domain fulfills two biological roles: a dimer interface manifested in the crystalline state, and a functional role manifested in the soluble state: target recognition (domain II), oligomerization (domain III) and pore formation (domain I)[18]. Differences in the size and composition of the three packing interfaces explain shape and size differences between Cry11Aa and Cry11Ba nanocrystals. Structure-guided site-directed mutagenesis verifies which residues affect crystal size, pH sensitivity of the crystal, and toxin folding. Our results elucidate the Cry11Aa and Cry11Ba bioactivation
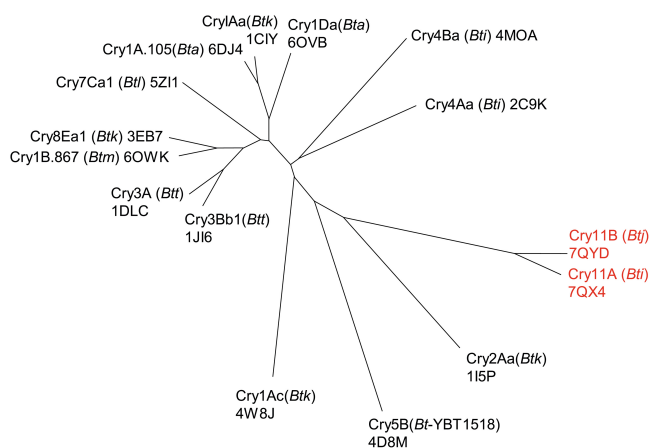

Fig. 1 Phylogentic tree of delta-endotoxins. Tree plot showing structural relatedness of Cry11Aa and Cry11Ba to the family of 13 delta-endotoxins reported in the PDB. Structural similarity is represented on a phylogenetic tree plot. Shorter connecting lines signify closer structural similarity. Cry11Aa and Cry11Ba are structurally similar to each other, but distant from the other 13 delta-endotoxins. The closest structural homolog of Cry11Aa and Cry11Ba is Cry2Aa. The PDB ID codes are reported next to the toxin name. Species of origin are annotated in parentheses. *Bta, Bti, Btj, Btk, Btm* and *Btt* abbreviate insectidal *Bacillus thuringiensis* subspecies *aizawai, israelensis, jegathesan, kurstaki, medellin* and *tenebrionis*, respectively. *Bt*-YBT1518 is a strain displaying nematocidal activity.
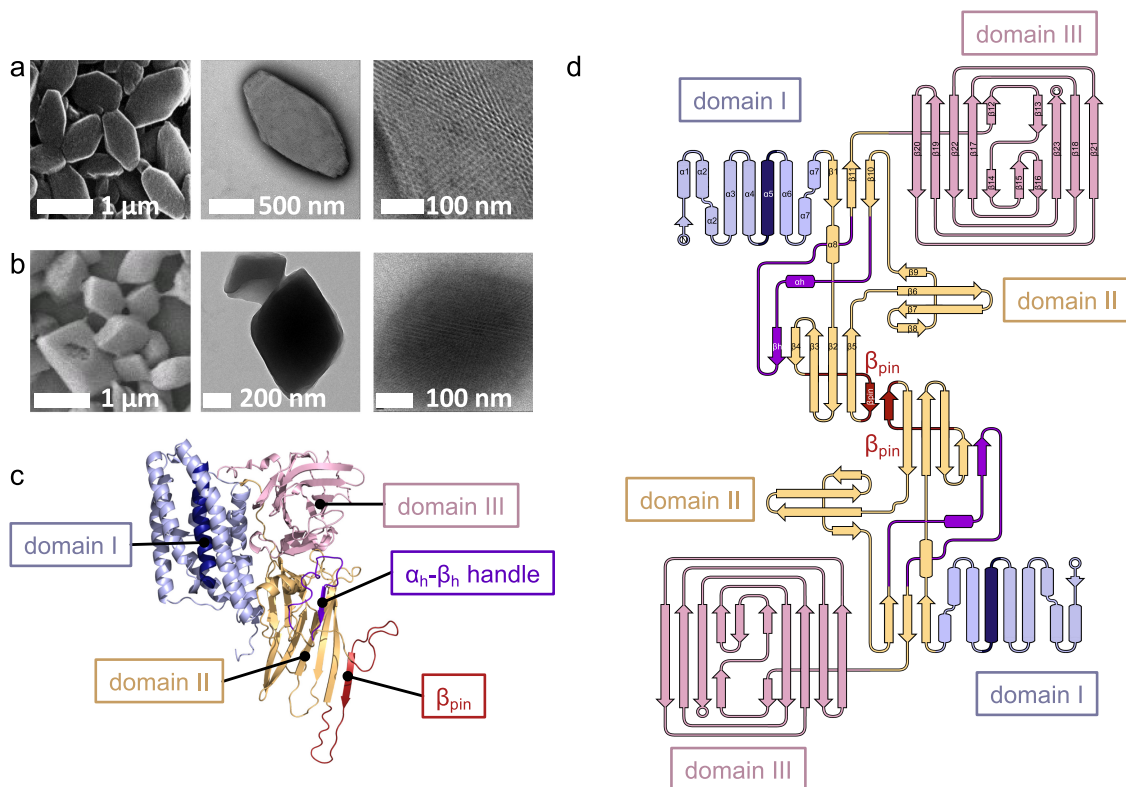
Fig. 2 Crystals and overall fold of Cry11 toxins. a, **b**, Scanning (left; SEM) and transmission (middle, right; TEM) electron micrographs of gold plated and negatively-stained Cry11Aa (**a**) and Cry11Ba (**b**) crystals, respectively. The right panels show a close-up view of the crystal surface. SEM and TEM experiments on Cry11Aa were repeated 7 and 2 times, respectively, while those on Cry11Ba were repeated 3 and 16 times, respectively. **c** Cry11Aa crystal structure, depicted as cartoon. Domain I is shown in light blue, except for central helix α5 which is shown in dark blue; domain II is shown in orange except for the $\alpha_h\beta_h$-handle and β_{pin} which are shown in purple and red, respectively; domain III is shown in pink. **d** Topology diagram of a Cry11Aa dimer with similar color code as in (**c**) and with labeling of secondary structure elements in one of the two monomers. The two monomers in a dimer assemble via the β_{pin}, resulting in the formation of a large β-sheet. The short helices α9, α10 and α11, respectively located in the β11-β12, β15-β16 and β21-β22 loops, are not shown in the diagram.

cascade and enable development of new, rational strategies for improved mosquito control.

Results

De novo phasing of Cry11Aa and Cry11Ba structures by SFX. In vivo-grown crystals of Cry11Aa and Cry11Ba protoxins exhibit distinct morphologies, which initially concealed a surprising conservation of their crystal packing patterns. Cry11Aa crystallizes as hexagonal plates and Cry11Ba crystallizes as larger bipyramidal crystals (Fig. 2a-b) as reported earlier[4]. These morphological distinctions cannot be attributed to differences in crystallization mechanisms in their parent organisms, *Bti* and *Btj*, since both protoxins were recombinantly produced in the same host organism, an acrystalliferous strain of *Bti* (4Q7). Cry11Aa and Cry11Ba protoxins are expected to share structural resemblance to each other since the two sequences share 54% identity; however, 46% non-identity at the molecular level could easily produce large differences at the macroscopic level of crystal morphology. Moreover, the sequence of Cry11Ba is extended by 77-residues at its C-terminus, potentially also affecting differences in crystal packing (Supplementary Fig. 1). Interestingly, this extension has been identified as a low complexity region (LCR) by both CAST[19] and SEG[20] computational methods, which implicates the extension in the mechanism of crystal nucleation. At this point in our studies, the balance of evidence suggested that sequence divergence was likely to have erased the crystal packing pattern that early ancestors of today's Cry11Aa and Cry11Ba presumably once shared.

Our diffraction experiments yielded the first hint that Cry11Aa and Cry11Ba shared a conserved crystal packing pattern. We collected diffraction data from Cry11Aa and Cry11Ba nanocrystals injected in the vacuum chamber of the CXI-SC3 microfocused beamline at the Stanford Linear Accelerator Center (SLAC) Linac Coherent Light Source (LCLS)[21] using a microfluidic electrokinetic sample holder (MESH)[22] (Cry11Ba crystals) or a gas-dynamic virtual nozzle (GDVN)[23] (Cry11Aa crystals). The underlying similarity in the packing of Cry11Aa and Cry11Ba became evident when their diffraction patterns were indexed, revealing similarly sized unit cells (a ~ 58; b ~ 155; c~171 Å; α = β = γ=90°), albeit belonging to two different space groups: $I222$ and $P2_12_12$, respectively (Table 1). Conservation of unit cell parameters hinted that this crystal packing pattern is special, evolved to perform a function more intricate than just storing protein.

To gain further insight into Cry11Aa and Cry11Ba crystal packing, we depended on de novo methods to solve the crystallographic phase problem. Initial attempts to acquire phases from homologous structures by molecular replacement (MR) failed, suggesting Cry11Aa and Cry11Ba contained novel features, not present in the PDB. Our search models included structures of Cry δ-endotoxins homologs (exhibiting up to 26% sequence identity to our two targets) and homology models produced using Robetta[24] (http://robetta.bakerlab.org/) and SwissProt[25] (https://www.ebi.ac.uk/uniprot/). After MR failed, we turned to de novo phasing methods. We soaked Cry11 nanocrystals with conventional heavy atom derivatives including gadolinium (Gd),

Table 1 Data collection and refinement statistics of Cry11Aa and Cry11Ba.

	Cry11Aa pH 7	Cry11Aa-TBXO4 pH7	Cry11Ba pH 6.5	Cry11Ba pH 10.4
PDB ID	7QX4		7QYD	7R1E
Data collection				
Space group	$I\ 2\ 2\ 2$	$I\ 2\ 2\ 2$	$P\ 2_1\ 2_1\ 2$	$P\ 2_1\ 2_1\ 2$
Cell dimensions (Å)	57.64 ± 0.19	57.64 ± 0.15	168.18 ± 0.19	167.50 ± 0.29
	155.69 ± 0.80	156.29 ± 0.73	158.45 ± 0.26	157.99 ± 0.47
	171.14 ± 0.54	170.75 ± 0.40	57.51 ± 0.08	57.43 ± 0.14
Wavelength (Å)	1.27	1.27	1.30	1.30
X-ray beam focus (µm)	5	5	1	1
No. collected frames	792623	558747	813133	990643
No. indexed frames	48652	77373	19708	15689
No. merged crystals	50613	88511	19708	15689
Resolution range (Å)	33.55–2.60	33.51–2.55	42.06–2.40	35.72–2.65
	(2.66–2.60)	(2.61–2.55)	(2.44–2.40)	(2.70–2.65)
No. observations	8253629 (365007)	14069217 (640046)	3541082 (51048)	3482475 (67984)
No. unique reflections	24198 (1583)	48634 (3297)	61141 (2980)	45243 (2204)
$\langle I/\sigma\ (I)\rangle$	9.50 (1.16)	11.23 (1.62)	4.73 (0.90)	3.98 (1.02)
R_{split} (%)	10.73 (95.40)	7.97 (70.58)	14.5 (84.9)	22.4 (84.0)
$CC_{1/2}$ (%)	99.3 (37.7)	99.6 (67.5)	99.1 (35.9)	98.5 (15.2)
Completeness (%)	99.9 (100.0)	100.0 (100.0)	100.0 (100.0)	99.4 (100.0)
Multiplicity	341.09 (230.58)	289.29 (194.13)	57.9 (17.1)	77.0 (30.9)
Anomalous data				
Completeness (%)		100.0 (100.0)		
CCano		0.26 (0.00)		
CRDano		1.35 (1.01)		
Refinement				
Resolution range (Å)	33.55–2.60		42.10–2.40	35.72–2.65
No. reflections	24196		55060	45228
R_{work}/R_{free}[a]	17.2/24.1		18.7/23.1	23.7/24.7
No. atoms				
Protein	5080		10083	9900
Water	261		623	119
B-factors (Å^2)				
Main chain	50.47		49.0/48.7[b]	46.4/43.7
Side chain	51.44		53.3/53.1	47.0/44.8
Water	46.17		48.4	36.5
R.m.s.d.				
Bonds lengths (Å)	0.004		0.008	0.001
Bonds angles (°)	0.633		1.324	0.408

[a]R_{free} is calculated using 5 or 10% (Cry11Aa or Cry11Ba, respectively) of random reflections excluded from refinement.
[b]Average B-factor for chain A / chain B.

gold (Au), platinum (Pt), and mercury (Hg) salts, but they failed to produce interpretable isomorphous or anomalous difference Patterson peaks. Finally, a recently introduced caged-terbium compound[16,17], Tb-Xo4, produced a successful derivative of Cry11Aa (after a 30 h soak at 10 mM concentration), and phases were determined by the single-wavelength anomalous dispersion (SAD) method at 2.55 Å resolution (using anomalous signal up to 3.5 Å). Two Tb-Xo4 molecules were identified bound to the single Cry11Aa molecule in the asymmetric unit (isomorphous peaks at 23 and 9 σ, and anomalous peaks at 33 and 8.1 σ, respectively; Supplementary Fig. 2a). The success of Tb-Xo4 can be partly ascribed to the dramatically high anomalous dispersion signal (i.e., f' and f") of terbium, but likely also stems from stronger binding of Tb-Xo4 to the protein owing to presence of an organic cage; indeed, f' and f" of Gd and Tb are similar at the X-ray energy used for data collection (9 keV). Regardless, phases were of sufficient quality to reveal all Cry11Aa residues from N13 to the C-terminal K643.

The Cry11Ba structure was thereafter phased successfully by MR using the Cry11Aa structure as a search model, revealing *a posteriori* that the Gd, Pt and Au ions had successfully bound to the crystalline Cry11Ba, despite anomalous and isomorphous signals being too weak to enable phasing (Supplementary Fig. 2b-c

and Methods section). Our MR-phase 2.4 Å resolution map reveals two Cry11Ba molecules in the asymmetric unit. All residues are visible except for the N-terminus (residues M1-I11), two loops (residues S332-C335, and G354-S359) and the C-terminal extension (residues S659-K724). The lack of order in this extension is not surprising given the low complexity of its sequence.

Cry11 domain organization is similar to δ-endotoxins, but exhibits some non-canonical features. Cry11Aa and Cry11Ba structures maintain the three-domain organization characteristic of Cry δ-endotoxins[13,26] (Fig. 2c and Supplementary Fig. 3). Domain I is involved in formation of a pore in the target membrane. Like in other Cry δ-endotoxins, it forms a seven-α-helix bundle; at the center of the bundle is α5 (residues 146–170), surrounded by the remaining six helices. Domain II is involved in the recognition of mosquito-specific receptors. It forms a β-prism composed of three-β-sheets, wherein the first two β-sheets (β4-β3-β2-β5 and β8-β7-β6-β9) each adopts a Greek-key topology while the third β-sheet is three-stranded (β1-β10-β11). Domain III is involved in oligomerization. It forms a β-sandwich of two antiparallel five-stranded β-sheets (viz. β20 − β19 − β22 − β17 − β12/β14 and β15 − β13/β16 − β23 − β18 − β21) forming a

jelly-roll topology, whereby β12/β14 and β13/β16 are interrupted β-strands contributed by two non-consecutive shorts β-strands, which appose and intercalate one after the other onto β17 and between β15 and β23, respectively (Fig. 2d).

The closest homolog of known structure to Cry11 toxins is *Bt kurstaki* (*Btk*) Cry2Aa (PDBid: 1i5p), with a sequence identity of 26.6 and 23.6 % and main-chain r.m.s.d. of 3.7 and 4.0 Å, with respect to Cry11Aa and Cry11Ba, respectively (Fig. 1). As with Cry2Aa, the Cry11 toxins feature a long insert (27 residues in Cry2Aa; 21 residues in the Cry11 toxins) between strands β10 and β11, which together with domain-I β1, form the third β-sheet of the domain-II β-prism. This insert, which features a short α-helix (α_h) and a β-strand (β_h), folds like a handle, and is therefore referred to as the α_hβ_h-handle, throughout the manuscript (Fig. 2c, Supplementary Fig. 3). The α_hβ_h-handle fastens domain II onto domain III through direct (e.g., in Cry11Aa, D443(OD2)-R502(NH2); D443(O)-R502(NH1); L447(N)-S503(O)) and water mediated H-bonds (T446(OG1)/T448(O)-Wat869(O)-R502(N); T448(OG1)/V499(O)-Wat744(O)-D501(OD1); T448(N)/L447(N)-Wat774(O)-S503(OG)/(O))) (Supplementary Fig. 4), and enables the burying of domain-II α8 at an interface formed by α_hβ_h, α6-α7 (domain I), β10-β11 (domain II), β15 and the β13-β14 and β15-β16 loops (domain III), and the α9 helix connecting domain II and domain III (D469-K478 in Cry11Aa). The firm hold of α8 enables the three domains to be more tightly packed in Cry2Aa and Cry11 toxins than in other Cry toxins (e.g., *Bt tenebrionis* (*Btt*) Cry3Aa or *Btk* Cry1Ac). Additionally, strand β_h lays aside strand β4 thereby expanding – and consequently, stabilizing—the first β-sheet of domain II (β_h-β4-β3-β2-β5). Also, alike Cry2Aa, the Cry11 toxins feature a smaller β-prism due to deletions in the second constitutive β-sheet, namely between β7 and β8 (6 and 10 residues missing in Cry2Aa and Cry11 toxins, respectively), and between β9 and β10 (14 and 15 residues missing in Cry2Aa and Cry11 toxins, respectively; Supplementary Fig. 3).

The Cry11 toxin structures are distinguished by a 36 to 38 residue insertion that is observed between strands β4 and β5. The insertion appends a β-strand at the edge of the first β-sheet of domain II – hereafter referred to as the β_pin (Fig. 2c). The β_pin forms the center of a two-fold symmetric dimer interface with the β_pin of another toxin molecule. The interface features approximately twelve backbone hydrogen bonds, merging two β-sheets into a large, antiparallel, intermolecular β-sheet (β_h-β4-β3-β2-β_pin – β_pin-β2-β3-β4-β_h) which fastens chain A to C and B to D (interface #3, see below) and stabilizes the tetramer (Fig. 3b, e). We note that the buried surface area (BSA) at the tetramerization interface is 33% lower in Cry11Ba, pointing to higher flexibility; this hypothesis is supported by the absence of interpretable electron density for residues at the N-terminus (332–335) and C-terminus (354–359) of the β_pin in the Cry11Ba structure. Also noteworthy is that Cry11 toxins feature a conserved N/D-DDLGITT insertion between β21 and β22, and deletions (>3 residues) between α3 and α4 (−5 and −8 residues with respect to *Btk* Cry2Aa and *Btt* Cry3Aa), and β20 and β21 (−10 and −9 residues with respect to *Btk* Cry2Aa and *Btt* Cry3Aa). Altogether, these changes render Cry11 toxins uniquely large from the structural standpoint, with predicted radii of gyration of 27.5 and 26.7 Å for Cry11Aa and Cry11Ba, compared to 25.0 and 25.6 Å for *Btk* Cry2Aa and *Btt* Cry3Aa, respectively.

All domains engage in producing the in vivo crystal lattice. Examination of packing interfaces reveal that all three domains are involved in the formation and stabilization of Cry11Aa and Cry11Ba nanocrystals. The in vivo crystallization pathway can be best trailed from Cry11Aa crystals, which feature a single monomer per asymmetric unit and build on six packing

interfaces burying a cumulated surface area (BSA) of 3515 Å², corresponding to 13.1 % of the total protein surface area. The main building block of Cry11Aa crystals consists of a tetramer with a total BSA of 9663 Å² and a predicted binding energy of −12.5 kcal.mol⁻¹ at pH 7 by PISA[27] (Fig. 3a-b).

The tetramer comprises two principle dimers: dimer A-B, and dimer C-D. A two-fold symmetry axis (vertical in Fig. 3b) relates chain A to B and chain C to D. At the dimer interface domain II contacts domain II (α_hβ_h-handle residues P433-P457 and strand β4) and domain III contacts domain III (interface #1; Fig. 3b). Perpendicular to this axis, another two-fold symmetry axis relates dimer A-B to dimer C-D creating A-C and B-D interfaces. These interfaces involve β_pin strands in domains II as mentioned in the previous paragraph (interface #3; Fig. 3b). The tetramer is further stabilized by minor contacts between apices of domain II (interfaces AD or BC; interface #6). Crystals grow by packing such tetramers in a brick-wall fashion via face-to-back contacts between domains I (interface #2; Fig. 3c). Cry11Aa crystals are further cemented by two additional minor interfaces. The first involves the apex of the second β-sheet of domain II (interface #5) from monomers in each dimer of the tetramer (AD or BC). The second occurs between the α3-α4 loop of domain I in one tetramer and the apex of the second β-sheet of domain II in another tetramer (interface #4).

Cry11Ba crystals assemble from tetrameric building blocks analogous to those in Cry11Aa crystals, as judged by the similarity of their crystal packing patterns (Fig. 3d, e, f). However, the tetramer in Cry11Ba is not as readily identified as an autonomous unit by PISA as it was in Cry11Aa crystals. Our measurements of BSA in the crystal packing interfaces suggest an explanation. We assign crystal packing forces to two types: those that associate monomers into dimers and tetramers and those that assemble tetramers into a crystal. The BSA which associates monomers into a dimer (interface #1) is 38% lower than the homologous interface in Cry11Aa (1009 Å² Fig. 2e; vs. 1631 Å²; Fig. 3g), explaining PISA's failure to identify the tetramer. However, the packing of tetramers into crystals is 53% higher in Cry11Ba than in Cry11Aa (1429 Å² at interface #2; Fig. 3f vs. 934 Å²; Fig. 3g). Thus, the relative contributions of the two types of crystal packing interfaces differ between Cry11Aa and Cry11Ba, but the sum of the contributions is nearly the same. Cry11Aa exhibits only slightly more BSA per monomer (3515 Å²) than Cry11Ba (3385 Å²), corresponding to 13.1% of the total protein surface area of Cry11Aa and 12.6 % of Cry11Ba. The mechanism by which Cry11Ba evolved stronger tetramer-tetramer interfaces could have exploited the extra degrees of freedom afforded by having two molecules in the asymmetric unit rather than just one as does Cry11Aa. Moreover, the emphasis on lattice-forming associations (rather than associations within a tetramer unit) could explain the larger crystal size achieved by Cry11Ba. Regardless, the Cry11 toxins structures shows that each domain functions to assemble and stabilize in vivo-grown nanocrystals. These functions must have evolved alongside domain specific functions: pore formation (domain I), receptor-recognition and membrane-insertion (domain II), and oligomerization and stabilization of the toxic pore conformation (domain III)[26].

Drastic conformational changes drive crystal dissolution. We sought to characterize the conformational changes that ensue pH elevation, preceding dissolution of the crystals in the mosquito larvae gut[28]. As the crystals are naturally labile at pH 11, we aimed at collecting data from crystals soaked at slightly lower pH than 11, hypothesizing that early conformational changes would show but the crystal packing still hold. In the case of Cry11Aa crystals, diffraction quality was decreased
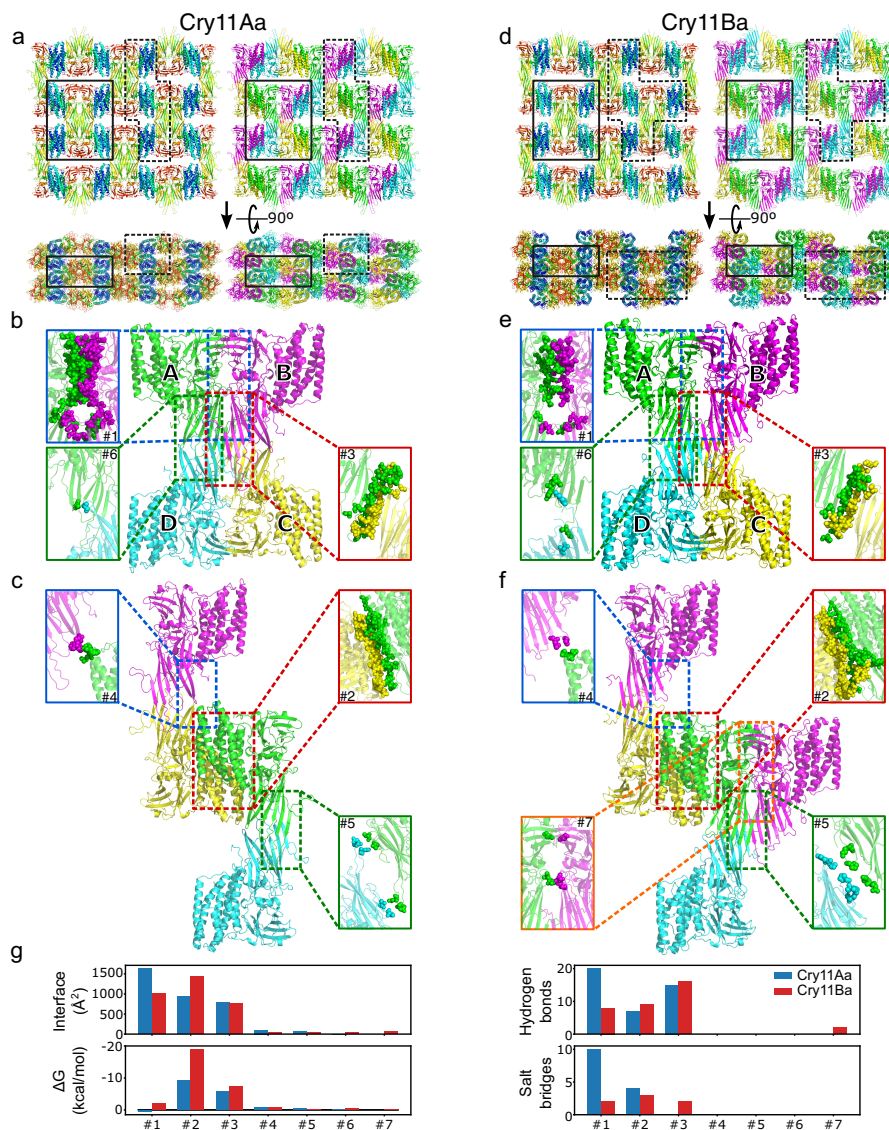
Fig. 3 Monomer interactions in Cry11Aa and Cry11Ba. a Cry11Aa crystal packing, colored according to sequence (from blue to red) indicating the domain-based assembly; and colored according to tetramer assembly (see panel (**b**)). The highlighted areas indicate the regions shown in (**b**) (full line) and (**c**) (dashed line). **b** Cry11Aa tetramer with zoom on each of the three interfaces identified by PISA (interface #1, #3 and #6), with the involved residues depicted as spheres. Supplementary Fig. 5 shows the residues involved in hydrogen bonds and salt bridges. **c** Cry11Aa crystal assembly by interactions between neighboring tetramers, formed by interface #2, #4 and #5, visualized as in (**b**). **d**, Cry11Ba crystal packing, colored as in (**a**). **e** Cry11Ba tetramer with zoom on the interfaces as in (**b**). **f** Cry11Ba crystal assembly, visualized as in (**c**). As compared to Cry11Aa, Cry11Ba crystals contain an additional interface #7 between an A-B pair from two neighboring tetramers. **g** Interface statistics as identified by PISA for Cry11Aa (blue) and Cry11Ba (red).

dramatically at pH values of 9.5 (CAPS buffer, glycerol 30%) and above, preventing collection of a sufficiently large number of diffraction patterns to produce a high-pH dataset. Hence, large conformational changes occur in Cry11Aa at pH as low as 9.5, opposing diffraction quality, despite crystals dissolving as of pH 11 only (Fig. 4a). In the case of Cry11Ba, 2.65 Å diffraction was preserved up to pH 10.4 (Table 1). Comparison between the refined 'pH10.4' and "pH6.5" structures points to large inter-domain rearrangements induced by pH increase. Detailed analysis of structural changes at the side chain level was yet prevented by the non-isomorphism of the "pH6.5" and "pH10.4" datasets. A 1 % unit-cell contraction, and hence tighter crystal packing, was observed in the "pH10.4" crystals in comparison to the "pH6.5" crystals. However, because a higher glycerol concentration was used for injection of Cry11Ba crystals at pH 10.4, we cannot exclude that unit cell contraction might be caused by crystal dehydration.

Crystals are made of full-sized monomers of Cry11 protoxins. In both Cry11Aa and Cry11Ba toxins, the β_{pin} (residues E339-Q350 and I341-Y350, respectively) is a ~10-residue long β-strand that hydrogen-bonds with a twofold related symmetry mate, contributing the interface that assembles dimers (AC and BD) into tetramers. This strand is bordered on each side by the only two loops that have disordered electron density in Cry11Ba (missing residues S332-C335 and G354-S359) and are comparatively difficult to interpret in Cry11Aa (F330-D334 and Q350-E355), respectively. As Cry11Aa N335-Y349 and Cry11Ba I341-N351 regions match the enzymatic cleavage site known to generate the two activated fragments of ~32 and ~36 kDa[29,30] upon proteolytic activation in the mosquito larvae gut, we asked whether disorder in the F330-D334 (G330-E340) and Q350-E355 (D352-I358) loops serves the purpose of enabling facilitated access of proteases to Cry11Aa (Cry11Ba) cleavage sites or if each monomer occurs in natural crystals as two polypeptide chains
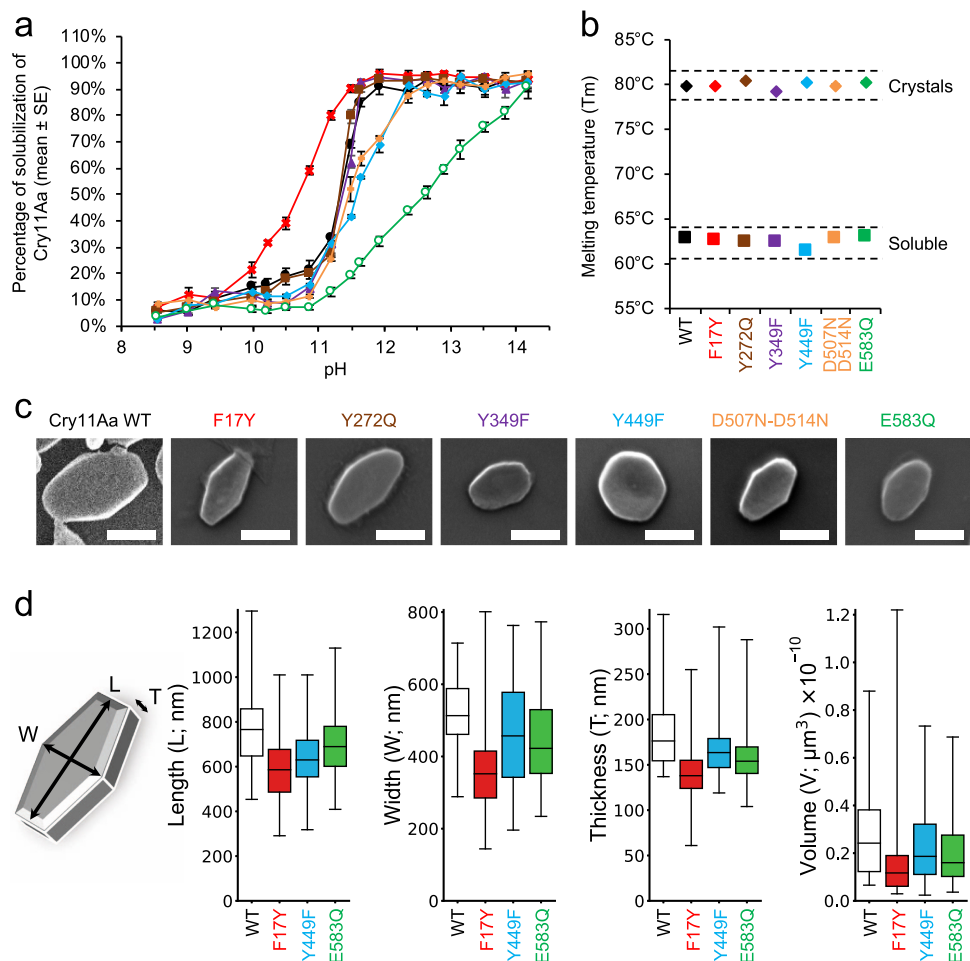
Fig. 4 Point-mutations of Cry11Aa affect the shape, size and pH-sensitivity of in vivo-grown nanocrystals. a Crystals from mutants exhibit similar sigmoidal patterns of crystal solubilization as a function of pH, except F17Y and E583Q that are more and less sensitive to pH, respectively ($n = 3$ independent measurements, data are presented as mean values ± SEM; Cry11Aa WT: black circles, F17Y: red crosses, Y272Q: brown squares, Y349F: purple triangles, Y449F: blue diamonds, D507N-D514N: orange plus symbols and E583Q: empty green circles). **b** Cry11Aa WT and mutants exhibit similar heat stability. As expected, toxins are more stable (+17.5 ± 0.3 °C) in their crystalline (diamonds) than soluble form (squares), irrespective of the mutation. The dashed lines are a guide to the eye. **c** Visualization of a representative crystal for Cry11Aa WT and mutants F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q by SEM (scale bar = 500 nm). **d** Crystals of F17Y ($n = 93$ crystals), Y449F ($n = 60$ crystals) and E583Q ($n = 94$ crystals) imaged by AFM were all smaller in length (L), width (W), thickness (T) and volume than WT ($n = 45$ crystals) highlighting a perturbation of the intrinsic crystal organization induced by these mutations. In each graph, the boxes represent the lower and upper quartiles around the median. The whiskers indicate the minimum and maximum values.

cleaved prior or during crystal formation. SDS-PAGE analysis of Cry11Aa (12% gels, heating at 95 °C for 5 min, presence of DTT and SDS; Supplementary Fig. 6) resulted in a major band ~70 kDa, in line with previous reports[31–33]. As the denaturing treatment would have broken any disulfide-bridge or non-covalent interactions that could maintain cleaved fragments together, this result suggests that Cry11Aa occurs in crystals as a full monomer. We further verified this hypothesis by use of MALDI TOF mass spectrometry. In MALDI mass spectra collected after direct solubilization of the natural crystals in sinapinic acid matrix in presence or absence of DTT, we observed main peaks at m/z of 72246 and 72235 (mass error: ± 100 Da) and 36154 and 36129 Da, respectively, in agreement with expected molecular masses for singly- and doubly- charged ions of a full-size monomer (expected mass: 72.349 kDa) (Uniprot accession number: P21256; Supplementary Fig. 7). However, because proteolytic activation is as well expected to yield a 36 kDa fragment, in addition to a 32 kDa fragment for which a minor peak was present in the MALDI-TOF mass spectra, we resorted to native mass spectrometry to assert that the ~72.240 and ~36.140 kDa

peaks originated from the same species – rather than being indicative of the crystallization of proteolytic products. With this approach, we could confirm that upon dissolution of Cry11Aa crystals, a 72.345 kDa fragment is released, corresponding to the full-size monomer (Supplementary Fig. 8a). Moreover, both incubation of solubilized toxin at room temperature (RT) for 2 h (Supplementary Fig. 8b) and use of increased collision energy (Supplementary Fig. 8c-d) failed at yielding a signature for the two polypeptides that would have been generated if cleavage at position 329 had occurred. We conclude that natural crystals of Cry11Aa, and possibly Cry11Ba (Supplementary Figs. 6b and 7c), grow from the addition of full-size monomers, and that disorder in the F330-D334 (G330-E340) and Q350-E355 (D352-I358) loops could serve the purpose of enabling facilitated access of proteases to Cry11Aa (Cry11Ba) cleavage sites. Considering proteinase K as a surrogate analogue for mosquito larvae gut proteases[34], one would expect the βpin to be released upon proteolytic activation, suggesting that the role of the latter is to promote in vivo crystallization, and that its cleavage would ensure irreversibility of crystal dissolution (Supplementary Fig. 3). We

note that other cleavage sites are predicted, which would release the first six residues and last two β-strands (β22-β23), as well as rescind the covalent association between domain I and domains II and III, thereby leaving non-covalent interactions surfaces as the sole glue between them.

Mutagenesis to alter crystal formation and dissolution. We proposed earlier that the packing of Cry11Ba into slightly larger crystals than Cry11Aa could stem from differences in the extent and nature of the interfaces which support dimerization, tetramerization and packing of tetramers into crystals (Fig. 3). Considering recent evidence linking LCR regions with diverse functions including chaperoning[35] and reversible oligomerization, we further asked whether or not presence of the 77-residue LCR region at the C-terminus of Cry11Ba plays a complementary role in the promotion of crystal formation. A chimera was therefore designed, coined C11AB, wherein the LCR of Cry11Ba was fused to the C-terminal end of Cry11Aa (Methods section; Supplementary Fig. 9a). C11AB was produced at the expected size but at a lower yield than Cry11Aa WT (Supplementary Fig. 9b). Atomic force micrographs (AFM) revealed the presence of multiple needle-like inclusions in the parasporal envelope encasing the crystals, suggesting that presence of Cry11Ba-LCR at the C-terminal end of Cry11Aa favors nucleation, but not crystal growth (Supplementary Fig. 9c).

Seven Cry11Aa mutants and four Cry11Ba were additionally designed with the aim to probe the involvement of intra- and inter-molecular interfaces in toxin stability, crystal formation and dissolution. Each mutant was designed to challenge a specific interface and served as a coarse proxy to evaluate its pH sensitivity and putative participation in the crystal dissolution mechanism. First, we asked whether the intra-chain stabilization of α8 at an interface contributed by the three domains (namely, $\alpha_h\beta_h$, α6-α7, α9, β10, β11, β15 and the β13-β14 and β15-β16 loops) could play a role in crystal dissolution. Residues central to this interface in Cry11Aa are Y272, D514 and D507, which H-bond to one another and to Y203, R222, T249, S251 through direct and water-mediated interactions (W253 and W267), connecting the three domains (Supplementary Fig. 10a, Supplementary Table 1a). Upon pH elevation, Y272, D514 and D507 are all expected to be deprotonated, which should result in electrostatic repulsion and thence dissociation of the three domains. All these residues and their interactions are conserved in Cry11Ba (viz. Y273, D518, D511, Y203, R222, T249, S251, W253 and W268), suggesting that a similar mechanism could be at play. To test the hypothesis, we first produced three Cry11Aa mutants intended to eliminate pH sensitivity of the above-described H-bonds. Neither did the Y272Q nor D507N-D514N mutations impact the overall stability of the toxin, in the soluble or crystalline form (Fig. 4b), but their combination in the triple mutant Y272Q-D507N-D514N resulted in an unexpected abolishment of the ability of Cry11Aa to form crystals in vivo – possibly due to improper folding (Supplementary Fig. 11). The Y272Q mutation had no effect on the pH sensitivity of Cry11Aa crystals, while only a minor effect was seen with the D507N-D514N mutant (Fig. 4a). Thus, the increased stabilization of α8 at the interface between the three domains does not result in an increased tolerance of Cry11Aa crystals to pH elevation. We therefore probed the opposite question in Cry11Ba, and disrupted the hydrogen bond between Y273(OH) and D518(O) by engineering the Y273F mutation (Supplementary Fig. 12a). We found Cry11Ba-Y273F crystals to dissolve at a lower pH, indicating that destabilization of α8 at the interface between the three domains effects in increasing pH sensitivity (Supplementary Fig. 13). Thus, we could increase the pH sensitivity of Cry11Ba

crystals by tampering with interactions between α8 and the three domains, suggesting that dissociation of domains is an important step in the pH induced dissolution of Cry11 crystals. However, decreasing the pH sensitivity of Cry11Aa crystals by stabilization of this region was not successful.

We then focused on Cry11Aa E583, a residue sitting at the intramolecular interface between domain I and domain III. This β21 residue, condemned to be anionic at higher pH, takes part in the water-mediated hydrogen bond network that connects α6 and α7 from domain I with domain III (Supplementary Fig. 10b, Supplementary Table 1b). We asked whether or not suppression of the pH sensitivity of the network would stabilize the monomer at high pH, thereby increasing the tolerance of crystals to pH elevation. This was indeed the case, with an SP_{50} (pH at which 50% of crystals are dissolved) of 12.6 ± 1.0 for crystals of the E583Q mutant, compared to 11.2 ± 1.0 for WT Cry11Aa crystals (Fig. 4a). The dissolution profile was also altered, showing a reduced slope and no visible plateau up to pH 14. Thus, the alteration of the protonation state of residues and water molecules at the intramolecular interface between domain I and domain III may be involved in the early step of Cry11Aa crystal dissolution. In Cry11Ba, which displays a similar SP_{50} of 11.9 ± 1.0 (Supplementary Fig. 13), this residue is substituted for glycine (G587) suggesting a different mechanism of pH-induced intramolecular separation of domain I and domain III – or at least the involvement of additional residues at the interface. We tested this hypothesis by engineering the Y241F mutation in Cry11Ba, effectively suppressing the H-bond between this residue, at the junction between domain I and domain II, and domain III residue D590 (Supplementary Fig. 12b). A reduced SP_{50} value was observed, confirming that the interface between domain I and domain III is central to the tuning of pH sensitivity in Cry11Ba. Considering that Y241 and D590 are strictly conserved in Cry11Aa (Y241 and D586, respectively), this conclusion could be valid for the two toxins.

Crystal contacts were also investigated. We first tampered with the interface enabling the brick-wall packing of Cry11Aa tetramers (Fig. 3c, interface #2) by introducing a F17Y substitution, intended to induce electrostatic repulsion with the negatively charged D180 (distance D180(OD1)—F17(CZ) of 3.3 Å) due to deprotonation of its hydroxyl group upon pH increase (Supplementary Fig. 10c). As expected, crystals of the F17Y mutant were found to be more sensitive to increases in pH, with crystals starting to dissolve at pH as low as ~9.5 and an SP_{50} of 10.6 ± 1.0 (Fig. 4a). The dissolution profile of F17Y crystals is again characterized by a reduced slope, as compared to WT crystals, explaining that the plateau is nonetheless reached at the same pH (~pH 11.6). Nevertheless, the result suggests that dissolution of Cry11Aa crystals can be accelerated by separation of the tetramers associated through interface #2. The F17Y mutation was also found to challenge crystal formation, yielding crystals far smaller than their WT counterparts. We note that F17, D180 and the H-bond between them are strictly conserved in Cry11Ba; hence, the importance of interface 2 for crystal formation and dissolution could be extendable to crystals formed by Cry11Ba.

Next, we challenged the role of the dimerization interface (Fig. 3b, interface #1) by mutating a residue positioned in the central part of the interface, viz. Y449 in Cry11Aa, corresponding to Y453 in Cry11Ba. Recall that the two toxins differ greatly at this interface contributed by domain III from facing monomers, with only 8 hydrogen bonds and 2 salt bridges to support the interface in Cry11Ba, compared to 20 hydrogen bonds and 10 salt bridges in Cry11Aa, and a 38% lower BSA in Cry11Ba than in Cry11Aa. Y449 is not involved in direct H-bonding to other protein residues but supports a large H-bond network that interconnects waters

Table 2 Data collection and refinement statistics of the Cry11Aa mutants.

	Cry11Aa-F17Y pH 7	Cry11Aa-Y449F pH 7	Cry11Aa-E583Q pH 7
PDB ID	7QX7	7QX5	7QX6
Data collection			
Space group	*I 2 2 2*	*I 2 2 2*	*I 2 2 2*
Cell dimensions (Å)	57.72 ± 0.35	57.73 ± 0.24	57.76 ± 0.24
	155.39 ± 1.49	155.55 ± 1.21	155.51 ± 0.98
	171.66 ± 0.64	171.52 ± 0.57	171.51 ± 0.58
Wavelength (Å)	1.33	1.33	1.33
X-ray beam focus (μm)	1.3	1.3	1.3
No. collected frames	3150500	5993679	3523741
No. indexed frames	28227	104359	21833
No. merged crystals	28811	111014	22760
Resolution range (Å)	23.17–3.40	23.78–3.10	23.50–3.30
	(3.48–3.40)	(3.17–3.10)	(3.38–3.30)
No. observations	2908715	20279640	3210163
	(141787)	(1092683)	(154933)
No. unique reflections	10990 (707)	14447 (950)	12014 (787)
$\langle I/\sigma\,(I)\rangle$	6.31 (1.67)	9.95 (1.35)	5.64 (1.52)
R_{split} (%)	19.74 (76.86)	11.79 (89.56)	21.11 (80.18)
$CC_{1/2}$ (%)	95.9 (20.7)[a]	99.8 (59.3)	98.7 (30.7)
Completeness (%)	99.6 (100.0)	99.7 (100.0)	99.6 (100.0)
Multiplicity	265.7 (200.5)	1403.7 (1150.2)	267.2 (196.8)
Refinement			
Resolution range (Å)	23.17–3.40	23.18–3.10	23.08–3.30
No. reflections	10986	14442	12008
R_{work}/R_{free}[b]	21.2/25.1	22.4/25.2	21.5/25.4
No. atoms			
Protein	4970	4965	4970
Water	5	13	6
B-factors (Å²)			
Main chain	54.6	43.1	45.4
Side chain	54.2	42.7	45.3
Water	52.9	59.3	36.0
R.m.s.d.			
Bonds lengths (Å)	0.002	0.002	0.003
Bonds angles (°)	0.448	0.441	0.489

[a]The low value for $CC_{1/2}$ for the outer resolution shell can be justified by the other favorable data quality statistics at this resolution.
[b]R_{free} is calculated using 5% of random reflections excluded from refinement.

and residues from facing monomers in the dimer (Supplementary Fig. 10d, Supplementary Table 1c). Furthermore, the two facing Y449 engage in edge-to-edge aromatic-dimer[36] interactions (centers of the rings are 5.5 Å apart, and the angle between them is 88°; Supplementary Fig. 10d). Contrastingly, the facing Y453 are 10.4 Å apart (center-to-center distance) in the Cry11Ba dimer, but the hydroxyl oxygen is H-bonded to T318(O) (Supplementary Fig. 12c). The Y449F mutation only had a minor effect on Cry11Aa crystal dissolution (Fig. 3a), indicating that deprotonation of its hydroxyl does not play a major role in the dissolution mechanism. Nonetheless, the mutation was detrimental to the protein stability (Fig. 3b) and resulted in the growth of crystals of different size and shape (Fig. 3c), likely owing to destabilization of the H-bond network at the dimer interface. The Y453F mutation, however, resulted in Cry11Ba crystals that dissolved at a lower pH (Supplementary Fig. 13), demonstrating the importance of this H-bond for the stability of the Cry11Ba dimer and the pH-sensing properties of the crystals.

Finally, we introduced a Y349F mutation in the β_{pin}, hypothesizing that suppression of its pH sensitive H-bond to E295(OE1) in the adjacent strand β2 would disturb the β_{pin} fold and destabilize the tetramerization interface (Fig. 3b, interface #3, Supplementary Fig. 10e, Supplementary Table 1d), thereby increasing sensitivity to pH elevation. This expected effect was not observed, with crystals of the mutant displaying the same pH-induced dissolution profile as those of the WT. Nonetheless, smaller crystals were observed that were characterized by a decreased thermal stability (Fig. 4 and Supplementary Fig. 14), indicating that reduced stabilization of the turn preceding the β_{pin} not only impacts folding and stability of the toxin, but as well its packing into crystals – probably due to reduced tetramerization. As Y349 is conserved in Cry11Ba, where it H-bonds to P362(O) (Supplementary Fig. 12d) due to the shorter side chain of D296 compared to Cry11Aa E295, we produced the analogous Y350F mutant and found that it solubilizes at a lower pH (Supplementary Fig. 13). Hence, in absence of a glutamic acid facing the tyrosine hydroxyl (E295), the expected effect on crystal dissolution is seen. This observation suggests that while Cry11Aa E295 and Y349 likely serve the same goal of inducing electrostatic repulsion upon deprotonation, pH sensing mostly depends on the that of E295.

Of all the single and double Cry11Aa mutants we investigated, the Y349F mutation is that which results in the smallest crystals, closely followed by F17Y and E583Q. The Y449F mutant, however, exhibits the most noticeable change in shape compared to WT Cry11Aa. To evaluate the significance of these changes, we characterized the distribution in size of crystals of Cry11Aa-WT, Y449F, F17Y and E583Q using AFM (Fig. 4d). All three mutants had a significantly reduced volume compared to WT Cry11Aa, due to a reduced thickness of the crystals.

Probing crystalline order of the Cry11Aa mutants by SFX. The presence of crystals does not necessarily infer that molecules are well ordered within them. We therefore used SFX to assess the level of crystalline order in crystals of the mutants that displayed modified solubilization or shape. Data were collected at the SPB/SFX beamline of the EuXFEL (Hamburg, Germany) from crystals delivered across the X-ray beam using a liquid microjet focused through a gas-dynamic virtual nozzle GDVN[23] (Table 2). All crystals were kept in water at pH 7 for the GDVN injection, and pulses were delivered at the MHz repetition rate (1.1 MHz)[37,38] using 10 Hz trains of 160 pulses, with a spacing of 880 ns apart. Data was collected on the AGIPD detector at its maximum rate of 3.52 kHz[39]. With the notable exception of Y349F, crystals of all four single point mutants diffracted, yet unequal amounts of data were collected from each, and none from WT crystals, due to technical difficulties that arose during the experiment. This impeded a thorough comparison of the diffraction power of the various mutants, and prevented structure determination for the Y272Q mutant. The structures of the other three mutants were determined, using the WT structure as a molecular replacement model for the phasing of diffraction data. We found that neither overall packing, tertiary structure nor interface formation is affected in the tested mutants at neutral pH (Supplementary Fig. 15). Of important note, these data demonstrate the feasibility of macromolecular nano-crystallography at MHz pulse rate using the brilliant micro-focused beam available at the SPB/SFX beamline of the EuXFEL.

The needle shape inclusions formed by C11AB were also investigated by SFX and found to present some crystalline order, as evidenced by diffraction rings up to ~6 Å resolution in the powder diagram calculated from the maximum projection of 395,656 hits (Supplementary Fig. 9d). It is clear, however, that a high-resolution structure is not readily practicable with these crystals, either because their small size makes them unsuitable for diffraction using a micro-focused XFEL beam or due to intrinsic disorder.

Discussion

We here report the structures of Cry11Aa and Cry11Ba, the two most potent Cry δ-endotoxins expressed by mosquitocidal *Bti* and *Btj*, respectively. Both toxins occur as natural nanocrystals that are produced during the sporulation phase of the bacteria, and dissolve upon elevation of pH in the mosquito larvae gut. Proteolytic activation enables binding to their specific receptors[40], including a membrane embedded alkaline phosphatase[41] but as well the co-delivered Cyt1Aa[12,42–44], triggering insertion in gut cell membranes and subsequent oligomerization into pores that will eventually kill the cells. Both toxins are of industrial interest due to their environmental safety, explained by the multi-step activation outlined above, and their high stability as crystals. Our results shed light on the mechanisms of in vivo crystallization, pH-induced dissolution and proteolytic activation, and on structural features that support the toxins specificity with respect to other Cry toxins. Thereby, our work offers a foundation for further improvement of the toxic activity or crystal size by rational design. Additionally, we demonstrate the feasibility of de novo structure determination of a previously-unknown protein-structure by SFX, from nanocrystals featuring only ~75,000 unit-cells, using a single caged-terbium (Tb-Xo4) derivative. Below, we recapitulate these findings and discuss their implications.

In vivo crystallization pathway of Cry11 toxins. The building block of Cry11Aa and Cry11Ba crystals is a tetramer formed by the interaction of two dimers, via their domain II. The dimers are themselves assembled from the interaction of two monomers, via their domains II and III. Crystals form from the brick-wall packing of tetramers, as enabled by the face-to-back interaction of domain I from symmetry-related tetramers (Fig. 3). Thus, all three domains are involved in the in vivo crystal packing of Cry11 toxins, each contributing a twofold axis. This observation contrasts with other toxin structures determined from in vivo-grown crystals, wherein either propeptide(s) (e.g., *Lysinibacillus sphaericus* BinAB[28] and *Bti* Cyt1Aa[12]) or a specific domain (e.g., domain I in *Btt* Cry3Aa[45,46]) serves as the major contributor to crystallization. Expanding to previously determined Cry δ-endotoxins[5–7,13–15,46,47] structures, solved from in vitro-grown macrocrystals obtained following dissolution of the natural crystals at high pH, the same trend is observed—i.e., crystallization mostly depends on a dedicated portion of the protein, either it be a N-terminal and/or C-terminal propeptide (e.g., the ~650 C-terminal residues in *Btk* Cry1Ac) or a specific domain (e.g., domain II in *Btk* Cry2Aa). Thus, the Cry11Aa and Cry11Ba structures illustrate a yet unobserved pathway for in vivo crystallization, wherein all domains act on a specific step of the coalescence process, viz. dimerization (domains II and III from two Cry11 monomers), tetramerization (domain II from two Cry11 dimers) and tetramer packing (domain I in each tetramer). With Cry11Aa featuring a larger dimerization interface, and Cry11Ba a larger interface between piled tetramers, the two structures underline different levels of tradeoff between packing *into* tetramers and packing *of* the tetramers.

The difference in thickness of Cry11Aa and Cry11Ba crystals is of interest. Considering that all crystals were produced in *Bti*, we could exclude the possibility that the slightly larger size of Cry11Ba crystals originates from a more efficient crystallization machinery in *Btj* than in *Bti*. Puzzled by the presence of a 77-residue long low complexity region at the C-terminus of Cry11Ba (LCR-Cry11Ba), which is absent in Cry11Aa, we asked whether or not a C-terminal fusion of LCR-Cry11Ba with Cry11Aa would result in larger crystals. LCR regions have indeed been shown to support a variety of functions, including chaperoning[35] and reversible oligomerization[48,49] so that the role of the LCR-Cry11Ba could

be found either in crystal nucleation or in crystal growth. Support of the first, but not the second hypothesis was obtained. Indeed, the C11AB chimera, consisting of a fusion of LCR-Cry11Ba to the C-terminus of Cry11Aa, yields smaller crystals that poorly diffract, even when exposed to high intensity XFEL pulses. This observation is in line with previous results which showed that substitution of Cry11Ba domain III for that of Cry11Aa leads to limited expression and comparatively small inclusions[50]. Thus, the LCR region of Cry11Ba is unlikely to account for the difference in size between Cry11Aa and Cry11Ba crystals. Instead, we favor the hypothesis that it is the larger surface of interaction between piled tetramers that accounts for the larger size of the Cry11Ba crystals. Given the absence of electron density for LCR-Cry11Ba residues in the Cry11Ba structure, and the abundance of needle-like inclusions in the parasporal body enveloping the C11AB crystals, it is reasonable to assume that they do not engage in structurally important interactions with functional domains, but rather favor nucleation of crystals. This aid-to-nucleation would be required for Cry11Ba, which features a reduced dimerization interface, but not for Cry11Aa, wherein this interface is 62 % larger. In line with this hypothesis, four regions are predicted to form short adhesive motifs of the Low Complexity, Amyloid-like Reversible Kinked Segments (LARKS) type (Supplementary Fig. 3).

Cry11 toxins depart from the canonical Cry δ-endotoxins architecture. The structures of Cry11Aa and Cry11Ba shed light on features that would not have been predicted based on sequence alignments (i.e., by homology modelling), and which largely deviate from the canonical organization observed in other Cry δ-endotoxins[5–7,13–15,28,45–47]. The most notable difference is the presence of a ~36 to 38 residue insertion between strands β4 and β5 in domain II, which results in an extra β-strand, coined β_{pin}. The β_{pin} not only participates in the formation of a modified β-prism, but contributes to a two-fold axis that supports tetramerization of Cry11 toxins through formation of two large β_h-β4-β3-β2-β_{pin}—β_{pin}-β2-β3-β4-β_h sheets between symmetry-related dimers into a tetramer. The observation of proteolytic cleavage sites at both the N- and C-termini of the β_{pin} suggests that it is removed upon activation by mosquito gut proteases, in line with the observation of ~32 and ~36 kDa fragments upon proteolytic activation of the Cry11 toxins[30–32]. If true, the unique role of the β_{pin} would be to support in vivo crystallization and its removal would entail the dissociation of tetramers into dimers and eventually monomers. While mutagenesis results indicate that this interface does not play a major role in crystal dissolution (see below), it seems likely that upon pH elevation and deprotonation of tyrosines and acidic groups, electrostatic repulsion will occur between Y349(OH) and E295(OE2) in Cry11Aa, and between Y350(OH) and P362(O) in Cry11Ba. Increased disorder of these regions could facilitate the access of proteases, and thus favor the activation of the Cry11Aa and Cry11Ba toxins. This hypothesis would rationalize the reluctance of the two toxins to recrystallize in vitro after pH induced dissolution, due to an impossibility to re-form tetramers – or at least, to re-match the exact positioning of the β_{pin}. The Cry11 toxins also differ from other Cry δ-endotoxins by the presence of a conserved N/D-DDLGITT insertion between β21 and β22, contributing a short helix, and by deletions of ~5–10 residues in the α3-α4 and β20-β21 loops, respectively. Compilation of these changes likely explains failures to phase the Cry11 structures by the molecular replacement method, even when *Btk* Cry2Aa, which also features a $\alpha_h\beta_h$-handle, was used as a starting model.

Mapping the interfaces involved in crystal dissolution. Our efforts to determine the structures of Cry11Aa and Cry11Ba at

alkaline pH were unsuccessful, due to high sensitivity of crystals diffraction quality to pH increase. In the case of Cry11Aa we could not collect data, while in the case of Cry11Ba, we obtained a non-isomorphous structure which, while showing possible inter-domain rearrangements, did not inform on specific side chain rearrangements. Therefore, we resorted to site-specific mutagenesis to obtain information regarding the crystal dissolution pathway. We found that in Cry11Aa, the crystal interface most sensitive to pH elevation is the one enabling the honey-comb brick-wall packing of Cry11 tetramers, with the Cry11Aa-F17Y mutant displaying increased pH sensitivity (with an SP_{50} of 10.6 ± 1.0 compared to 11.2 ± 1.0 for WT Cry11Aa crystals). In contrast, the dimerization (Y349F mutant) and tetramerization interfaces (Y449F mutant) appear to be less pH-sensitive. At the monomer level, we found that the three-domain interface to which α8 and the $\alpha_h\beta_h$-handle contribute is not very sensitive to pH increase (Y272Q and D507N-D514N mutants), possibly due to burying of mutated residues at the interface, preventing bulk solvent to access these sites and therefore retarding pH-sensing. Alternatively, interaction of Cry11 toxins with their membrane-bound receptors[40] could be a required step to expose α8, shown to play a major role in binding and toxicity[51]. Supporting this hypothesis is the observation that disruption of the H-bond between Y273(OH) and D518(O) in the Cry11Ba Y273F mutant increases pH sensitivity, demonstrating that Y273 plays a more important role in structuring the interface between α8 and the three domains than in triggering dissolution by electrostatic repulsion upon pH elevation.

The intramolecular domain I vs. domain III interface was found to be important for the pH-induced crystal dissolution, with the Cry11Aa E583Q and Cry11Ba Y241F mutants displaying reduced and increased sensitivity to pH (SP_{50} of 12.6 ± 1.0 and 11.3 ± 1.0, respectively). Yet unlike the other tested interfaces, which are overall well conserved, the domain I vs. domain III interface differs in Cry11Aa and Cry11Ba, suggesting that caution is advised upon reflecting results obtained from Cry11Aa mutants onto Cry11Ba. Indeed, while Y241 and its H-bonding partner (D586 and D590 in Cry11Aa and Cry11Ba, respectively) are conserved in the two toxins (Supplementary Fig. 16), E583 is substituted for glycine in Cry11Ba (G587), suggesting that further mutagenesis experiments will be needed to further characterize the residues involved in the pH-induced separation of domain I and domain III. Amidst candidate residues to further tune the pH sensitivity is Cry11Ba E247, whose substitution for a glutamine would be expected to reduce electrostatic repulsion of V494 (β14) upon pH elevation. Likewise, E234 H-bonds to Q625(NE2; 2.6 Å) in Cry11Aa, and to K629(NZ; 2.8 Å) and R553(NH1; 2.8 Å) in Cry11Ba, suggesting that a E234Q mutation would reduce pH sensitivity in the two toxins whilst not affecting their folding. Inversely, the mutation into a glutamic acid of Q511/Q515, squeezed between a tryptophan (W584/W588), an arginine (R549/R553) and a glutamic acid (E234), would be expected to increase the pH sensitivity of the domain I vs. domain III intramolecular interface in both Cry11Aa and Cry11Ba – and by extension, that of their crystals.

The Cry11 structures afford rationalization of previous mutagenesis results. Prior to our work, the roles of select domains of Cry11 toxins (α3, α3-α4 loop, β1-α8 loop, α8, β3) have been investigated by mutagenesis. In light of the structures, the effect of single-point mutations can tentatively be clarified. Specific to Cry11Aa, one mutation was shown to fully abrogate crystal formation (V104E), whereas seven others were reported to reduce or suppress toxicity against *Aedes aegypti* mosquitoes (Supplementary Table 2)[31,51–54]. The α3 residue V104 is not

present at a crystal interface, but rather exposed at the surface of domain I[31]. Thus, the V104E mutation is more likely to challenge crystal formation by affecting the folding of domain I. Given the immediate environment of this residue, we see two possible explanations; either electrostatic repulsion of E104 by α4 residues L129 and A132, resulting in destabilization of domain I; or the formation of a salt-bridge to R136, replacing the native salt-bridge to α3 E97 (Supplementary Fig. 17a). Supporting the latter hypothesis is the report that the E97A mutation also produces a non-toxic variant of Cry11Aa[31]. Further pushing forward the centrality of the R136-E97 salt bridge between α3 and α4 is the report that the R90E mutation also leads to a non-toxic Cry11Aa variant[31]. Indeed, replacement of R90 by a glutamic acid would force the Q139 side chain to flip, resulting in electrostatic repulsion of E97 and thereby disruption of the H-bonding pattern between α3 and α4 (Supplementary Fig. 17a). More difficult to understand is, however, how the mutation into an aspartic acid of V142, facing R90, results in a non-toxic Cry11Aa variant[54]. One would indeed expect that higher stabilization of the two helices by a direct salt bridge between R90 and D142 would result in a more potent toxin. Two other mutations introduced in α3 led to non-toxic Cry11Aa variants, namely Y98E and S105E[31]. The structure shows that Y98 fills a hydrophobic groove contributed by F63, P68, M71, L94, I101, I102, F140, L152 and M156 (Supplementary Fig. 17b); while S105(OH) H-bonds to main chain α3 atoms I101(O) and I102(O) and S105(Cβ) plugs another hydrophobic groove contributed by F55, L59, F108 and L129 (Supplementary Fig. 17c). It thus seems plausible that replacement of either residue by a glutamic acid would result in a destabilized domain I. The other Cry11Aa region that was targeted by mutations[51–53] is helix α8 and preceding residues in the β1-α8 loop, both of which are sandwiched between the β-handle in domain II and α6 and α7 in domain I. Briefly, full suppression of toxicity to *Aedes aegypti* mosquitoes was observed for V262E. Concerning the P261A, V262A and E266A mutations, results are contradictory in the literature, either mentioning a reduced toxicity or no difference with the wild-type Cry11Aa (Supplementary Table 2)[51–53]. If confirmed, the reduced activity of the P261A mutant would highlight the importance of the kink between α8 and the β1-α8 loop for the activity of Cry11Aa (Supplementary Fig. 17d). Interestingly, in Cry11Ba whose β1-α8 loop sequence differs, a proline residue is also found (P265) yet three residues downstream. The V262 residue fits in a hydrophobic groove contributed by C211, L215, I260, W267, V271, C432 and L438 (Supplementary Fig. 17d). Hence, it is plausible that introduction of a glutamic acid in this groove (V262E) would completely disrupt the interface between domain I and domain II, while replacement by the shorter side chain of an alanine (V262A) would loosen it. Last, the α8 residue E266 is exposed to the solvent and H-bonds to both the side chain and main chain nitrogen atoms of N263, in the β1-α8 loop. Mutation of E266 into an alanine will result in disruption of these H-bonds, and therefore in a destabilization of the interface between α8 and the β1-α8 loop (Supplementary Fig. 17e).

In Cry11Ba, most mutations have been made in triplet, making it difficult to pinpoint the contribution of each residue to the observed phenotypes[55]. Nonetheless, five single-point mutations have been reported, which led to Cry11Ba variants with reduced or suppressed toxicity against *Aedes aegypti*, *Anopheles stephensis* and *Culex quinquefasciatus* mosquitoes. Echoing the work described above, three of these have been introduced in α8 and the β1-α8 loop. Residue I263 is structurally equivalent to V262 in Cry11Aa and likewise, its side chain fits into a groove contributed by L211, L215, I260, W268, L272, C436 and L442 at the interface between domain I and domain II (Supplementary Fig. 18a). Hence, we propose that the I263A mutant exerts its detrimental

effect on Cry11Ba toxicity in the same fashion than the V262A mutation in Cry11Aa. Likewise, the exposed side chains of S264 and K269 contribute to structuring α8 and the β1-α8 loop through interaction with D267(N) and P265(O), respectively (Supplementary Fig. 18b). Hence it is possible that the S264A and K269A mutations exert their effects in the same fashion as the E266A mutation in Cry11Aa, i.e., a destabilization of the region encompassing α8 and the β1-α8 loop. Since the effect of the three above mutations on the three mosquito species is not the same (Supplementary Table 2), it may be proposed that α8 and the β1-α8 loop are part of the binding epitope on the midgut microvillar receptor. The two other Cry11Ba residues that have been challenged by single point mutations are G257 and I306, both mutated into alanine (G257A, I306A). The former is involved in the conserved turn between β1 and the β1-α8 loop (the equivalent residue in Cry11Aa is also G257), hence the results highlight the necessity of a tight turn at the end of β1 to achieve full toxicity (Supplementary Fig. 18a). The latter is located in the middle of domain II β-prism, where its side chain fits into a hydrophobic grove contributed by F289, Y291, V309, W400, I402, L410 and I466 (Supplementary Fig. 18c). It is presumable that the I306 side chain replacement by the shorter side chain of an alanine effects in loosening the β-prism, which in turn would result in a reduced interaction with the receptor. Again, the effect of the two latter mutations varies depending on the considered mosquito species, suggesting an involvement in the binding to the receptors (Supplementary Table 2).

Implication for the future of nanocrystallography using SFX. In this study, de novo phasing was required—not because of the absence of homologous structures, but because none of those available were sufficiently close to serve as a search model for molecular replacement. Using Tb-Xo4, a caged terbium compound, we could phase the *Bti* Cry11Aa structure by SAD, from ~77,000 diffraction patterns collected on crystals consisting of ~75,000 unit cells on average – an achievement to compare to the determination of the *Ls* BinAB structure from >370,000 patterns (native and three derivatives) collected on crystals with ~100,000 unit cell[28]. Our success in phasing the Cry11Aa structure likely stemmed from a combination of the use of a dramatically powerful phasing agent[16,17] and advances in SFX data processing tools over the last five years, including the Xgandalf[56] indexing algorithm and improvements in data handling and integration in CrystFEL[57]. It should offer hope to investigators seeking to determine the structure of proteins of which no known structural homologue exists and that have to resort to SFX due to smallness of their crystals. It is foreseeable, however, that de novo structure determination will be helped by recent advances in comparative and ab initio modelling and the availability of programs such as RosettaFold[58] and AlphaFold2[59], capable of producing a decently-accurate structure for virtually all proteins and thus a good model for phasing of crystallographic data by molecular replacement. Latest releases of the two programs were published in the final stage of the writing of this manuscript, hence we asked whether or not the availability of these tools would have facilitated our journey towards the Cry11 toxins structures, and submitted the sequence of Cry11Aa to the two servers. For RosettaFold, the r.m.s.d. to the final refined structure of the five best models was over 4 Å, with discrepancies observed mostly in domain II. For AlphaFold2, however, the two first models displayed r.m.s.d. of 1.2 and 1.0 Å to the final structure, respectively. Using the worst of these two models, we could find a molecular replacement solution using Phaser, and a partial model featuring 95% of the residues in sequence was obtained after 20 cycles of automatic iterative model-building and refinement using

Bucanneer[60] and Refmac5[61]. Thus, a problem which occupied five crystallographers for several years could have been solved in less than an hour using the new tools recently made available to the structural biology community. Based on our results, it is tantalizing to claim that the phase problem in crystallography has been solved, or that experimental structural biology will be abandoned, but such assertions would likely be shortsighted. Rather, we encourage investigators to challenge AlphaFold2 and RosettaFold as much as humanly possible, but to not forsake de novo phasing as it may remain the only route to success in difficult cases where molecular replacement based on such models does not work[62]. It must also be emphasized that in the case of Cry11 toxins and, more generally, naturally-crystalline proteins, the issue is not just phasing, but packing. For such proteins, crystal formation and dissolution serve function, hence characterization of packing interfaces is central to finely comprehend their bioactivation cascades. Without the naturally-occurring crystals and the atomic resolution experimental SFX data, it would not have been possible to make predictions on potential mutations affecting Cry11Aa crystal formation or dissolution.

Methods

Crystal production and purification. Crystals of Cry11Aa and Cry11Ba were produced by electrotransformation of the plasmids pWF53 and pPFT11S[63] into the acrystalliferous strain 4Q7 of *Bacillus thuringiensis* subsp. *israelensis* (*Bti*; The Bacillus Genetic Stock Center (BGSC), Columbus OH, USA), respectively[64]. Colonies were selected on LB agar medium supplemented with erythromycin (25 µg/mL) and used to inoculate precultures of LB liquid medium. For Cry11Aa production, precultures were spread on T3 sporulation medium. After incubation at 30 °C for 4 days, spores/crystals suspensions were collected using cell scrapers and resuspended in ultrapure water. After sonication-induced cell lysis and subsequent centrifugation at 4,000 g for 45 min to discard cell and medium debris, pellets were resuspended in water and crystals were purified by ultracentrifugation (23,000 × g, 4 °C, 16 h) on discontinuous sucrose gradient (67-72-79%). After ultracentrifugation, crystals were recovered and several rounds of centrifugation/resuspension in ultrapure water allowed discarding as much sucrose as possible for proper downstream application. Crystal purity was verified by SDS-PAGE on 12% gels. Purified crystals were conserved in ultrapure water at 4 °C until use. For Cry11Ba, a glycerol stock of the 4Q7/pPFT11S was streaked onto 25 µg/mL erythromycin Nutrient Agar plates. From here a single colony was selected and added to a Glucose-Yeast-Salts (GYS) media culture and allowed to grow continuously at 30 °C, 250 rpm for 5 days. This culture was then spun down, resuspended in ultrapure water, and the lysate was sonicated for 3 min at 50% duty. The sonicated lysate was added to the 30–65% discontinuous sucrose gradient (35-40-45-50-55-60-65 %) and spun down for 70 min at 20,000 rpm and 4 °C. The sucrose gradient was then hand fractionated with Cry11Ba crystals collected around 57–60% and dialyzed into ultrapure water. Crystal characterization and purity was completed by phase contrast light microscopy, X-ray powder diffraction, transmission electron microscopy, and 4–12% SDS-PAGE gels. The pure Cry11Ba crystals were stored at 4 °C in ultrapure water.

Cry11Aa mutagenesis. Based on the SFX structure of Cry11Aa, a total of 7 mutants of Cry11Aa were constructed to challenge the different crystal packing and intramolecular interfaces. The rationale behind these mutations is illustrated in Supplementary Fig. 10 and discussed in the main text. Point-mutations were inserted into *cry11aa* gene by Gibson assembly using pWF53 as a backbone[64]. Two different primer couples were used for each mutation to amplify two fragments that were complementary by their 15–18 bp overlapping 5′ and 3′ overhangs with a target Tm of 50 °C. Point mutations were inserted in the complementary part of the overhangs of the two fragments spanning the *cry11aa* region to be mutated. The double mutant D507N-D514N was successfully constructed in a single-step by respectively adding the D507N mutation on the non-overlapping overhang region of the forward primer, and the D514N on the non-overlapping overhang of the reverse one. The triple mutant Y272Q-D507N-D514N was constructed by using the primers containing the Y272Q mutation and the plasmid pWF53-D507N-D514N as a backbone. In addition to the point mutants, a Cry11Aa-Cry11Ba chimeric toxin—coined C11AB—was also constructed. For this, the sequence of the *cry11aa* gene was fused with the 234 bp extra 3′ extension of *cry11ba* gene, which is suggested to feature a low complexity region (LCR) based on sequence analysis using the LCR-eXXXplorer web platform (http://repeat.biol.ucy.ac.cy/fgb2/gbrowse/swissprot)[65], which implements the CAST[19] and SEG[20] computational methods to identify LCR. The C11AB chimera was constructed by Gibson assembly following a "1 vector, 2 fragments" approach. The plasmid pWF53 containing the *cry11aa* gene was used as a backbone and the *cry11ba* 3′ fragment was amplified from the extracted and purified plasmid of the WT strain of *Btj* containing the

cry11ba gene. The list of primers used for plasmids construction is available in Supplementary Table 3. For each plasmid construction, the fragments with overlapping overhangs were assembled using the NEBuilder HiFi DNA Assembly (New England BioLabs) as previously described[12]. Briefly, after 90 min incubation at 50 °C, the constructed plasmids were transformed by heat shock into chemically competent Top10 *Escherichia coli* (New England BioLabs). Plasmids were extracted from colonies selected on LB agar medium containing ampicillin (100 μg/mL) using the NucleoSpin Plasmid extraction kit (Macherey-Nagel) following the manufacturer's instructions. The successful construction of each plasmid was assessed by double digestion (EcoRI and BamHI) followed by migration on 1% agarose gel stained with SYBR Safe (Invitrogen) and by Sanger sequencing of the region containing the mutation at the Eurofins Genomics sequencing platform. Of note, the *cry11aa* gene was also fully sequenced to validate its sequence for mutagenesis primer design and for comparing the expected toxin size to the observed one in mass spectrometry analyses. All mutants were produced as crystals in *Bti*, as described above. The presence of the mutated *cry11aa* gene sequence in the transformed *Bti* colony used for production was verified by colony PCR using specific primers and Sanger sequencing at the Eurofins Genomics sequencing platform. Crystals from all mutants were analyzed by SDS-PAGE on 12% gels. For C11AB, its proper size was confirmed by using the "gel analysis" module implemented in the software ImageJ v1.51k ($N = 7$)[66].

Cry11Ba mutagenesis. gBLOCK gene sequences composed of 2877 bp harboring the open reading frame of Cry11Ba, and targeted point mutations resulting in single amino acid replacements (Y241F, Y273F, Y350F, Y453F) in cry11Ba, expressed under the cyt1Aa BtI, BtII, and BtIII promoters and featuring its transcription termination sequence[67], were synthesized at Integrated DNA Technologies (IDT, San Diego; Supplementary Table 4). These constructs were designed for cloning into the *E. coli* - *Bt* shuttle vector pHT3101[68], and contained homology sequences at the 5′ end (gaccatgattacgaatt) and 3′ end (gcatgcaagcttggc) for directional assembly in pHT3101 linearized with EcoR1 and Pst1 using the Choo-Choo Cloning Kit (Molecular Cloning Laboratories, MCLAB, San Francisco), according to manufacturer's protocol. Recombinant plasmids were propagated in E. coli DH5a, purified using the Wizard Plus Miniprep Kit (Promega), and point mutations were confirmed by sequencing (Genomic Core Facility, University of California Riverside) using internal forward and reverse primers that flanked the sites of interest (forward: DB11BseqF1 5′-GAATTTAGGAGGAAGCGATTGGGGA-3′ and DB11BseqF2 5′GTATTGTACTGAAAAGAATTTGGACGG-3′; reverse: DB11Seq11R 5′CTGGTGTATCTTCTAAGAATGATCTAT-3′). The acrystalliferous *B. thuringiensis* subsp. *israelensis* strain 4Q7 was transformed with the recombinant plasmids by electroporation as described previously[67]. Transformants were selected on Nutrient Agar supplemented with erythromycin (25 mg/ml) and the presence of crystals were initially monitored by phase contrast microscopy.

Crystal visualization by scanning electron microscopy (SEM). Purified crystals of Cry11Aa WT and of the 7 mutants were visualized using either a Zeiss LEO 1530 scanning electron microscope from the SEM facility of the European Synchrotron Radiation Facility (ESRF, Grenoble, France), a Thermo Fisher Quanta 650 FEG environmental SEM (ESEM) available for users at the European XFEL (EuXFEL, Hamburg, Germany) or a JEOL JSM-6700M FE-SEM (UCLA, Los Angeles, USA). For SEM at ESRF, samples were coated with a 2 nm thick gold layer with the Leica EM ACE600 sputter coater before imaging. For ESEM at the European XFEL, samples were diluted (1/1000) and mixed with 25 mM of ammonium acetate. Samples were then coated with a thin gold layer as described above using a Leica EM ACE600 sputter coater as well. Images were recorded at 15 kV acceleration voltage by collecting secondary electrons using an Everhart-Thornley-Detector (ETD detector) in high-vacuum mode. For SEM at UCLA, samples were diluted (1/5) in ultrapure H_2O. They were then added to 300 mesh Cu F/C grids that were positively glow discharged. These samples were then wicked away and washed with ultrapure water, wicked, and allowed to dry overnight to ensure all moisture had evaporated inside of a dessicator. These were then attached to a holder with carbon tape and coated with an Anatech Hummer VI sputter coater with approximately 2 nm of thick gold layer. Images were recorded at 5 kV acceleration voltage by collecting secondary electrons using a Lower secondary electron (LEI) or Upper secondary electron in-lens (SEI) detector.

Crystal visualization by transmission electron microscopy (TEM). Nonpurified crystals of Cry11Aa WT were visualized using a Thermofisher TF20 electron microscope from the IBS electron microscopy platform. For negative staining TEM, samples were diluted 5 times in H_2O and 4 μL of the diluted sample was introduced to the interface of an amorphous carbon film evaporated on a mica sheet. The carbon film was then floated off the mica sheet in ~200 μL 2% sodium silicotungstate (SST) solution. The carbon film with the crystal sample was then recovered onto a Cu 300 mesh TEM grid after 30 s, let dry, and imaged at 200 keV. Images were recorded on a Gatan OneView CMOS detector. Non-purified crystals of Cry11Ba WT were visualized using a FEI Tecnai T12 electron microscope within the UCLA California Nanoscience Institute, EICN facility. For negative staining TEM, samples were prepared by adding 5 μL of pure crystal fractions in 10 μL ultrapure H_2O. In total, 2.5 μL of this sample was added to 300 mesh Cu F/C grids

that were positively glow discharged. These samples were then wicked away using Whatman 1 filter paper; washed with 2.5 μL ultrapure H_2O, wicked; and negatively stained with 2.5 μL 2% uranyl acetate, wicked. These were allowed to dry overnight to ensure all moisture had evaporated and imaged at 120 keV. Images were recorded on a Gatan 2 K x 2 K CCD.

Crystal characterization by atomic force microscopy (AFM). Crystals of Cry11Aa were visualized by AFM as previously described[12]. Briefly, 5 μL of crystals suspended in ultrapure water were deposited on freshly cleaved mica. After 30 min in a desiccation cabinet (Superdry cabinet, 4% relative humidity), crystals were imaged on a Multimode 8, Nanoscope V (Bruker) controlled by the NanoScope software (Bruker, Santa Barbara, CA). Imaging was done in the tapping mode (TAP) with a target amplitude of 500 mV (about 12 nm oscillation) and a variable setpoint around 70% amplitude attenuation. TESPA-V2 cantilevers (k = 42 Nm⁻¹, Fq = 320 kHz, nominal tip radius = 7 nm, Bruker probes, Camarillo, CA, USA) were used and images were collected at ~1 Hz rate, with 512- or 1024-pixel sampling. Images were processed with Gwyddion[69], and if needed stripe noise was removed using DeStripe[70]. Measurements were performed on Cry11Aa WT and on mutants selected on the basis of their aspect in eSEM images (Y449F) or their solubilization pattern (F17Y and E583Q). Size measurements were performed on AFM images using Gwyddion[69] in a semi-automated protocol. A classical height threshold was applied to each image to select as many individual crystals as possible. Sometimes, partially overlapping crystals were individualized using the manual edition of the mask of selected crystals by adding a separation line. Finally, a filter was applied to remove very small selections (artefacts) or crystals touching the edge of the image. Measures were obtained using the 'distribution of grains' feature in Gwyddion where the crystal thickness (T) is the returned mean value, the volume (V) is the Laplacian background basis volume, and the length and width are the major and minor semi-axes of equivalent ellipses, respectively. The total number of crystals measured are: 45 for WT, 93 for F17Y, 60 for Y449F, and 94 for E583Q.

Data collection history. The Cry11Aa/Cry11Ba structure determination project was initiated in 2015. Data were collected at five different occasions, in two XFEL facilities, namely at the Linac Coherent Light Source (LCLS), Stanford (USA) and EuXFEL, Hamburg (Germany). During our first LCLS-SC3 beamtime (cxi04616), we collected data from native Cry11Ba (2.3 Å resolution), and in our second beamtime (LO91), we collected data from native Cry11Aa (2.8 Å resolution). Nanocrystals grown by recombinant expression in the modified acrystalliferous 4Q7 strain of *Bti* were injected by a microfluidic electrokinetic sample holder (MESH) device[22] in the microfocus chamber of LCLS-SC3[21]. After data reduction using cctbx.xfel and dials (hit-finding through merging)[71–74], we attempted phasing of both datasets by molecular replacement (MR), using sequence-alignment based multi-model approaches implemented in Mr Bump (based on MR by Molrep[75]) as well as custom-scripts testing models produced by Rosetta[24] (using the Robetta server; http://robetta.bakerlab.org/) and SwissProt[25] (https://www.ebi.ac.uk/uniprot/) servers (based on MR by Phaser[76]). Failure to find a homologue of a sufficiently-close structure led us to attempt de novo phasing of the Cry11 nanocrystalline proteins.

Initially, we aimed at obtaining experimental phases for Cry11Ba, considering that its larger crystals would produce a stronger diffraction signal which in turn would facilitate phasing. Hence, we collected derivative data on Cry11Ba, from crystals soaked with Gd, Pt and Au salts (P127 experiment) before injection using a MESH device[22]. Unfortunately, the data did not allow phase determination, as indicated by very weak and absent peaks in the isomorphous and anomalous difference maps, respectively (Supplementary Fig. 2), due to low occupancy of the soaked metal ions. Hence, we shifted focus to Cry11Aa crystals soaked with a recently introduced caged-terbium compound, Tb-Xo4[16,17] (P125 experiment). Crystals were injected using a GDVN[23] liquid microjet in the microfocus chamber of LCLS-SC3[21]. Online data processing was performed using NanoPeakCell[77] and CASS[78]. Offline data processing with NanoPeakCell[77] (hit finding) and CrystFEL[57] (indexing and merging) revealed a strong anomalous signal that enabled determination of the substructure and phasing of the SFX data, using Crank2[79] and its dependencies in the CCP4 suite[80] (see below for more details). The Cry11Aa structure was thereafter used to phase the Cry11Ba datasets by molecular replacement. *A posteriori*, we discovered that two of the heavy atom compounds that we used for soaking actually did bind Cry11Ba (Supplementary Fig. 2b-c). Difference Fourier maps revealed 7–8 σ peaks indicating Pt bound near Met 19 and 200, and Gd bound near Asp83 and Asp427 (Supplementary Fig. 2b). Surprisingly, however, there were no peaks in the anomalous difference Fourier maps. We speculate, however, that if we had achieved higher heavy-atom occupancy and/or higher multiplicity in our measurements, the anomalous signal would have been strong enough to detect and perhaps used for phasing. We note that an alternative strategy could have been to first obtain experimental phases (either by seleno-methionylation or soaking with heavy metals) from in vitro-grown macro-crystals obtained by isolation, dissolution, recrystallization and derivatization, which could have allowed phasing by molecular replacement. However, as we could not exclude that Cry11 toxins would undergo large structural changes upon pH-induced activation, which would have complicated molecular replacement, we opted for the current strategy.

We last attempted data collection on Cry11Aa and Cry11Ba crystals soaked at elevated pH and injected by a MESH device (P141 experiment). Only Cry11Ba

crystals could sustain the pH jump and yielded usable data. From the comparative analysis of the Cry11Aa and Cry11Ba structures, we nonetheless designed mutations aimed at increasing or decreasing the resilience of crystals; these were introduced in the *cry11aa* gene, and crystals were produced by recombinant expression in *Bti*. From these, SFX data were collected at the MHz pulse rate, during experiment P2545 at the SPB/SFX beam line of EuXFEL where a GDVN was used to inject crystals. The data were also processed with NanoPeakCell[77] (hit finding) and CrystFEL (indexing and merging).

It might be asked whether or not differences in data quality, arising from the use of different injection methods, could have played a role in the success in phasing Cry11Aa data, but not Cry11Ba data. Indeed, the use of a GDVN device, compatible with injection of a colloidal suspension of crystals in pure water, enables constant background in the diffraction patterns. This is less straightforward to achieve using a MESH device as the smaller size of the jet (reducing sample consumption by 5–10 fold) results in decreased stability (requiring to reposition the jet more often) and in the necessity to add highly viscous solvents to the crystal slurry (to avoid freezing in the vacuum chamber). To conclude on this point, a systematic study would be needed, whereby datasets assembled from the same number of images collected with either type of injector would be compared.

Data collection and processing, and structure refinement. During the P125 beamtime at LCLS, where the SAD data used for the phasing of the Cry11Aa structure were collected, the X-ray beam was tuned to an energy of 9800 eV (i.e., a wavelength of 1.27 Å), a pulse duration of 50 fs, a repetition rate of 120 Hz, and a focal size of 5 μm. SAD data were collected from nanocrystals soaked for 30 h with Tb-Xo4 at 10 mM in water, prior to GDVN injection[23]. Of 558747 images collected using the 5 μm beam available at LCLS-SCC, 76687, 292, 217 and 177 were indexed (i.e., a total of 77373 images) using Xgandalf[56], Dirax[81], taketwo[82] and Mosflm[83], respectively, in CrystFEL v.0.8.0[57]. Post-refinement was not attempted, but images were scaled one to another using the 'unity' model in CrystFEL *partialator*, yielding a derivative dataset extending to 2.55 Å resolution. *A posteriori*, we found that simple Monte Carlo averaging using the 'second-pass' option in CrystFEL *process_hkl* would have yielded data of similar quality, most probably because of the large number of indexed images. A native dataset was also collected and processed in the same fashion yielding, from 792,623 collected patterns of which 48,652 were indexed, a dataset extending to 2.60 Å resolution. The substructure of the derivative dataset was easily determined by ShelxD (figure of merit (FOM): 0.22), prompting us to try automatic methods for structure determination. Using Crank2[79] and its dependencies (ShelxC, ShelxD, Solomon, Bucanneer, Refmac5, Parrot) in CCP4 Online[84], the FOM was 0.52 after density modification, and rose to 0.88 upon building of 613 residues. This first model was characterized by R_{work}/R_{free} of 27.7/32.1% and was further improved by automatic and manual model building in phenix.autobuild[85] and Coot[86] until 630 residues were correctly built. This model was then used to phase the native data. Final manual rebuilding (using Coot[86]) and refinement (using phenix.refine[87] and Refmac5[61]) afforded a native model characterized by R_{work}/R_{free} of 17.2/24.1 % and consisting of most of the 643 residues. Only the first 12 N-terminal residues are missing (Table 1).

Cry11Ba data were collected during the cxi04616 and P141 beamtimes at LCLS-CXI. At both occasions, the photon energy was 9503 eV (i.e., a wavelength of 1.30 Å), a pulse duration of 50 fs, a repetition rate of 120 Hz, and a focal size of 1 μm—i.e., a similar standard configuration (pulse length, repetition rate) to that used for Cry11Aa, notwithstanding the beam size and wavelength. Data were collected from crystals at pH 6.5 (30% glycerol in pure water; cxi04616) and pH 10.4 (30% glycerol in 100 mM CAPS buffer; P141), presented to the X-ray beam using a MESH injector[22]. Of 813133 images collected for the pH 6.5 dataset, 19708 were indexed and scaled, post-refined, and merged using cctbx.xfel[71–74] and PRIME[88], yielding a dataset extending to 2.4 Å resolution. The Cry11Aa structure was used as a starting model to phase the Cry11Ba pH 6.5 dataset by molecular replacement using Phaser[76] with initial R_{work}/R_{free} being 34.4/40.4 %. Manual model building (using Coot[86]) and refinement (using Refmac5[61], phenix.refine[87] and Buster[89]) afforded a model characterized by R_{work}/R_{free} of 18.7/23.1 % (Table 1).

We used the refined Cry11Ba pH 6.5 structure as the starting model for the Cry11Ba pH 10.4 structure. We used Refmac5[61] to perform rigid body refinement and then Refmac5[61], phenix.refine[87], and Buster[89] to perform individual atomic refinement at a resolution of 2.65 Å. We performed manual model building with Coot[86]. The R_{work}/R_{free} of the final model was 23.7/24.7 %. The structural changes between the pH 6.5 and pH 10.4 models were difficult to interpret. No major peaks were observed in the difference Fourier difference map obtained by subtracting the pH 10.4 structure factors from the pH 6.5 structure factors. Consistent with this observation, there were no significant local structural changes, only a small contraction in the separation between subdomains. This contraction is consistent with a 1% shrinkage of the unit cell volume at pH 10.4. We ascribe this shrinkage to crystal dehydration caused by the use of a higher glycerol concentration during injection of the pH 10.4 sample. The conformational changes arising from elevated glycerol confound our interpretation of pH-related structural changes. Hence, we do not further discuss it in our manuscript.

Diffraction data on the Cry11Aa mutants at pH 7.0 was acquired on the SPB/SFX beamline at EuXFEL during our P002545 beamtime allocation, using a GDVN injector and X-ray energy and focal size of 9300 eV (1.33 Å) and 1.3 μm (FWHM),

respectively. Technical problems allowed us to collect only a limited number of diffraction pattern of the Cry11Aa-Y349F mutant. 3150500; 5993679 and 3523741 images were collected for the F17Y, Y449F and E583Q mutant, respectively, of which 28227, 104359 and 21833 could be processed using CrysFEL0.8.0[57] and MonteCarlo based scaling and merging. The three structures were solved using MR with Phaser[76], using the Cry11Aa WT structure as input model. The structures were refined with only two B-factors per residue and secondary structure restraints in Phenix.refine[87] and Coot[86], with final R_{work}/R_{free} values of 21.2/25.1 % for Cry11Aa-F17Y, 22.4/25.1 % for Cry11Aa-Y449F and 21.5/25.4 % for Cry11Aa-E583Q (Table 2).

Structure analysis. Figures were prepared using PyMOL v. 2.5[90] (Figs. 2, 3 and Supplementary Fig. 4, 5, 10, 14, 15, 16, 17) and aline (Supplementary Fig. 3)[91]. Radii of gyration were predicted using the PyMOL script rgyrate (https://pymolwiki.org/index.php/Radius_of_gyration). Interfaces were analyzed with PISA[27] and root mean square deviations (r.m.s.d.) among structures were calculated using PyMOL using the 'super' algorithm. Sequence based alignment—performed using EBI laglign and ClustalW[92]—was challenged by the large gaps between *Bti* Cry11Aa, *Btj* Cry11Ba, *Btk* Cry2Aa and *Btt* Cry3Aa, while structure-based alignment—performed using SSM[93]—was blurred by the varying size of secondary structure elements in the three domains of the various toxins. Hence, for Supplementary Fig. 1, 3, the alignment of *Bti* Cry11Aa, *Btj* Cry11Ba, *Btk* Cry2Aa and *Btt* Cry3Aa was performed using strap[94] which takes into account both sequence and structural information. Specifically, the online version of the program was used (http://www.bioinformatics.org/strap/)[95]. To generate the phylogenetic tree in Fig. 1, We used the CCP4 program lsqkab to compute all pairwise superpositions of the 15 delta-endotoxins, and the r.m.s.d. of aligned alpha carbons. We uploaded a 15 × 15 matrix of r.m.s.d. values to the T-REX phylogenetic tree plot server www.trex.uqam.ca[96].

Structure prediction using AlphaFold2 and RosettaFold. RosettaFold[58] predictions were obtained by submitting the sequence to the Rosetta structure-prediction server (https://robetta.bakerlab.org). AlphaFold2[59] predictions were obtained by use of the Collaboratory service from Google Research (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb). The mmseq2 method[97,98] was employed for the multiple-sequence alignment instead of the slower jackhmmer method[99,100] used in Ref. [59]. Structural alignments were performed using the *align* tool in PyMOL[90]. Molecular replacements trials were carried out with Phaser[76]. Using the best five RosettaFold models, all characterized by an overall r.m.s.d. to the final structure superior to 4 Å, no molecular replacement solution could be found, due to inaccurate prediction of domain II β~pin~ and α~h~-β~h~ regions, resulting in clashes. The best AlphaFold2 model was yet successful at predicting the domain II structure, which enabled successful phasing by molecular replacement, yielding a model characterized by R_{free} and R_{work} values of 0.322 and 0.292, respectively. This model was further used as a starting model for automatic model building and refinement using the Buccaneer pipeline in CCP4, resulting in a model characterized by R_{free} and R_{work} values of 0.245 and 0.215, respectively, after only five automatic cycles of iterative model-building, refinement and density modification using Buccaneer[60] and Refmac5[61] in the CCP4 suite[80].

Crystal solubilization assays. The solubility of crystals of Cry11Aa WT and of the mutants F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q was measured at different pH values as previously described[12]. Briefly, crystal suspensions were centrifuged at 11,000 × g for 2 min and resuspended in 18 different buffers with pH ranging from 8.6 to 14.2. After 1 h incubation in each buffer, crystals were centrifuged and the supernatant was collected. The concentration of soluble toxin in the supernatant was quantified using a Nanodrop 2000 (Thermo Fisher Scientist) by measuring the OD at 280 nm and by using the molar extinction coefficient and toxin size (102,930 M^{-1} cm^{-1} and 72.349 kDa, respectively, as calculated with the ProtParam tool of ExPASy (https://www.expasy.org) using the Cry11Aa protein sequence available under accession number "P21256"). Solubility was measured in triplicate for each toxin (Cry11Aa WT and mutants) and each pH. Data are normalized and represented as percentage of solubilization by dividing the concentration measured at a given pH by the maximum measured concentration.

For Cry11Ba and its mutants, the crystal suspensions were centrifuged at 13,300 × g for 3 min and ultrapure H$_2$O was removed from crystals. They were then resuspended in one of 18 buffers ranging from pH 7 to 14. These crystals were incubated for 1 h, afterwards the solution was centrifuged at 13,300 × g and the supernatant was separated from the crystal pellet. The concentration of the supernatant was then quantified by a ThermoFisher Nanodrop One (Thermo) by measuring the OD at 280 nm and using the molar extinction coefficient and toxin size (114600 M^{-1}.cm^{-1} and 81.344 kDa respectively) that were calculated with Expasy ProtParam using the Cry11Ba sequence available at Uniprot.org under accession number Q45730. Solubility was measured in triplicate for each toxin at each pH measured. Data are normalized and represented as percentage of solubilization by dividing the concentration measured at a given pH by the maximum measured concentration. For Cry11Ba WT, this was then further tested by conducting a turbidity assay by resuspending the crystal pellet in 150 μL

ultrapure H_2O and placed in a 96-well plate to be read on a NEPHELOstar Plus (BMG Labtech) nephelometer. These counts were normalized by subtracting the background signal and conducted in triplicate.

Solubility of Cry11Aa WT, Cry11Ba WT and their mutants was compared by calculating SP_{50} (pH leading to solubilization of 50% of crystals) as previously described[12], by globally fitting the data using a logistic regression model for binomial distribution using a script modified from Ref. [101]. Differences in SP_{50} between mutants were considered significant when 95% confidence intervals (CI), calculated using a Pearson's chi square goodness-of-fit test, did not overlap[102], and was confirmed by the individual fits of each of the triplicate measurement (Supplementary Fig. 19). All statistics were conducted using the software R 3.5.2[103].

Proteomic characterization. For SDS-PAGE experiments, samples heated to 95 °C were migrated on 12 % SDS-PAGE gels (1 h, 140 V) after addition of Laemmli buffer devoid of DTT. After staining by overnight incubation in Instant*Blue* (Sigma Aldrich, France), gels were washed twice in ultrapure water and migration results were digitalized using a ChemiDoc XRS + imaging system controlled by Image Lab software version 6.0.0 (BioRad, France).

MALDI TOF mass spectrometry. MALDI TOF mass spectra on Cry11Aa were acquired on an Autoflex mass spectrometer (Bruker Daltonics, Bremen, Germany) operated in linear positive ion mode. External mass calibration of the instrument, for the *m/z* range of interest, was carried out using as calibrants the monomeric (66.4 kDa) and dimeric (132.8 kDa) ions of bovine serum albumin (reference 7030, Sigma Aldrich). Just before analysis, crystals of Cry11Aa were firstly dissolved in acetonitrile/water mixture (70:30, *v/v*). For samples under reducing condition, DTT was added at a final concentration of 10 mM. The obtained solutions were therefore directly mixed in variable ratios (1:5, 1:10, 1:20, *v/v*) with sinapinic acid matrix (20 mg/mL solution in water/acetonitrile/trifluoroacetic acid, 70:30:0.1, *v/v/v*, Sigma Aldrich) to obtain the best signal-to-noise ratio for MALDI mass spectra. 1 to 2 µL of these mixtures were then deposited on the target and allowed to air dry (at room temperature and pressure). Mass spectra were acquired in the 10 to 160 kDa *m/z* range and data processed with Flexanalysis software (v.3.0, Bruker Daltonics).

MALDI TOF mass spectra on Cry11Ba were collected at the USC Mass Spectrometry Core Facility, Los Angeles, CA, USA. Purified Cry11Ba protein was dissolved in water (~5 mg/mL) and heated at 70 °C for 10 min to facilitate dissolution. One microliter of protein solution was spotted on a 384 Big Anchor MALDI target and let dry at room temperature. Crystallized protein was washed on-target twice with ultrapure water, on top of which 0.5 µL of 2,6 dihydroxyacetophenone (DHAP) solution (30 mg/mL in 50% acetonitrile:0.1% formic acid) was spotted and let dry at room temperature. Crystallized sample was then analyzed using Bruker Rapiflex® MALDI-TOF MS equipped with a Smartbeam 3D, 10 kHz, 355 nm Nd:YAG laser. The laser parameters were optimized as follows: scan range = 26 µm; number of shots per sample = 1000; laser frequency = 5000 Hz. The mass spectrometer was calibrated for high-mass range using Protein A and Trypsinogen standards under Linear Mode. Data were analyzed using FlexAnalysis software and plotted using Graphpad Prism.

In-gel digestion and peptide mass fingerprinting of Cry11Aa using MALDI. Selected bands were in-gel digested with trypsin as previously described[104]. MALDI mass spectra of the tryptic peptides were recorded on an Autoflex mass spectrometer (Bruker Daltonics, Bremen, Germany) in the reflectron positive ion mode. Before analysis samples were desalted and concentrated on RP-C18 tips (Millipore) and eluted directly with 2 µL of α-cyano-4-hydroxy cinnamic acid matrix (10 mg/mL in water/acetonitrile/trifluoroacetic acid: 50/50/0.1, *v/v/v*) on the target.

In-gel digestion and peptide mass fingerprinting of Cry11Ba using GeLC-MS/MS. Gel Liquid Chromatography tandem mass spectrometry spectra collected on Cry11Ba were acquired on a ThermoFisher Q-Exactive Plus (UCLA Molecular Instrumentation Center, Los Angeles, CA, USA). Before analysis, the Cry11Ba crystals were diluted at a 1:5 dilution with ultrapure H_2O and 4x SDS Loading Buffer Dye. These samples were then boiled for 3 min at 98 °C and were loaded on a 4–12% Bis-Tris SDS-PAGE gel. Protein embedded in gel bands were extracted and digested with 200 ng trypsin at 37 °C overnight. The digested products were extracted from the gel bands in 50% acetonitrile/49.9% H_2O/ 0.1% trifluoroacetic acid (TFA) and desalted with C18 StageTips prior to analysis by tandem mass spectrometry. Peptides were injected on an EASY-Spray HPLC column (25 cm × 75 µm ID, PepMap RSLC C18, 2 µm, ThermoScientific). Tandem mass spectra were acquired in a data-dependent manner with a quadrupole orbitrap mass spectrometer (Q-Exactive Plus Thermo Fisher Scientific) interfaced to a nanoelectrospray ionization source. The raw MS/MS data were converted into MGF format by Thermo Proteome Discoverer (VER. 1.4, Thermo Scientific). The MGF files were then analyzed by a MASCOT sequence database search.

Native mass spectrometry. Crystals of Cry11Aa were centrifuged for 5 min at $5000 \times g$ during the buffer wash and washed twice with ammonium acetate buffer (pH adjusted to 6.4 with acetic acid). Pelleted crystals were then dissolved in ammonium acetate buffer (pH adjusted to 11.5 using ammonium hydroxide).

Gold-coated capillary emitters were prepared as previously described[105]. The sample was analyzed on a Synapt G1 mass spectrometer (Waters Corporation). The instrument was tuned to preserve non-covalent interactions. Briefly, the capillary voltage was set to 1.60 kV, the sampling cone voltage was 20 V, the extraction cone voltage was 5 V, the source temperature was 80 °C, the trap transfer collision energy was 10 V, and the trap collision energy (CE) was set at 30 V. For MS/MS characterization, a particular charge state was isolated in the quadrupole and the complex was dissociated by application of 200 V of CE. The data collected were deconvoluted and analyzed using UniDec[106].

Heat stability and aggregation propensity. The thermal unfolding of Cry11Aa WT and mutants was measured by following changes as a function of temperature (15–95 °C) in tryptophan fluorescence leading to an increase of the F350/F330 ratio. Scattering was also monitored to address aggregation propensity of Cry11Aa WT and of the mutants F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q (Supplementary Fig. 14). All the measurements were performed on a Prometheus NT.48 (Nanotemper) following manufacturer's instructions.

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

Data availability

The data that support this study are available from the corresponding author upon reasonable request. Structures and structure factor amplitudes are available in the PDB databank under accession codes 7QX4 (Cry11Aa WT, pH 7.0), 7QX5 (Cry11Aa Y449F, pH 7.0), 7QX6 (Cry11Aa E583Q, pH 7.0), 7QX7 (Cry11Aa F17Y, pH 7.0), 7QYD (Cry11Ba WT, pH 6.5), 7R1E (Cry11Ba WT, pH 10.4). Raw image files are deposited in the Coherent X-ray Imaging Data Bank (CXIDB) under accession number 190 [https://doi.org/10.11577/1873154]. The source data for Fig. 4 and for Supplementary Figs. 6, 7, 8, 9, 12 and 13, as well as uncropped blot scans for Supplementary Figs. 6, 9 and 11 are provided in a combined Source Data file. Source data are provided with this paper.

References

1. Lacey, L. A. Bacillus thuringiensis serovariety israelensis and Bacillus sphaericus for mosquito control. *J. Am. Mosq. Control Assoc.* **23**, 133–163 (2007).
2. Vachon, V., Laprade, R. & Schwartz, J. L. Current models of the mode of action of Bacillus thuringiensis insecticidal crystal proteins: a critical review. *J. Invertebr. Pathol.* **111**, 1–12 (2012).
3. Delecluse, A., Rosso, M. L. & Ragni, A. Cloning and expression of a novel toxin gene from Bacillus thuringiensis subsp. jegathesan encoding a highly mosquitocidal protein. *Appl. Env. Microbiol* **61**, 4230–4235 (1995).
4. Federici, B. A., Park, H.-W. & Sakano, Y. Insecticidal Protein Crystals of Bacillus thuringiensis. in *Inclusions in Prokaryotes* (ed. Shively, J. M.) 195–236 (Springer, 2006).
5. Boonserm, P., Mo, M., Angsuthanasombat, C. & Lescar, J. Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from Bacillus thuringiensis at a 2.8-Angstrom resolution. *J. Bacteriol.* **188**, 3391–3401 (2006).
6. Boonserm, P., Davis, P., Ellar, D. J. & Li, J. Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications. *J. Mol. Biol.* **348**, 363–382 (2005).
7. Cohen, S. et al. Cyt1Aa toxin: crystal structure reveals implications for its membrane-perforating function. *J. Mol. Biol.* **413**, 804–814 (2011).
8. Gutierrez, P., Alzate, O. & Orduz, S. A theoretical model of the tridimensional structure of Bacillus thuringiensis subsp. medellin Cry 11Bb toxin deduced by homology modelling. *Mem. Inst. Oswaldo Cruz* **96**, 357–364 (2001).
9. Chapman, H. N. et al. Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).
10. Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2**, 246–255 (2015).
11. Boutet, S., Fromme, P. & Hunter, M. S. *X-ray Free Electron Lasers - A Revolution in Structural Biology*. vol. 1 (Springer International Publishing, 2018).
12. Tetreau, G. et al. Serial femtosecond crystallography on in vivo-grown crystals drives elucidation of mosquitocidal Cyt1Aa bioactivation cascade. *Nat. Commun.* **11**, 1153 (2020).
13. Li, J. D., Carroll, J. & Ellar, D. J. Crystal structure of insecticidal delta-endotoxin from Bacillus thuringiensis at 2.5 A resolution. *Nature* **353**, 815–821 (1991).
14. Morse, R. J., Yamamoto, T. & Stroud, R. M. Structure of Cry2Aa Suggests an Unexpected Receptor Binding Epitope. *Structure* **9**, 409–417 (2001).
15. Grochulski, P. et al. Bacillus thuringiensis CrylA(a) Insecticidal Toxin: Crystal Structure and Channel Formation. *J. Mol. Biol.* **254**, 447–464 (1995).

16. Engilberge, S. et al. Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem. Sci.* **8**, 5909–5917 (2017).

17. Engilberge, S. et al. Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722–731 (2019).

18. Xu, C., Wang, B. C., Yu, Z. & Sun, M. Structural insights into Bacillus thuringiensis Cry, Cyt and parasporin toxins. *Toxins* **6**, 2732–2770 (2014).

19. Promponas, V. J. et al. CAST: an iterative algorithm for the complexity analysis of sequence tracts. *Bioinformatics* **16**, 915–922 (2000).

20. Wootton, J. C. & Federhen, S. Statistics of local complexity in amino acid sequences and sequence databases. *Comput. Chem.* **17**, 149–163 (1993).

21. Liang, M. et al. The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *J. Synchrotron Radiat.* **22**, 514–519 (2015).

22. Sierra, R. G. et al. Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).

23. DePonte, D. P. et al. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys. -Appl. Phys.* **41**, 7 (2008).

24. Leaver-Fay, A. et al. Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).

25. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. *Nucleic Acids Res.* **27**, 49–54 (1999).

26. Pardo-Lopez, L., Soberon, M. & Bravo, A. Bacillus thuringiensis insecticidal three-domain Cry toxins: mode of action, insect resistance and consequences for crop protection. *FEMS Microbiol. Rev.* **37**, 3–22 (2013).

27. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

28. Colletier, J.-P. et al. De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature* **539**, 43–47 (2016).

29. Oestergaard, J., Ehlers, R. U., Martinez-Ramirez, A. C. & Real, M. D. Binding of Cyt1Aa and Cry11Aa toxins of Bacillus thuringiensis serovar israelensis to brush border membrane vesicles of Tipula paludosa (Diptera: Nematocera) and subsequent pore formation. *Appl. Environ. Microbiol.* **73**, 3623–3629 (2007).

30. Yamagiwa, M., Sakagawa, K. & Sakai, H. Functional Analysis of Two Processed Fragments of Bacillus thuringiensis Cry11A Toxin. *Biosci. Biotechnol. Biochem.* **68**, 523–528 (2014).

31. Munoz-Garay, C. et al. Oligomerization of Cry11Aa from Bacillus thuringiensis Has an Important Role in Toxicity against Aedes aegypti. *Appl. Environ. Microbiol.* **75**, 7548–7550 (2009).

32. Fernandez, L. E. et al. Cry11Aa toxin from Bacillus thuringiensis binds its receptor in Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514 (2005).

33. Carmona, D. et al. Dominant Negative Phenotype of Bacillus thuringiensis Cry1Ab, Cry11Aa and Cry4Ba Mutants Suggest Hetero-Oligomer Formation among Different Cry Toxins. *PLoS ONE* **6**, e19952 (2011).

34. Du, J. P., Knowles, B. H., Li, J. & Ellar, D. J. Biochemical characterization of Bacillus thuringiensis cytolytic toxins in association with a phospholipid bilayer. *Biochem. J.* **338**, 185–193 (1999).

35. Ntountoumi, C. et al. Low complexity regions in the proteins of prokaryotes perform important functional roles and are highly conserved. *Nucleic Acids Res.* **47**, 9998–10009 (2019).

36. Lanzarotti, E., Defelipe, L. A., Marti, M. A. & Turjanski, A. G. Aromatic clusters in protein–protein and protein–drug complexes. *J. Cheminformatics* **12**, 30 (2020).

37. Wiedorn, M. O. et al. Megahertz serial crystallography. *Nat. Commun.* **9**, 4025 (2018).

38. Grünbein, M. L. et al. Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nat. Commun.* **9**, 3487 (2018).

39. Henrich, B. et al. The adaptive gain integrating pixel detector AGIPD a detector for the European XFEL. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **633**, S11–S14 (2011).

40. Likitvivatanavong, S. et al. Multiple Receptors as Targets of Cry Toxins in Mosquitoes. *J. Agric. Food Chem.* **59**, 2829–2838 (2011).

41. Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A. & Soberon, M. A GPI-anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in Aedes aegypti larvae. *Biochem. J.* **394**, 77–84 (2006).

42. Bourgouin, C., Delecluse, A. & Rapoport, G. Specificity and Synergism of Bacillus thuringiensis israelensis Toxins on Mosquito Larvae. in *Bacterial Protein Toxins* (ed. Fehrenbach) 225–226 (Gustav Fisher, 1988).

43. Lopez-Diaz, J. A., Emiliano Canton, P., Gill, S. S., Soberon, M. & Bravo, A. Oligomerization is a key step in Cyt1Aa membrane insertion and toxicity but not necessary to synergize Cry11Aa toxicity in Aedes aegypti larvae. *Environ. Microbiol.* **15**, 3030–3039 (2013).

44. Pérez, C. et al. Bacillus thuringiensis subsp. israelensis Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. USA* **102**, 18303–18308 (2005).

45. Sawaya, M. R. et al. Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proc. Natl. Acad. Sci.* **111**, 12769–12774 (2014).

46. Galitsky, N. et al. Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of Bacillus thuringiensis. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1101–1109 (2001).

47. Adalat, R., Saleem, F., Crickmore, N., Naz, S. & Shakoori, A. R. In vivo crystallization of three-domain Cry Toxins. *Toxins Basel* **9**, (2017).

48. Hughes, M. P. et al. Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. *Science* **359**, 698–701 (2018).

49. Salinas, N., Colletier, J. P., Moshe, A. & Landau, M. Extreme amyloid polymorphism in Staphylococcus aureus virulent PSMα peptides. *Nat. Commun.* **9**, 3512 (2018).

50. Sun, Y. et al. Construction and characterization of the interdomain chimeras using Cry11Aa and Cry11Ba from Bacillus thuringiensis and identification of a possible novel toxic chimera. *Biotechnol. Lett.* **36**, 105–111 (2014).

51. Fernandez, L. E. et al. Cry11Aa toxin from Bacillus thuringiensis binds its receptor in Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514 (2005).

52. Perez, C. et al. Bacillus thuringiensis subsp israelensis Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. USA* **102**, 18303–18308 (2005).

53. Perez, C. et al. Bacillus thuringiensis ssp israelensis Cyt1Aa enhances activity of Cry11Aa toxin by facilitating the formation of a pre-pore oligomeric structure. *Cell. Microbiol.* **9**, 2931–2937 (2007).

54. Carmona, D. et al. Dominant negative phenotype of bacillus thuringiensis Cry1Ab, Cry11Aa and Cry4Ba mutants suggest hetero-oligomer formation among different cry toxins. *PLoS ONE* **6**, e19952 (2011).

55. Likitvivatanavong, S., Aimanova, K. G. & Gill, S. S. Loop residues of the receptor binding domain of Bacillus thuringiensis Cry11Ba toxin are important for mosquitocidal activity. *FEBS Lett.* **583**, 2021–2030 (2009).

56. Gevorkov, Y. et al. XGANDALF - extended gradient descent algorithm for lattice finding. *Acta Crystallogr. Sect. Found. Adv.* **75**, 694–704 (2019).

57. White, T. A. Processing serial crystallography data with CrystFEL: a step-by-step guide. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 219–233 (2019).

58. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

59. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

60. Cowtan, K. Completion of autobuilt protein models using a database of protein fragments. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 328–335 (2012).

61. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 355–367 (2011).

62. Banneville, A.-S. et al. Structural and functional characterization of DdrC, a novel DNA damage-induced nucleoid associated protein involved in DNA compaction, *Nucleic Acids Res.* (2022) https://doi.org/10.1093/nar/gkac563.

63. Park, H. W., Delecluse, A. & Federici, B. A. Construction and characterization of a recombinant Bacillus thuringiensis subsp israelensis strain that produces Cry11B. *J. Invertebr. Pathol.* **78**, 37–44 (2001).

64. Wu, D. & Federici, B. A. Improved production of the insecticidal CryIVD protein in Bacillus-thuringiensis using CryLA(c) promoters to express the gene for an associated 20-KDa protein. *Appl. Microbiol. Biotechnol.* **42**, 697–702 (1995).

65. Kirmitzoglou, I. & Promponas, V. J. LCR-eXXXplorer: a web platform to search, visualize and share data for low complexity regions in protein sequences. *Bioinformatics* **31**, 2208–2210 (2015).

66. Girish, V. & Vijayalakshmi, A. Affordable image analysis using NIH Image/ImageJ. *Indian J. Cancer* **41**, 47 (2004).

67. Sakano, Y., Park, H. W., Bideshi, D. K., Ge, B. & Federici, B. A. Contributions of 5'-UTR and 3'-UTR cis elements to Cyt1Aa synthesis in Bacillus thuringiensis subsp. israelensis. *J. Invertebr. Pathol.* **149**, 66–75 (2017).

68. Lereclus, D., Arantès, O., Chaufaux, J. & Lecadet, M. Transformation and expression of a cloned delta-endotoxin gene in Bacillus thuringiensis. *FEMS Microbiol Lett.* **51**, 211–217 (1989).

69. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Cent. Eur. J. Phys.* **10**, 181–188 (2012).

70. Chen, S. W. & Pellequer, J.-L. DeStripe: frequency-based algorithm for removing stripe noises from AFM images. *BMC Struct. Biol.* **11**, 7 (2011).

71. Waterman, D. G. et al. Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. Sect. Struct. Biol.* **72**, 558–575 (2016).

72. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1274–1282 (2013).

73. Hattne, J. et al. Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nat. Methods* **11**, 545–548 (2014).

74. Winter, G. et al. DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 85–97 (2018).

75. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).

76. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

77. Coquelle, N. et al. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallogr. D. Biol. Crystallogr.* **71**, 1184–1196 (2015).

78. Foucar, L. et al. CASS—CFEL-ASG software suite. *Comput. Phys. Commun.* **183**, 2207–2213 (2012).

79. Skubák, P. & Pannu, N. S. Automatic protein structure solution from weak X-ray data. *Nat. Commun.* **4**, 1–6 (2013).

80. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 (2011).

81. Duisenberg, A. J. M. Indexing in single-crystal diffractometry with an obstinate list of reflections. *J. Appl. Crystallogr.* **25**, 92–96 (1992).

82. Ginn, H. M. et al. ıt TakeTwo: an indexing algorithm suited to still images with known crystal parameters. *Acta Crystallogr. Sect. D.* **72**, 956–965 (2016).

83. Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 48–57 (2006).

84. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for macromolecular crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 143–151 (2018).

85. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D. Biol. Crystallogr.* **64**, 61–69 (2008).

86. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Biol. Crystallogr* **60**, 2126–2132 (2004).

87. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).

88. Uervirojnangkoorn, M. et al. Enabling X-ray free electron laser crystallography for challenging biological systems from a limited number of crystals. *eLife* **4**, e05421 (2015).

89. Blanc, E. et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2210–2221 (2004).

90. *The PyMOL Molecular Graphics System*. (Schrödinger, LLC.).

91. Bond, C. S. & Schüttelkopf, A. W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr Biol. Crystallogr* **65**, 510–512 (2009).

92. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).

93. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Biol. Crystallogr* **60**, 2256–2268 (2004).

94. Gille, C. & Frömmel, C. STRAP: editor for STRuctural Alignments of Proteins. *Bioinformatics* **17**, 377–378 (2001).

95. Gille, C., Fähling, M., Weyand, B., Wieland, T. & Gille, A. Alignment-Annotator web server: rendering and annotating sequence alignments. *Nucleic Acids Res.* **42**, W3–W6 (2014).

96. Boc, A., Diallo, A. B. & Makarenkov, V. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res.* **40**, W573–W579 (2012).

97. Mirdita, M., Steinegger, M., Breitwieser, F., Söding, J. & Levy Karin, E. Fast and sensitive taxonomic assignment to metagenomic contigs. *Bioinformatics* **37**, 3029–3031 (2021).

98. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).

99. Prakash, A., Jeffryes, M., Bateman, A. & Finn, R. D. The HMMER Web Server for Protein Sequence Similarity Search. *Curr. Protoc. Bioinforma.* **60**, 3.15.1–3.15.23 (2017).

100. Potter, S. C. et al. HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).

101. Savi, M. K., Mangamana, E. T., Deguenon, J. M., Hounmenou, C. G. & Kakaï, R. G. Determination of Lethal Concentrations Using an R Software Function Integrating the Abbott Correction. *J. Agric. Sci. Technol. A* **7**, 25–30 (2017).

102. Wheeler, M. W., Park, R. M. & Bailer, A. J. Comparing median lethal concentration values using confidence interval overlap or ratio tests. *Env. Toxicol. Chem.* **25**, 1441–1444 (2006).

103. R. R: *a language and environment for statistical computing*. (R Foundation for Statistical Computing, 2011).

104. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **203**, 173–179 (1992).

105. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).

106. Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

Author contributions

G.T., A.-So.B., D.B., E.L., F.L., L.D., H.-W.P., B.F. designed and constructed and transformed WT and mutants plasmids; G.T., E.A.A., A.-So.B., N.A.S., D.B., N.Z., H.-W.P. and B.F. produced crystals in vivo; L.S., E.J.C. and M.M.A. performed MALDI-TOF mass spectrometry experiments; P.Q. and A.L. performed native MS/MS mass spectrometry experiments; G.T., E.A.A. and N.A.S. performed solubilization assays; E.A.A. performed heat stability assays; N.A.S., M. B., W.L.L. and I.G. conducted transmission electron microscopy imaging; G.T., E.A.A., A.-So.B., N.A.S., R.S. and I.S. performed crystal visualization by SEM; G.T., J.-M.T., D.F. and J.-L.P. performed crystal visualization and size measurements by AFM; G.T., J.-L.P. and E.D.Z. performed the statistical analysis of solubilization data and G.T. and J.-L.P. performed the statistical analyses of AFM data; A.-So.B., M.B., M.W. and J.-P.C. secured beamtime at the ESRF for crystal screening; M.R.S., N.S., J.R., B.F., and D.C. secured beamtime at the APS for crystal screening; M.R.S., A.-So.B., M.S.H., J.R., M.W., N.K.S., B.F., D.C., I.S., J.-P.C. secured beamtime at the LCLS for data collection; I.S. and J.P.C. secured beamtime at the EuXFEL for data collection; S.E., E.G., A.R., C.C., F.R. and O.M. synthesized Tb-Xo4; G.T. derivatized Cry11Aa crystals for injection at LCLS; G.T., M.R.S., E.A.A., A.-So.B. and N.A.S. prepared crystals for data collection at XFEL and synchrotrons; R.G.S. developed and operated the MESH-on-a-stick injector; R.L.S. and R.B.D. developed the GDVN injector; E.A.A., G.S., M.G., G.N.-K., M.K., G.S., M.S., R.L.S. and R.B.D. operated the GDVN injector; G.T., M.R.S., E.A.A., A.-So.B., N.A.S., A.S.B., M.G., G.N.-K., M.S.H., M.K., R.G.S., G.S., M.S., I.D.Y., A.G., A.B., S.B., T.M.R.S., J.R., R.L.S., R.B.D., M.W., N.K.S., D.C., I.S. and J.-P.C. performed serial data collection at the LCLS; G.T., E.A.A., A.-So.B., N.C., M.G., G.N.-K., M.S.H., M.K., R.G.S., G.S., A.G., M.H., L.F., J.B., R.B., R.W., T.S., H.K., R.L., A.M., T.R.M.B., R.L.S., R.B.D., I.S. and J.-P.C. performed serial data collection at the EuXFEL; E.D.Z., N.C., A.S.B., I.D.Y. and N.K.S. produced new processing tools or devices; E.D.Z., N.C., A.S.B., A.G., I.D.Y., N.K.S. and J.-P.C. performed serial data processing; E.D.Z. and J.P.C. phased the structural data; M.R.S., E.D.Z., N.A.S. and J.-P.C. performed atomic model building, refinement and structure interpretation; G.T., M.R.S., E.D.Z. and J.-P.C. prepared figures and tables and wrote the manuscript with input from E.A.A., A.-So.B., N.A.S., A.S.B., M.L.G., D.B., S.E., L.S., R.S., W.L.L., J.-L.P., A.L., R.L.S., D.C. and I.S; J.-P.C. designed and coordinated the project; E. A. A., A.-S. B. and N. A. S. contributed equally.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-31746-x.

Correspondence and requests for materials should be addressed to Jacques-Philippe Colletier.

Peer review information *Nature Communications* thanks Mike Hough, Azadeh Shahsavar and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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.

[1]Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, 71 Avenue des martyrs, F-38000 Grenoble, France. [2]UCLA-DOE Institute for Genomics and Proteomics, Department of Biological Chemistry, University of California, Los Angeles, CA 90095-1570, USA. [3]Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany. [4]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA. [5]Large-Scale Structures Group, Institut Laue-Langevin, F-38000 Grenoble, France. [6]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [7]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA. [8]European XFEL GmbH, Holzkoppel 4, 22869 Schenefeld, Germany. [9]Department of Chemistry, Texas A&M University, College Station, TX 77845, USA. [10]Department of Entomology and Institute for Integrative Genome Biology, University of California, Riverside, CA 92521, USA. [11]Department of Biological Sciences, California Baptist University, Riverside, CA 92504, USA. [12]European Synchrotron Radiation Facility (ESRF), BP 220, 38043 Grenoble, France. [13]Mass Spectrometry Core Facility, School of Pharmacy, University of Southern California, Los Angeles, CA 90089, USA. [14]Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, CA 90089, USA. [15]Univ. Grenoble Alpes, CNRS, LECA, F-38000 Grenoble, France. [16]Univ. Lyon, ENS de Lyon, CNRS UMR 5182, Université Claude Bernard Lyon 1, Laboratoire de Chimie, F-69342 Lyon, France. [17]Polyvalan SARL, 15 parvis René Descartes, 69342 Lyon, France. [18]These authors contributed equally: Guillaume Tetreau, Michael R. Sawaya, Elke De Zitter. ✉email: colletier@ibs.fr