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

Authors:

Guillaume Tetreau[1,18], Michael R. Sawaya[2,18], Elke De Zitter[1,18], Elena A. Andreeva[1,3,19], Anne-Sophie Banneville[1,19], Nathalie Schibrosky[2,4,19], 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*].

Abstract (150 words)

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.

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.

[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

[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.

[7]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025.

[8]European XFEL GmbH, Holzkoppel 4, 22869, Schenefeld, Germany.

[9]Department of Chemistry, Texas A&M University, College Station, TX 77845.

[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, California 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, F69342 Lyon, France.

[17]Polyvalan SARL, 15 parvis René Descartes, 69342 Lyon

[18]These authors contributed equally

[19]These authors contributed equally

[*]Correspondance : colletier@ibs.fr (J-P.C.)

64 **Introduction (797 words)**

65 The most commonly used biological insecticide for controlling mosquito vector
66 populations is produced by the bacterium *Bacillus thuringiensis* subsp. *israelensis* (*Bti*) [1]. Its
67 highly potent mosquitocidal activity is due to three nanocrystalline forms of four protoxins, viz.
68 Cyt1Aa, Cry11Aa and co-crystallized Cry4Aa and Cry4Ba. These are produced during
69 sporulation and are remarkably stable in a variety of conditions, but dissolve after ingestion
70 under the high alkaline pH levels characteristic of the larval mosquito midgut [2]. Solubilized
71 protoxins are activated by insect gut proteases enabling binding to gut cell membranes,
72 subsequent oligomerization, and ultimately gut cell lysis leading to larval death [2]. *Bti* toxins are
73 environmentally safe because they are much more specific for target mosquitoes than broad-
74 spectrum chemical larvicides.

75 The most potent of the four *Bti* toxins is Cry11Aa, but it is poorly understood, in large
76 part because unlike Cry4Aa, Cry4Ba, and Cyt1Aa, its structure is unknown. A related toxin
77 produced by *Bt* subsp. *jegathesan* (*Btj*) is Cry11Ba, which is from seven to thirty-seven times
78 more toxic than Cry11Aa against major mosquito vector species belonging to the genera
79 *Aedes*, *Anopheles*, and *Culex* [3], and in some bacterial hosts appears to form slightly larger
80 crystals. Cry11Ba's structure is also unknown, although it has been used in recombinant
81 strains of *Bti* to improve mosquito activity significantly [3,4]. Thus, our goal was to determine the
82 structures of Cry11Aa and Cry11Ba protoxins to help understand their mechanisms of
83 crystallization that result in environmental stability and which could possibly yield structural
84 insights for increasing the efficacy of these proteins for mosquito control. Structure
85 determination of Cry11Aa and Cry11Ba protoxins from natural nanocrystals requires cutting-
86 edge technology. Conventional crystallography is limited to projects in which crystals are
87 sufficiently large to mount and oscillate individually in a synchrotron X-ray beam. In the past,
88 crystals of activated Cry4Aa [5], Cry4Ba [6] and Cyt1Aa [7] attained sufficient size by growing these
89 *in vitro* from toxins dissolved from natural nanocrystals and activating the toxins enzymatically.
90 However, Cry11Aa and Cry11Ba do not recrystallize *in vitro* from dissolved nanocrystals [8].
91 Moreover, enzymatic activation is unwanted since our goal is to understand the pH-controlled
92 mechanism of natural crystal dissolution. To observe the protoxin state in natural nanocrystals
93 produced in bacterial cells, we applied serial femtosecond crystallography (SFX) at X-ray free
94 electron lasers (XFEL) [9–11]. In the SFX experiment, high brilliance XFEL beam pulses, each
95 lasting only ~10-50 fs, intercept a series of nanocrystals, one pulse-per-crystal, eliciting the
96 strongest possible diffraction signal from each tiny crystal before it vaporizes, and producing a
97 series of diffraction snapshots, later assembled into a full data set. Feasibility of this strategy
98 had been demonstrated by the recent elucidation of the full bioactivation cascade of Cyt1Aa
99 [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 ten structurally-determined members of the three-domain Cry δ-endotoxin family [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 2.3 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 explains 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 cascade and enable development of new, rational strategies for improved mosquito control.

128 **Results (4407 words)**

129 *De novo phasing of Cry11Aa and Cry11Ba structures by SFX*

130 *In vivo*-grown crystals of Cry11Aa and Cry11Ba protoxins exhibit distinct morphologies,
131 which initially concealed a surprising conservation of their crystal packing patterns. Cry11Aa
132 crystallizes as hexagonal plates and Cry11Ba crystallizes as larger bipyramidal crystals (Fig.
133 1 a,b) as reported earlier [4]. These morphological distinctions cannot be attributed to differences
134 in crystallization mechanisms in their parent organisms, *Bti* and *Btj*, since both protoxins were
135 recombinantly produced in the same host organism, an acrystalliferous strain of *Bti* (4Q7).
136 Cry11Aa and Cry11Ba protoxins are expected to share structural resemblance to each other
137 since the two sequences share 54% identity; however, 46% non-identity at the molecular level
138 could easily produce large differences at the macroscopic level of crystal morphology.
139 Moreover, the sequence of Cry11Ba is extended by 77-residues at its C-terminus, potentially
140 also affecting differences in crystal packing (Supplementary Fig. 1). Interestingly, this
141 extension has been identified as a low complexity region (LCR) by both CAST [19] and SEG [20]
142 computational methods, which implicates the extension in the mechanism of crystal nucleation.
143 At this point in our studies, the balance of evidence suggested that sequence divergence was
144 likely to have erased the crystal packing pattern that early ancestors of today's Cry11Aa and
145 Cry11Ba presumably once shared.

146 Our diffraction experiments yielded the first hint that Cry11Aa and Cry11Ba shared a
147 conserved crystal packing pattern. We collected diffraction data from Cry11Aa and Cry11Ba
148 nanocrystals injected in the vacuum chamber of the CXI-SC3 micro-focused beamline at the
149 Stanford Linear Accelerator Center (SLAC) Linac Coherent Light Source (LCLS) [21] using a
150 microfluidic electrokinetic sample holder (MESH) [22] (Cry11Ba crystals) or a gas-dynamic virtual
151 nozzle (GDVN) [23] (Cry11Aa crystals). The underlying similarity in the packing of Cry11Aa and
152 Cry11Ba became evident when their diffraction patterns were collected and indexed, revealing
153 similarly sized unit cells (a~58; b~155; c~171 Å; α=β=γ=90°), albeit belonging to two different
154 space groups: *I*222 and $P2_12_12$, respectively (Table 1). Conservation of unit cell parameters
155 hinted that this crystal packing pattern is special, evolved to perform a function more intricate
156 than just storing protein.

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

gadolinium, gold, platinum, and mercury 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 30h soak at 10 mM concentration), and phases were determined by the single wavelength anomalous dispersion (SAD) method at 2.6 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 TbX04 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. *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 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. Our MR-phase 2.3 Å resolution map reveals two Cry11Ba molecules in the asymmetric unit. All residues are visible except for the N-terminus (residues M1-N16), two loops (residues G330-E340, and D352-I358) and the C-terminal extension (residues T654-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. 1c, Supplementary Fig. 3). Domain I is implicated to form a pore in the target membrane. Like 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 implicated to recognize 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 implicated to oligomerize. It forms a β-sandwich of two antiparallel five-stranded β-sheets (viz. $\beta20 - \beta19 - \beta22 - \beta17 - \beta4 - {}^{\beta12}/_{\beta14}$ and $\beta15 - {}^{\beta13}/_{\beta16} - \beta23 - \beta18 - \beta21$) forming a jelly-roll

201 topology, whereby $^{\beta12}/_{\beta14}$ and $^{\beta13}/_{\beta16}$ are interrupted β-strands contributed by two non-
202 consecutive shorts β-strands, which appose and intercalate one after the other onto β4 and
203 between β15 and β23, respectively (Figure 1d).
204 The closest homolog of known structure to Cry11 toxins is *Bt kurstaki* (*Btk*) Cry2Aa
205 (PDBid: 1i5p), with a sequence identity of 26.6 and 23.6 % and main-chain rmsd of 3.7 and
206 4.0 Å, with respect to Cry11Aa and Cry11Ba, respectively. As with Cry2Aa, the Cry11Aa toxins
207 feature a long insert (27 residues in Cry2Aa; 21 residues in the Cry11 toxins) between strands
208 β10 and β11, which together with domain-I β1, form the third β-sheet of the domain-II β-prism.
209 This insert, which features a short α-helix (α_h) and a β-strand (β_h), folds like a handle, and is
210 therefore referred to as the $\alpha_h\beta_h$-handle, throughout the manuscript (Fig. 1c, Supplementary
211 Fig. 3). The αhβh-handle fastens domain II onto domain III through direct (e.g. in Cry11Aa,
212 D443(OD2)-R502(NH2); D443(O)-R502(NH1); L447(N)-S503(O)) and water mediated H-
213 bonds (T446(OG1)/T448(O)-Wat72(O)-R502(N); T448(OG1)/V499(O)-Wat308(O)-
214 D501(OD1); T448(N)/L447(N)-Wat65(O)-S503(OG)/(O)), and enables the burying of domain-
215 II α8 at an interface formed by $\alpha_h\beta_h$, α6-α7 (domain I), β10-β11 (domain II), β15 and the β13-
216 β14 and β15-β16 loops (domain III), and the α9 helix connecting domain II and domain III
217 (D469-K478 in Cry11Aa) (Supplementary Fig. 4). The firm hold of α8 enables the three
218 domains to be more tightly packed in Cry2Aa and Cry11 toxins than in other Cry toxins (e.g.
219 *Btt* Cry3Aa or *Btk* Cry1Ac). Additionally, strand β_h lays aside strand β4 thereby expanding –
220 and consequently, stabilizing – the first β-sheet of domain II (β_h-β4-β3-β2-β5). Also, alike
221 Cry2Aa, the Cry11 toxins feature a smaller β-prism due to deletions in the second constitutive
222 β-sheet, namely between β7 and β8 (6 and 10 residues missing in Cry2Aa and Cry11 toxins,
223 respectively), and between β9 and β10 (14 and 15 residues missing in Cry2Aa and Cry11
224 toxins, respectively; Supplementary Fig. 3). The Cry11 toxin structures are, however, specific
225 in that a 36 to 38 residue insertion is observed between strands β4 and β5, contributing an
226 additional β-strand to the first β-sheet of domain-II – thereafter referred to as the β_{pin} (Fig. 1c).
227 As the β_{pin} lays along a two-fold axis, two large β_h-β4-β3-β2-β_{pin} – β_{pin}-β2-β3-β4-β_h sheet are
228 formed between symmetry related dimers (AC or BD, interface #3), yielding the crystallizing
229 tetramer (Fig. 2b, e). We noted earlier that the BSA at the tetramerization interface is 33%
230 lower in Cry11Ba, pointing to higher flexibility; this hypothesis is supported by the absence of
231 interpretable electron density for residues at the N-terminus (330-340) and C-terminus (352-
232 360) of the β_{pin} in the Cry11Ba structure. Also noteworthy is that Cry11 toxins feature a
233 conserved N/D-DDLGITT insertion between β21 and β22, and deletions (>3 residues) between
234 α3 and α3 (-5 and -8 residues with respect to *Btk* Cry2Aa and *Bt tenebrionis* (*Btt*) Cry3Aa),
235 and β20 and β21 (-10 and -9 residues with respect to *Btk* Cry2Aa and *Btt* Cry3Aa). Altogether,
236 these changes render Cry11 toxins uniquely large from the structural standpoint, with predicted

237 radii of gyration of 27.5 and 26.7 Å for Cry11Aa and Cry11Ba, compared to 25.0 and 25.6 Å
238 for *Btk* Cry2Aaa and *Btt* Cry3aa, respectively.
239
240 *All domains engage in producing the in vivo crystal lattice.*
241 Examination of packing interfaces reveal that all three domains are involved in the
242 formation and stabilization of Cry11Aa and Cry11Ba nanocrystals. The *in vivo* crystallization
243 pathway can be best trailed from Cry11Aa crystals, which feature a single monomer per
244 asymmetric unit and build on six packing interfaces burying a cumulated surface area (BSA)
245 of 3514.5 $Å^2$, corresponding to 13.1 % of the total protein area. The main building block of
246 Cry11Aa crystals consists of a tetramer with a total BSA of 9663.0 $Å^2$ and a predicted binding
247 energy of -12.5 $kcal.mol^{-1}$ at pH 7 by PISA [27] (Fig. 2a,b). Supported by two of the six packing
248 interfaces, the tetramer builds from the cross-association (AC or BD; interface #3) of two
249 dimers (AB or CD) along the 2-fold axis contributed by domain II (Fig. 2b). Each dimer is itself
250 composed of monomers associated along another 2-fold axis contributed by domain III and by
251 strand β4 and the β10-β11 hairpin (P433-P457) in domain II. The tetramer is further stabilized
252 by a minor interface between apices from domain II (interfaces AD or BC; interface #6).
253 Crystals grow from the piling in a honeycomb brick-wall fashion of such tetramers, as a result
254 of a face-to-back interaction between domains I from symmetry-related molecules (interface
255 #2; Fig. 2c). Cry11Aa crystals are further cemented by two additional minor interfaces. The
256 first involves the apex of the second β-sheet of domain II (interface #5) from facing monomers
257 in each dimer (AD or BC) of the stable tetramer. The second occurs between the α3-α4 loop
258 of domain I in one tetramer and the apex of the second β-sheet of domain II in another tetramer
259 (interface #4).
260 The similarity between the packing of Cry11Aa and Cry11Ba crystals makes it
261 reasonable to propose that the latter also forms from the assembly of tetramers (Fig. 2d, e, f),
262 despite failure of PISA to identify the tetramer as a (meta-)stable building block for Cry11Ba
263 crystals. In these, two molecules are found in the asymmetric unit, associated through the face-
264 to-back interface between domains described above for Cry11Aa monomers (interface #2; Fig.
265 2f). The BSA at this interface is 1429.0 $Å^2$, *i.e.* 53% higher than in Cry11Aa (Fig. 2g). However,
266 BSAs at the interface formed by domains III and II, which associates monomers into a dimer
267 is 38% lower than the homologous interfaces in Cry11Aa (Fig. 2e, g). Thus, the intermolecular
268 contact between piled tetramers is larger in the Cry11Ba crystals, despite an overall looser
269 packing of monomers in the crystals, with an average BSA at crystal contacts of 3385.1 $Å^2$ per
270 monomer, corresponding to 12.6 % of the total protein area. The increased BSA between
271 tetramers (contributed by the large face-to-back interface between domains I), and the
272 presumably higher flexibility in the Cry11Ba tetramers, could be at the origin of Cry11Ba
273 packing into larger three-dimensional crystals. Regardless, our structures evidence that each

274 domain exhibits a dual role in Cry11 toxins, namely the formation and stabilization of *in vivo*-
275 grown nanocrystals, and execution of a domain specific function. The latter comprises pore
276 formation (domain I), receptor-recognition and membrane-insertion (domain II), and
277 oligomerization and stabilization of the toxic pore conformation (domain III) [26] .
278
279 *Drastic conformational changes drive crystal dissolution*
280 We sought to characterize the conformational changes that ensue pH elevation,
281 preceding dissolution of the crystals in the mosquito larvae gut [28]. As the crystals are naturally
282 labile at pH 11, we aimed at collecting data from crystals soaked at a lower pH, hypothesizing
283 that early conformational changes would show but the crystal packing still hold. In the case of
284 Cry11Aa crystals, diffraction quality was decreased dramatically at pH values of 9.5 (CAPS
285 buffer, glycerol 30%) and above, preventing collection of a sufficiently large number of
286 diffraction patterns to produce a high-pH dataset. Hence, large conformational changes occur
287 in Cry11Aa at pH as low as 9.5, opposing diffraction quality, despite crystals dissolving as of
288 pH 11 only (Fig. 3a). In the case of Cry11Ba, ~3 Å diffraction was preserved up to pH 10.4
289 (Table 1). Comparison between the refined 'pH10.4' and 'pH6.5' structures points to large inter-
290 domain rearrangements induced by pH increase. Detailed analysis of structural changes at
291 the side chain level was yet prevented by the limited resolution of the 'pH10.4' dataset. A 1%
292 unit cell contraction, and hence tighter crystal packing, was observed in the 'pH10.4' crystals
293 in comparison to the pH 6.5 crystals. However, because a higher glycerol concentration was
294 used for injection of Cry11Ba crystals at pH 10.4, we cannot exclude that unit cell contraction
295 might be caused by crystal dehydration.
296
297 *Crystals are made of full-sized monomers of Cry11 toxins*
298 In both Cry11Aa and Cry11Ba toxins, the β_{pin} (residues E339-Q350 and I341-Y350,
299 respectively) is a ~10-residue long β-strand that hydrogen-bonds with a two-fold related
300 symmetry mate, contributing the interface that assembles dimers (AC and BD) into stable
301 tetramers. This strand is bordered on each side by the only two loops that have disordered
302 electron density in Cry11Ba (missing residues G330-E340 and D352-I358) and are
303 comparatively difficult to interpret in Cry11Aa (F330-D334 and Q350-E355), respectively. As
304 Cry11Aa N335-Y349 and Cry11Ba I341-N351 regions match the enzymatic cleavage site
305 known to generate the two activated fragments of ~32 and ~36 kDa [29,30] upon proteolytic
306 activation in the mosquito larvae gut, we asked whether disorder in the F330-D334 (G330-
307 E340) and Q350-E355 (D352-I358) loops serves the purpose of enabling facilitated access of
308 proteases to Cry11Aa (Cry11Ba) cleavage sites or if each monomer occurs in natural crystals
309 as two polypeptide chains cleaved prior or during crystal formation. SDS PAGE analysis of
310 Cry11Aa (12% gels, heating at 95°C for 5 min, presence of DTT and SDS; Supplementary Fig.

311 5) resulted in a major band ~70 kDa, in line with previous reports [31–33]. As the denaturing
312 treatment would have broken any disulfide-bridge or non-covalent interactions that could
313 maintain cleaved fragments together, this result suggests that Cry11Aa occurs in crystals as
314 a full monomer. We further verified this hypothesis by use of MALDI TOF mass spectrometry.
315 In MALDI mass spectra collected after direct solubilization of the natural crystals in sinapinic
316 acid matrix in presence or absence of DTT, we observed main peaks at *m/z* of 72246 and
317 72235 (mass error: ± 100 Da) and 36154 and 36129 Da, respectively, in agreement with
318 expected molecular masses for singly- and doubly- charged ions of a full-size monomer
319 (expected mass: 72.349 kDa) (Uniprot accession number: P21256; Supplementary Fig. 6).
320 However, because proteolytic activation is as well expected to yield a 36 kDa fragment, in
321 addition to a 32 kDa fragment for which a minor peak was present in the MALDI-TOF mass
322 spectra, we resorted to native mass spectrometry to assert that the ~72.240 and ~36.140 kDa
323 peaks originated from the same species – rather than being indicative of the crystallization of
324 proteolytic products. With this approach, we could confirm that upon dissolution of Cry11Aa
325 crystals, a 72.345 kDa fragment is released, corresponding to the full-size monomer
326 (Supplementary Fig. 7a). Moreover, both incubation of solubilized toxin at room temperature
327 (RT) for 2 h (Supplementary Fig. 7b) and use of increased collision energy (Supplementary
328 Fig. 7c, d) failed at yielding a signature for the two polypeptides that would have been
329 generated if cleavage at position 329 had occurred. We conclude that natural crystals of
330 Cry11Aa, and possibly Cry11Ba, grow from the addition of full-size monomers, and that
331 disorder in the F330-D334 (G330-E340) and Q350-E355 (D352-I358) loops could serve the
332 purpose of enabling facilitated access of proteases to Cry11Aa (Cry11Ba) cleavage sites.
333 Considering proteinase K as a surrogate analogue for mosquito larvae gut proteases [34], one
334 would expect the β_{pin} to be released upon proteolytic activation, suggesting that the role of the
335 latter is to promote *in vivo* crystallization. We note that other cleavage sites are predicted,
336 which would release the first six residues and last two β-strands (β22-β23), as well as rescind
337 the covalent association between domain I and domains II and III, thereby leaving non-covalent
338 interactions surfaces as the sole glue between them.
339
340 *Mutagenesis to alter crystal formation and dissolution*
341 We proposed earlier that the packing of Cry11Ba into slightly larger crystals than
342 Cry11Aa could stem from differences in the extent and nature of the interfaces which support
343 dimerization, tetramerization and piling of tetramers into crystals (Fig. 2). Considering recent
344 evidence linking LCR regions with diverse functions including chaperoning [35] and reversible
345 oligomerization, we further asked whether or not presence of the 77-residue LCR region at the
346 C-terminus of Cry11Ba plays a complementary role in the promotion of crystal formation. A
347 chimera was therefore designed, coined C11AB, wherein the LCR region of Cry11Ba was

348 fused to the C-terminal end of Cry11Aa (Material and Methods; Supplementary Fig. 8a).
349 C11AB was produced at the expected size but at a lower yield than Cry11Aa WT
350 (Supplementary Fig. 8b). Atomic force micrographs (AFM) revealed the presence of multiple
351 needle-like inclusions in the parasporal envelope enchasing the crystals, suggesting that
352 presence of Cry11Ba-LCR at the C-terminal end of Cry11Aa favors nucleation, but not crystal
353 growth (Supplementary Fig. 8c).

354 Seven Cry11Aa mutants were additionally designed with the aim to probe the
355 involvement of Cry11Aa intra- and inter-molecular interfaces in toxin stability, crystal formation
356 and dissolution. Each mutant was designed to challenge a specific interface and served as a
357 coarse proxy to evaluate its pH sensitivity and putative participation in the crystal dissolution
358 mechanism. First, we asked whether the intra-chain stabilization of α8 at an interface
359 contributed by the three domains (namely, $\alpha_h\beta_h$, α6-α7, α9, β10, β11, β15 and the β13-β14
360 and β15-β16 loops) could play a role in crystal dissolution. Residues central to this interface
361 are Y272, D514 and D507, which H-bond to one another and to Y203, R222, T249, S251
362 through direct and water-mediated interactions (W253 and W267), connecting the three
363 domains (Supplementary Fig. 9a, Supplementary Table 1a). Upon pH elevation, Y272, D514
364 and D507 are all expected to be deprotonated, which should result in electrostatic repulsion
365 and thence dissociation of the three domains. To test the hypothesis, we produced three
366 Cry11Aa mutants intended to eliminate pH sensitivity of the above-described H-bonds. Neither
367 did the Y272Q nor D507N-D514N mutations impact the overall stability of the toxin, in the
368 soluble or crystalline form (Fig. 3b), but their combination in the triple Y272Q-D507N-D514N
369 mutant resulted in an unexpected abolishment of the ability of Cry11Aa to form crystals *in vivo*
370 – possibly due to improper folding (Supplementary Fig. 10). The Y272Q mutation had no effect
371 on the pH sensitivity of Cry11Aa crystals, while only a minor effect was seen with the
372 D507N+D514N mutant (Fig. 3a). Thus, pH-induced deprotonation of amino acids involved in
373 the stabilization of α8 at the interface between the three domains does not play a role in the
374 initial steps of crystal dissolution, possibly because of their deep burial at the interface. We
375 note that the above-mentioned residues and their interactions are all strictly conserved in
376 Cry11Ba (viz. Y273, D518, D511, Y203, R222, T249, S251, W253 and W268).

377 We then focused on Cry11Aa E583, a residue siting at the intramolecular interface
378 between domain I and domain III. This β21 residue, condemned to be anionic at higher pH,
379 takes part in the water-mediated hydrogen bond network that connects α6 and α7 from domain
380 I with domain III (Supplementary Fig. 9b, Supplementary Table 1b). We therefore asked
381 whether or not suppression of the pH-sensitivity of the network would stabilize the monomer
382 at high pH, thereby reducing the pH sensitivity of the crystals. This was indeed the case, with
383 an SP_{50} (pH at which 50% of crystals are dissolved) of 12.6 ± 1.0 for crystals of the E583Q
384 mutant, compared to 11.2 ± 1.0 for WT Cry11Aa crystals (Fig. 3a), and a dissolution profile

385 characterized by a reduced slope with no visible plateau up to pH 14. Thus, the alteration of
386 protonation state of residues and water molecules at the intramolecular interface between
387 domain I and domain III may be involved in the early step of Cry11Aa crystal dissolution. In
388 Cry11Ba (G587), which displays a similar SP50 of 11.5 (Supplementary Fig. 11), this residue
389 is substituted for glycine suggesting a different mechanism of pH-induced intramolecular
390 separation of domain I and domain III, in Cry11Ba – or at least the involvement of additional
391 residues at the interface.

392 Crystal contacts were also investigated. We first tampered with the interface enabling
393 the piling in a honey-comb fashion of Cry11Aa tetramers (Fig. 2c, interface #2), by introducing
394 a F17Y substitution, intended to induce electrostatic repulsion with the negatively charged
395 D180 (distance D180(OD1) - F17(CZ) of 3.3 Å), due to deprotonation of its hydroxyl group
396 upon pH increase (Supplementary Fig. 9c). As expected, crystals of the F17Y mutant were
397 found to be more sensitive to increases in pH, with crystals starting to dissolve at pH as low
398 as ~9.5 and an SP_{50} of 10.6 ± 1.0 (Fig. 3a). The dissolution profile of F17Y crystals is again
399 characterized by a reduced slope, as compared to WT crystals, explaining that the plateau is
400 nonetheless reached at the same pH (~pH 11.6). Nevertheless, the result suggests that
401 dissolution of Cry11Aa crystals can be accelerated by separation of the tetramers associated
402 through interface #2. The F17Y mutation was also found to challenge crystal formation,
403 yielding crystals far smaller than their WT counterparts. We note that F17, D180 and the H-
404 bond between them are strictly conserved in Cry11Ba; hence, the importance of interface 2 for
405 crystal formation and dissolution could be extendable to crystals formed by Cry11Ba.

406 Next, we challenged the role of the dimerization interface (Fig. 2b interface #1). Recall
407 that BSA at this interface, contributed by domain III from facing monomers, is 38% lower in
408 Cry11Ba than in Cry11Aa. Furthermore, only 8 hydrogen bonds and 2 salt bridges support the
409 interface in Cry11Ba, compared to 20 hydrogen bonds and 10 salt bridges in Cry11Aa. Y449
410 is positioned in the central part of the interface, and while not involved in direct H-bonding to
411 other protein residues, supports a large H-bond network that interconnects waters and
412 residues from facing monomers in the dimer (Supplementary Fig. 9d, Supplementary Table
413 1c). Hence, we investigated whether deprotonation of Y449 in the middle of the interface would
414 significantly affect crystal dissolution by engineering of a Y449F mutation. Only a minor effect
415 on crystal dissolution was observed (Fig. 3a), yet the mutation was detrimental to the protein
416 stability (Fig. 3b), resulting in the growth of crystals of different size and shape (Fig. 3c).

417 Finally, we introduced a Y349F mutation in the β_{pin}, hypothesizing that suppression of
418 its pH-sensitive H-bond to E295(OE1) in the adjacent strand β2 would disturb the β_{pin} fold and
419 destabilize the tetramerization interface (Fig. 2b interface #3, Supplementary Fig. 9e,
420 Supplementary Table 1d), thereby triggering crystal dissolution. This expected effect was not
421 observed, with crystals of the mutant displaying the same pH-induced dissolution profile as

422 those of the WT. Nonetheless, smaller crystals were observed whose thermal stability was
423 affected (Figure 3 and Supplementary Fig. 12), indicating that reduced stabilization of the turn
424 preceding the β_{pin} not only impacts folding and stability of the toxin, but as well its piling into
425 crystals – probably due to reduced tetramerization. Of note, Y349 is conserved in Cry11Ba
426 where it H-bonds to P362(O).
427 Of all the single and double mutants we investigated, the Y349F mutation is that which
428 results in the smallest crystals, closely followed by F17Y and E583Q. The Y449F mutant,
429 however, exhibits the most noticeable change in shape compared to WT Cry11Aa. To evaluate
430 the significance of these changes, we characterized the distribution in size of crystals of
431 Cry11Aa-WT, Y449F, F17Y and E583Q using AFM (Fig. 3d). All three mutants had a
432 significantly reduced volume compared to WT Cry11Aa, due to a reduced thickness of the
433 crystals (Fig. 3d).
434
435 *Probing crystalline order of the Cry11Aa mutants by SFX*
436 The presence of crystals does not necessarily infer that molecules are well arranged
437 within them. We therefore used SFX to assess the level of crystalline order in crystals of the
438 mutants that displayed modified solubilization or shape. Data were collected at the SPB/SFX
439 beam line of the EuXFEL (Hamburg, Germany) from crystals delivered across the X-ray beam
440 using a liquid microjet focused through a gas-dynamic virtual nozzle GDVN [23] (Table 2). All
441 crystals were kept in water at pH 7 for the GDVN injection, and pulses were delivered at the
442 MHz repetition rate (1.1 MHz) [89,89,90] using 10 Hz trains of 160 pulses, with a spacing of 880 ns
443 apart. Data was collected on the AGIPD detector at its maximum rate of 3.52 kHz [36]. With the
444 notable exception of Y349F, crystals of all four single point mutants diffracted, yet unequal
445 amounts of data were collected from each, and none from WT crystals, due to technical
446 difficulties that arose during the experiment. This impeded a thorough comparison of the
447 diffraction power of the various mutants, and prevented structure determination for the Y272F
448 mutant. The structures of the other three mutants were determined, using the WT structure as
449 a molecular replacement model for the phasing of diffraction data. We found that neither overall
450 packing, tertiary structure nor interface formation is affected in the tested mutants at neutral
451 pH (Supplementary Fig. 13). Of important note, these data demonstrate the feasibility of
452 macromolecular nano-crystallography at MHz pulse rate using the brilliant micro-focused
453 beam available at the SPB/SFX beamline of the EuXFEL.
454 The needle shape inclusions formed by C11AB were also investigated by SFX and
455 found to present some crystalline order, as evidenced by diffraction rings up to ~6 Å resolution
456 in the powder diagram calculated from the maximum projection of 395656 hits (Supplementary
457 Fig. 8d). 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.

462 **Discussion (2219 words)**

463 We here report the previously-unknown structures of Cry11Aa and Cry11Ba, the two
464 most potent Cry δ-endotoxins expressed by mosquitocidal *Bti* and *Btj*, respectively. Both toxins
465 occur as natural nanocrystals that are produced during the sporulation phase of the bacteria,
466 and dissolve upon elevation of pH in the mosquito larvae gut. Proteolytic activation enables
467 binding to their specific receptors [37], including a membrane embedded alkaline phosphatase[38]
468 but as well the co-delivered Cyt1Aa [12,39–41], triggering insertion in gut cell membranes and
469 subsequent oligomerization into pores that will eventually kill the cells. Both toxins are of
470 industrial interests due to their environmental safety, explained by the multi-step activation
471 outlined above, and to their high stability as crystals. Our results shed light on the mechanisms
472 of *in vivo* crystallization, pH-induced dissolution and proteolytic activation, and on structural
473 features that support the toxins specificity with respect to other Cry toxins. Thereby, our work
474 offers a foundation for further improvement of the toxic activity or crystal size by rational design.
475 Additionally, we demonstrate the feasibility of *de novo* structure determination of a previously-
476 unknown protein-structure by SFX, from nanocrystals only 10,000 unit-cells across, using a
477 single caged-terbium (TbXo4) derivative. Below, we recapitulate these findings and discuss
478 their implications.

479

480 In vivo *crystallization pathway of Cry11 toxins*

481 The building block of Cry11Aa and Cry11Ba crystals is a tetramer formed by the interaction of
482 two dimers, via their domain II. The dimers are themselves assembled from the interaction of
483 two monomers, via their domains II and III. Crystals form from the honey-comb brick-wall piling
484 of tetramers, as enabled by the face-to-back interaction of domain I from symmetry-related
485 tetramers (Figure 2). Thus, all three domains are involved in the *in vivo* crystal packing of
486 Cry11 toxins, each contributing a two-fold axis. This observation contrasts with other toxin
487 structures determined from *in vivo* grown crystals, wherein either propeptide(s) (e.g.
488 *Lysinibacillus sphaericus* BinAB [28] and *Bti* Cyt1A [12]) or a specific domain (e.g. domain I in *Btt*
489 Cry3Aa from [42,43]) serves as the major contributor to crystallization. Expanding to previously
490 determined Cry δ-endotoxins [12,28,42,44] structures, solved from *in vitro* grown macrocrystals
491 obtained following dissolution of the natural crystals at high pH, the same trend is observed –
492 *i.e.*, crystallization mostly depends on a dedicated portion of the protein, either it be a N-
493 terminal and/or C-terminal propeptide (e.g., the ~650 C-terminal residues in *Btk* Cry1Ac) or a
494 specific domain (e.g. domain II in *Btk* Cry2Aa). Thus, the Cry11Aa and Cry11Ba structures
495 illustrate a yet unobserved pathway for *in vivo* crystallization, wherein all domains act on a
496 specific step of the coalescence process, viz. dimerization (domains II and III from two Cry11
497 monomers), tetramerization (domains II from two Cry11 dimers) and tetramer-piling (domains
498 I in each tetramer). With Cry11Aa featuring a larger dimerization interface, and Cry11Ba a

499 larger interface between piled tetramers, the two structures underline different levels of tradeoff
500 between packing *into* tetramers and packing *of* the tetramers,
501 The difference in thickness of Cry11Aa and Cry11Ba crystals is of interest. Considering that
502 all crystals were produced in *Bti*, we could exclude the possibility that the slightly larger size of
503 Cry11Ba crystals originates from a more efficient crystallization machinery in *Btj* than *Bti*.
504 Puzzled by the presence of a 77-residue long low complexity region at the C-terminus of
505 Cry11Ba (LCR-Cry11Ba), which is absent in Cry11Aa, we asked whether or not a C-terminal
506 fusion of LCR-Cry11Ba with Cry11Aa would result in larger crystals. LCR regions have indeed
507 been shown to support a variety of functions, including chaperoning [35] and reversible
508 oligomerization [45,46] so that a role in crystal nucleation and/or growth could not be excluded.
509 Support of the first, but not the second hypothesis was obtained. Indeed, the C11AB chimera,
510 consisting of a fusion of LCR-Cry11Ba to the C-terminus of Cry11Aa, yields smaller crystals
511 that poorly diffract, even when exposed to high intensity XFEL pulses. This observation is in
512 line with previous results which showed that substitution of Cry11Ba domain III for that of
513 Cry11Aa leads to limited expression and comparatively small inclusions [47]. Thus, the LCR
514 region of Cry11Ba is unlikely to account for the difference in size between Cry11Aa and
515 Cry11Ba crystals. Instead, we favor the hypothesis that it is the larger surface of interaction
516 between piled tetramers that accounts for the larger size of the Cry11Ba crystals. Given the
517 absence of electron density for LCR-Cry11Ba residues in the Cry11Ba structure, and the
518 abundance of needle-like inclusions in the parasporal body enveloping the C11AB crystals, it
519 is reasonable to assume that they do not engage in structurally important interactions with
520 functional domains, but rather favor nucleation of crystals. This aid-to-nucleation would be
521 required for Cry11Ba, which features a reduced dimerization interface, but not for Cry11Aa,
522 wherein this interface is 62 % larger. In line with this hypothesis, four regions are predicted to
523 form short adhesive motifs of the Low Complexity, Amyloid-like Reversible Kinked Segments
524 (LARKS) type (Supplementary Fig. 3).
525
526 *Cry11 toxins depart from the canonical Cry δ-endotoxins architecture*
527 The structures of Cry11Aa and Cry11Ba shed light on features that would not have
528 been predicted based on sequence alignments (*i.e.*, by homology modelling), and which
529 largely deviate from the canonical organization observed in other Cry δ-endotoxins [12,28,42,44].
530 The most notable difference is the presence of a ~36 to 38 residue insertion between strands
531 β4 and β5 in domain II, which results in an extra β-strand, coined β_{pin}. The β_{pin} not only
532 participates in the formation of a modified β-prism, but contributes to a two-fold axis that
533 supports tetramerization of Cry11 toxins through formation of two large β_h-β4-β3-β2-β_{pin} – β_{pin}-
534 β2-β3-β4-β_h sheets between symmetry-related dimers into a tetramer. The observation of
535 proteolytic cleavage sites at both the N- and C-termini of the β_{pin} suggests that it is removed

536 upon activation by mosquito gut proteases, in line with the observation of ~32 and ~36 kDa
537 fragments upon proteolytic activation of the Cry11 toxins [32]. If true, the unique role of the β_{pin}
538 would be to support *in vivo* crystallization and its removal would entail the dissociation of
539 tetramers into dimers and eventually monomers. While mutagenesis results indicate that this
540 interface does not play a major role in crystal dissolution (see below), it seems likely that upon
541 pH elevation and deprotonation of tyrosines and acidic groups, electrostatic repulsion will occur
542 between Y349(OH) and E295(OE2) in Cry11Aa, and between Y350(OH) and P362(O) in
543 Cry11Ba. Increased disorder of these regions could facilitate the access of proteases, and thus
544 favor the activation of the Cry11Aa and Cry11Ba toxins. This hypothesis would rationalize the
545 reluctance of the two toxins to recrystallize *in vitro* after pH induced dissolution, due to an
546 impossibility to re-form tetramers – or at least, to re-match the exact positioning of the β-pin.
547 The Cry11 toxins also differ from other Cry δ-endotoxins by the presence of a conserved N/D-
548 DDLGITT insertion between β21 and β22, contributing a short helix, and by deletions of ~5-10
549 residues in the α3-α4 and β20-β21 loops, respectively. Compilation of these changes likely
550 explains failures to phase the Cry11 structures by the molecular replacement method, even
551 when *Btk* Cry2Aa, which also features a $\alpha_h\beta_h$-handle, was used as a starting model.
552
553 *Mapping the interfaces involved in crystal dissolution.*
554 Our efforts to determine the structures of Cry11Aa and Cry11Ba at alkaline pH were
555 unsuccessful, due to high sensitivity of crystals diffraction quality to pH increase. In the case
556 of Cry11Aa we could not collect data, while in the case of Cry11Ba, we obtained a low-
557 resolution structure which, while showing possible inter-domain rearrangements, did not inform
558 on specific side chain rearrangements. Therefore, we resorted to site-specific mutagenesis to
559 obtain information regarding the crystal dissolution pathway. We found that the crystal interface
560 most sensitive to pH elevation is the one enabling the honey-comb piling of Cry11 tetramers,
561 with the Cry11Aa-F17Y mutant displaying increased pH sensitivity (with an SP_{50} of 10.6 ± 1.0
562 compared to 11.2 ± 1.0 for WT Cry11Aa crystals). In contrast, the dimerization (Y349F mutant)
563 and tetramerization interfaces (Y449F mutant) appear to be less pH-sensitive. At the monomer
564 level, we found that the three-domain interface to which α_8 and the $\alpha_h\beta_h$-handle contribute is
565 not very sensitive to pH increase (Y272Q and D507N+D514N mutants), possibly due to
566 burying of mutated residues at the interface, preventing bulk solvent to access these sites.
567 Alternatively, interaction of Cry11 toxins with its membrane-bound receptors [37] could be a
568 required step to expose α_8, shown to play a major role in binding and toxicity [48].
569 The intramolecular domain I vs. domain III interface was found to be important for the
570 pH-induced crystal dissolution, with the Cry11A E583Q mutant displaying a reduced sensitivity
571 to pH (SP_{50} of 12.6 ± 1.0). Yet unlike the other tested interfaces, which are overall well
572 conserved, the domain I vs. domain III interface differs in Cry11Aa and Cry11Ba, suggesting

573 that caution is advised upon reflecting results obtained from Cry11Aa mutants onto Cry11Ba.
574 Indeed, E583 is substituted for glycine in Cry11Ba (G587), suggesting a different mechanism
575 of pH-induced separation of domain I and domain III – or at least, the participation of other
576 residues. Structural analysis suggests that the substitution of Cry11Ba Q247 for a glutamic
577 acid could compensate for the absence of E583, enabling electrostatic repulsion of V494 (β14)
578 – found at the opposed end of this interface – upon pH elevation. Numerous other residues at
579 this interface, otherwise mostly conserved between Cry11Aa and Cry11Ba, remain as
580 candidates to further tune the pH sensitivity. For example, Y241(OH) is H-bonded to
581 D586(OD1; 2.6 Å) and D590(OD2; 2.8 Å) in Cry11Aa and Cry11Ba, respectively, suggesting
582 that mutation of this residue into a phenylalanine (Y241F) and/or of D586/D590 into
583 asparagines would reduce the pH sensitivity while not affecting stability. Likewise, E234 H-
584 bonds to Q625(NE2; 2.6 Å) in Cry11Aa, and to K629(NZ; 2.8 Å) and R553(NH1; 2.9 Å) in
585 Cry11Ba, suggesting that a E234Q mutation would reduce pH sensitivity in the two toxins
586 whilst not affecting their folding. Inversely, the mutation into a glutamic acid of Q511/Q515,
587 squeezed between a tryptophan (W584/W588), an arginine (R549/R553) and a glutamic acid
588 (E234), would be expected to increase the pH sensitivity of the domain I vs. domain III
589 intramolecular interface in both Cry11Aa/Cry11Ba – and by extension, that of their crystals.
590
591 *Implication for the future of nanocrystallography using SFX.*
592 In this study, *de novo* phasing was required – not because of the absence of
593 homologous structures, but because none of those available were sufficiently close to serve
594 as a search model for molecular replacement. Using Tb-Xo4, a caged terbium compound, we
595 could phase the *Bti* Cry11Aa structure by SAD, from ~77,000 diffraction patterns collected on
596 crystals only 10,000 unit cell across – an achievement to compare to the determination of the
597 structure Ls BinAB from > 370,000 patterns (native and three derivatives) collected on crystals
598 100,000 unit cell across [49]. Our success in phasing the Cry11Aa structure stemmed from a
599 combination of advances in SFX data processing tools over the last five years and the use of
600 a dramatically powerful phasing agent, and should offer hope to investigators seeking to
601 determine the structure of proteins of which no known structural homologue exists and that
602 have to resort to SFX due to smallness of their crystals. It is foreseeable, however, that *de-*
603 *novo* structure determination will be helped by recent advances in comparative and *ab-initio*
604 modelling and the availability of programs such as RosettaFold [50] and AlphaFold2 [51], capable
605 of producing a decently-accurate structure for virtually all proteins and thus a good model for
606 phasing of crystallographic data by molecular replacement. Latest releases of the two
607 programs were published in the final stage of the writing of this manuscript, hence we asked
608 whether or not the availability of these tools would have facilitated our journey towards the
609 Cry11 toxins structures, and submitted the sequence of Cry11Aa to the two servers. For

610 RosettaFold, the rmsd to the final refined structure of the five best models was over 4 Å, with
611 discrepancies observed mostly in domain II. For AlphaFold2, however, the two first models
612 displayed rmsd of 1.2 and 1.0 Å to the final structure, respectively. Using the worst of these
613 two models, we could find a molecular replacement solution using Phaser, and a partial model
614 featuring 95% of the residues in sequence was obtained after 20 cycles of automatic iterative
615 model-building and refinement using Bucanneer [52] and Refmac [53]. Thus, a problem which
616 occupied a handful of crystallographers for several years could have been solved in less than
617 an hour using the new tools recently made available to the structural biology community. Based
618 on our results, it is tantalizing to claim that the phase problem in crystallography has been
619 solved, or that experimental structural biology has lived, but such assertions would likely be
620 shortsighted. Rather, we encourage investigators to challenge AlphaFold2 and RosettaFold
621 as much as humanly feasible, but to not forsake *de novo* phasing as it may remain the only
622 route to success in difficult cases where molecular replacement based on such models does
623 not work [54]. It must also be emphasized that in the case of Cry11 toxins and, more generally,
624 naturally-crystalline proteins, the issue is not just phasing, but packing. For such proteins,
625 crystal formation and dissolution serve function, hence characterization of packing interfaces
626 is central to finely comprehend their bioactivation cascades. Without the naturally-occurring
627 crystals and the atomic resolution SFX structures, it would not have been possible to make
628 predictions on potential mutations affecting Cry11Aa crystal formation or dissolution.
629
630
631

632 **Materials and Methods (4643 words)**

633 **Crystal production and purification.** Crystals of Cry11Aa and Cry11Ba were produced by
634 electrotransformation of the plasmids pWF53 and pPFT11S [55] into the acrystalliferous strain
635 4Q7 of *Bacillus thuringiensis* subsp. *israelensis* (Bti; The Bacillus Genetic Stock Center
636 (BGSC), Columbus OH, USA), respectively [56]. Colonies were selected on LB agar medium
637 supplemented with erythromycin (25 µg/mL) and used to inoculate precultures of LB liquid
638 medium. For Cry11Aa production, precultures were spread on T3 sporulation medium. After
639 incubation at 30°C for 4 days, spores/crystals suspensions were collected using cell scrapers
640 and resuspended in ultrapure water. After sonication-induced cell lysis and subsequent
641 centrifugation at 4000 g for 45 min to discard cell and medium debris, pellets were
642 resuspended in water and crystals were purified using a discontinuous sucrose gradient (67-
643 72-79 %). After ultracentrifugation, crystals were recovered and several rounds of
644 centrifugation/resuspension in ultrapure water allowed discarding as much sucrose as possible
645 for proper downstream application. Crystal purity was verified by SDS-PAGE on 12% gels.
646 Purified crystals were conserved in ultrapure water at 4 °C until use. For Cry11Ba, a glycerol
647 stock of the 4Q7/pPFT11S was streaked onto 25 µg/mL erythromycin Nutrient Agar plates.
648 From here a single colony was selected and added to a Glucose-Yeast-Salts (GYS) media
649 culture and allowed to grow continuously at 30°C, 250 rpm for 5 days. This culture was then
650 spun down, resuspended in ultrapure water, and the lysate was sonicated for 3 min at 50%
651 duty. The sonicated lysate was added to the 30-65% discontinuous sucrose gradient (35-40-
652 45-50-55-60-65 %) and spun down for 70 min at 20,000 rpm and 4°C. The sucrose gradient
653 was then hand fractionated with Cry11Ba crystals collected around 57-60% and dialysed into
654 ultrapure water. Crystal characterization and purity was completed by phase contrast light
655 microscopy, X-ray powder diffraction, transmission electron microscopy, and 4-12% SDS-
656 PAGE gels. The pure Cry11Ba crystals were stored at 4°C in ultrapure water.

657

658 **Cry11Aa mutagenesis.** Based on the SFX structure of Cry11Aa, a total of 7 mutants of
659 Cry11Aa were constructed to challenge the different crystal packing and intramolecular
660 interfaces. The rationale behind these mutations is illustrated in Supplementary Fig. 9 and
661 discussed in the main text. Point-mutations were inserted into *cry11aa* gene by Gibson
662 assembly using pWF53 as a backbone [56]. Two different primer couples were used for each
663 mutation to amplify two fragments that were complementary by their 15-18 bp overlapping 5'
664 and 3' overhangs with a target Tm of 50°C. Point mutations were inserted in the
665 complementary part of the overhangs of the two fragments spanning the *cyt1aa* region to be
666 mutated. The double mutant D507N-D514N was successfully constructed in a single-step by
667 respectively adding the D507N mutation on the non-overlapping overhang region of the
668 forward primer, and the D514N on the non-overlapping overhang of the reverse one. The triple

669 mutant Y272Q-D507N-D514N was constructed by using the primers containing the Y272Q
670 mutation and the plasmid pWF53-D507N-D514N as a backbone. In addition to the point
671 mutants, a Cry11Aa-Cry11Ba chimeric toxin – coined C11AB – was also constructed. For this,
672 the sequence of the *cry11aa* gene was fused with the 234 bp extra 3' extension of *cry11ba*
673 gene, which is suggested to feature a low complexity region (LCR) based on sequence
674 analysis using the LCR-eXXXplorer web platform
675 (http://repeat.biol.ucy.ac.cy/fgb2/gbrowse/swissprot) [57], which implements the CAST [19] and
676 SEG [20] computational methods to identify LCR. The C11AB chimera was constructed by
677 Gibson assembly following a "1 vector, 2 fragments" approach. The plasmid pWF53 containing
678 the *cry11aa* gene was used as a backbone and the *cry11ba* 3' fragment was amplified from
679 the extracted and purified plasmid of the WT strain of *Btj* containing the *cry11ba* gene. The list
680 of primers used for plasmids construction is available in Supplementary Table 2. For each
681 plasmid construction, the fragments with overlapping overhangs were assembled using the
682 NEBuilder HiFi DNA Assembly (New England BioLabs) as previously described [12]. Briefly, after
683 90 min incubation at 50°C, the constructed plasmids were transformed by heat shock into
684 chemically competent Top10 *Escherichia coli* (New England BioLabs). Plasmids were
685 extracted from colonies selected on LB agar medium containing ampicillin (100 µg mL^{-1}) using
686 the NucleoSpin Plasmid extraction kit (Macherey-Nagel) following the manufacturer's
687 instructions. The successful construction of each plasmid was assessed by double digestion
688 (EcoRI and BamHI) followed by migration on 1% agarose gel stained with SYBR Safe
689 (Invitrogen) and by Sanger sequencing of the region containing the mutation at the Eurofins
690 Genomics sequencing platform. Of note, the *cry11aa* gene was also fully sequenced to
691 validate its sequence for mutagenesis primer design and for comparing the expected toxin size
692 to the observed one in mass spectrometry analyses. All mutants were produced as crystals in
693 *Bt*, as described above. The presence of the mutated *cry11aa* gene sequence in the
694 transformed *Bt* colony used for production was verified by colony PCR using specific primers
695 and Sanger sequencing at the Eurofins Genomics sequencing platform. Crystals from all
696 mutants were analyzed by SDS-PAGE on 12% gels. For C11AB, its proper size was confirmed
697 by using the "gel analysis" module implemented in the software ImageJ v1.51k (N = 7) [58].

698 **Crystal visualization by scanning electron microscopy (SEM).** Purified crystals of Cry11Aa
699 WT and of the 7 mutants were visualized using either a Zeiss LEO 1530 scanning electron
700 microscope from the SEM facility of the European Synchrotron Radiation Facility (ESRF,
701 Grenoble, France), a Thermo Fisher Quanta 650 FEG environmental SEM (ESEM) available
702 for users at the European XFEL (EuXFEL, Hamburg, Germany) or a JEOL JSM-6700M FE-
703 SEM (UCLA, Los Angeles, USA). For SEM at ESRF, samples were coated with a 2 nm thick
704 gold layer with the Leica EM ACE600 sputter coater before imaging. For ESEM at the

705 European XFEL, samples were diluted (1/1000) and mixed with 25 mM of ammonium acetate.
706 Samples were then coated with a thin gold layer as described above using a Leica EM ACE600
707 sputter coater as well. Images were recorded at 15 kV acceleration voltage by collecting
708 secondary electrons using an Everhart-Thornley-Detector (ETD detector) in high-vacuum
709 mode. For SEM at UCLA, samples were diluted (1/5) and ultrapure H2O.they were then added
710 to 300 mesh Cu F/C grids that were positively glow discharged. These samples were then
711 wicked away and washed with ultrapure water, wicked, and allowed to dry overnight to ensure
712 all moisture had evaporated inside of a dessicator. These were then attached to a holder with
713 carbon tape and coated with an Anatech Hummer VI sputter coater with approximately 2 nm
714 of thick gold layer. Images were recorded at 5 kV acceleration voltage by collecting secondary
715 electrons using a Lower secondary electron (LEI) or Upper secondary electron in-lens (SEI)
716 detector.

717 **Crystal visualization by transmission electron microscopy (TEM).** Non-purified crystals of
718 Cry11Aa WT were visualized using a Thermofisher TF20 electron microscope from the IBS
719 electron microscopy platform. For negative staining TEM, samples were diluted 5 times in H_2O
720 and 4 µl of the diluted sample was introduced to the interface of an amorphous carbon film
721 evaporated on a mica sheet. The carbon film was then floated off the mica sheet in ~200 µl
722 2% sodium silicotungstate (SST) solution. The carbon film with the crystal sample was then
723 recovered onto a Cu 300 mesh TEM grid after 30 s, let dry, and imaged at 200 keV. Images
724 were recorded on a Gatan OneView CMOS detector. Non-purified crystals of Cry11Ba WT
725 were visualized using an FEI Tecnai T12 electron microscope within the UCLA California
726 Nanoscience Institute, EICN facility. For negative staining TEM, samples were prepared by
727 adding 5 µL of pure crystal fractions in 10 µL ultrapure H2O. 2.5 µL of this sample was added
728 to 300 mesh Cu F/C grids that were positively glow discharged. These samples were then
729 wicked away using Whatman 1 filter paper; washed with 2.5 µL ultrapure H2O, wicked; and
730 negatively stained with 2.5 µL 2% uranyl acetate, wicked. These were allowed to dry overnight
731 to ensure all moisture had evaporated and imaged at 120 keV. Images were recorded on a
732 Gatan 2kX2k CCD.

733 **Crystal characterization by atomic force microscopy (AFM).** Crystals of Cry11Aa were
734 visualized by AFM as previously described[12]. Briefly, 5 µL of crystals suspended in ultrapure
735 water were deposited on freshly cleaved mica. After 30 min in a desiccation cabinet (Superdry
736 cabinet, 4% relative humidity), crystals were imaged on a Multimode 8, Nanoscope V (Bruker)
737 controlled by the NanoScope software (Bruker, Santa Barbara, CA). Imaging was done in the
738 tapping mode (TAP) with a target amplitude of 500 mV (about 12 nm oscillation) and a variable
739 setpoint around 70% amplitude attenuation. TESPA-V2 cantilevers (k = 42 Nm^{-1}, Fq = 320

740 kHz, nominal tip radius = 7 nm, Bruker probes, Camarillo, CA, USA) were used and images
741 were collected at ~1 Hz rate, with 512- or 1024-pixel sampling. Images were processed with
742 Gwyddion [59], and if needed stripe noise was removed using DeStripe [60]. Measurements were
743 performed on Cry11Aa WT and on mutants selected on the basis of their aspect in eSEM
744 images (Y449F) or their solubilization pattern (F17Y and E583Q). Size measurements were
745 performed on AFM images using Gwyddion [59] in a semi-automated protocol. A classical height
746 threshold was applied to each image to select as many individual crystals as possible.
747 Sometimes, partially overlapping crystals were individualized using the manual edition of the
748 mask of selected crystals by adding a separation line. Finally, a filter was applied to remove
749 very small selections (artefacts) or crystals touching the edge of the image. Measures were
750 obtained using the 'distribution of grains' feature in Gwyddion where the crystal thickness (T)
751 is the returned mean value, the volume (V) is the Laplacian background basis volume, and the
752 length and width are the major and minor semi-axes of equivalent ellipses, respectively. The
753 total number of crystals measured are: 45 for WT, 93 for F17Y, 60 for Y449F, and 94 for
754 E583Q.
755
756 **Data collection history.** The Cry11Aa/Cry11Ba structure determination project was initiated
757 in 2015. Data were collected at five different occasions, in two XFEL facilities, namely at the
758 Linac Coherent Light Source (LCLS), Stanford (USA) and EuXFEL, Hamburg (Germany).
759 During our first LCLS-SC3 beamtime (cxi04616), we collected data from native Cry11Ba (2.3
760 Å resolution), and in our second (LO91), we collected data from native Cry11Aa (2.8 Å
761 resolution). Nanocrystals grown by recombinant expression in the modified acrystalliferous
762 4Q7 strain of *Bti* were injected by a microfluidic electrokinetic sample holder (MESH) device[61]
763 in the microfocus chamber of LCLS-SC3 [62]. Crystals were. After data reduction using cctbx.xfel
764 and dials (hit-finding through merging) [63–66], we attempted phasing of both datasets by
765 molecular replacement (MR), using sequence-alignment based multi-model approaches
766 implemented in Mr Bump (based on MR by Molrep [67]) as well as custom-scripts testing models
767 produced by Rosetta [68] (using the Robetta server; http://robetta.bakerlab.org/) and SwissProt
768 [25] (https://www.ebi.ac.uk/uniprot/) servers (based on MR by Phaser [69]). Failure to find a
769 homolog of a sufficiently-close structure led us to attempt de novo phasing of the Cry11
770 nanocrystalline proteins. Initially, we aimed at obtaining experimental phases for Cry11Ba,
771 considering that its larger crystals would produce a stronger diffraction signal which in turn
772 would facilitate phasing. Hence, we collected derivative data on Cry11Ba, from crystals soaked
773 with Gd, Pt and Au salts (P127 experiment) before injection using a MESH device[61].
774 Unfortunately, the data did not allow phase determination, as indicated by very weak and
775 absent peaks in the isomorphous and anomalous difference maps, respectively
776 (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] (P125 experiment). Crystals were injected using a GDVN [23] liquid microjet in the microfocus chamber of LCLS-SC3 [62]. Online data processing was performed using NanoPeakCell [70] and CASS [71]. Offline data processing with NanoPeakCell [70] (hit finding) and CrystFEL [72] (indexing and merging) revealed a strong anomalous signal that enabled determination of the substructure and phasing of the SFX data, using Crank2 [73] and its dependencies in the CCP4 suite [74] (see below for more details). The Cry11Aa structure was thereafter used to phase the Cry11Ba datasets by molecular replacement, revealing *a posteriori* that the Gd, Pt and Au ions had successfully bound to the crystalline Cry11Ba in the various derivatives collected during P127, despite anomalous and isomorphous signals being too weak to enable phasing. 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 [70] (hit finding) and CrystFEL [72] (indexing and merging).

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 hours with Tb-Xo4 at 10 mM in water, prior to GDVN injection [23]. Of 558747 images collected using the 5 µm beam available at the at LCLS-SCC, a total of 77,373 images were indexed of which 76,687, 292, 217 and 177 using Xgandalf [75], Dirax [76], taketwo [77] and Mosflm [78], respectively, in CrystFEL v.0.8.0 [79]. 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. A native dataset was also collected and processed in the same fashion yielding, from 792623 collected patterns of which 48652 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 [73] and its dependencies (ShelxC, ShelxD, Solomon, Bucanneer, Refmac5, Parrot) in CCP4 Online [80], the FOM was 0.52 after density modification,

814 and rose to 0.88 upon building of 613 residues. This first model was characterized by R_{work}/R_{free}
815 of 27.7/32.1 % and was further improved by automatic and manual model building in
816 phenix.autobuild [81] and Coot [82] until 630 residues were correctly build. This model was then
817 used to phase the native data. Final manual rebuilding (using Coot [82]) and refinement (using
818 phenix.refine [83] and Refmac5 [53]) afforded a native model characterized by R_{work}/R_{free} of
819 17.2/24.1 % and consisting of most of the 643 residues. Only the first 12 N-terminal residues
820 are missing (Table 1).
821
822 Cry11Ba data were collected during the cxi04616 and P141 beamtimes at LCLS-CXI. At both
823 occasions, the photon energy was 9503 eV (i.e., a wavelength of 1.30 Å), a pulse duration of
824 50 fs, a repetition rate of 120 Hz, and a focal size of 1 µm – *i.e.,* a similar standard configuration
825 (pulse length, repetition rate) than that used for Cry11Aa, notwithstanding the beam size and
826 wavelength. Data were collected from crystals at pH 6.5 (30% glycerol in pure water; cxi04616)
827 and pH 10.4 (30% glycerol in 100 mM CAPS buffer; P141), presented to the X-ray beam using
828 a MESH injector [22]. Of 813133 images collected for the pH 6.5 dataset, 19708 were indexed
829 and scaled, post-refined, and merged using cctbx.xfel[63–66] and PRIME[84], yielding a dataset
830 extending to 2.2 Å resolution. The Cry11Aa structure was used as a starting model to phase
831 the Cry11Ba pH 6.5 dataset by molecular replacement using Phaser [69] with initial R_{work}/R_{free}
832 being 34.4/40.4 %. Manual model building (using Coot [82]) and refinement (using Refmac [53]
833 and Buster [85]) afforded a model characterized by R_{work}/R_{free} of 18.9/23.8 % (Table 1). Because
834 the 2.55 Å resolution Cry11Ba pH 10.4 data was of limited utility, in view of absence of major
835 peaks in the Fourier difference map calculated with the pH 6.5 data as a reference, and of a
836 1% change in the unit cell volume ascribable to the use of a different glycerol concentrations
837 during injection of the two samples, it was not included in our PDB and CXIDB depositions.
838 Diffraction data on the Cry11Aa mutants at pH 7.0 was acquired on the SPB/SFX beamline at
839 EuxFEL during our P002545 beamtime allocation, using a GVDN injector and X-ray energy
840 and focal size of 9300 eV (1.33 Å) and 1.3 µm (FWHM), respectively. Technical problems
841 allowed us to collect only a limited number of diffraction pattern of the Cry11Aa-Y349F mutant.
842 3,150,500; 5,993,679 and 3,523,741 images were collected for the F17Y, Y449F and E583Q
843 mutant, respectively, of which 28,227; 104,359 and 21,833 could be processed using
844 CrysFEL0.8.0 [79] and MonteCarlo based scaling and merging. The three structures were solved
845 using MR with Phaser [69], using the Cry11Aa WT structure as input model. The structures were
846 refined using Phenix.refine [83] and Coot [82], with final R_{work}/R_{free} values of 21.2/25.1 % for
847 Cry11Aa-F17Y, 22.4/25.1 % for Cry11Aa-Y449F and 21.5/25.4 % for Cry11Aa-E583Q (Table
848 2).
849

850 **Structure analysis.** Figures were prepared using pymol v. 2.5[86] (Fig. 1, 2 and Supplementary
851 Fig. 4, 9, 13) and aline (Supplementary Fig. 3) [87]. Radii of gyration were predicted using the
852 pymol script rgyrate (https://pymolwiki.org/index.php/Radius_of_gyration). Interfaces were
853 analyzed with PISA [88] and rmsd among structures were calculated using pymol using the
854 'super' algorithm. Sequence based alignment – performed using EBI laglign and ClustalW [89] –
855 was challenged by the large gaps between Bti Cry11Aa, Btj Cry11Ba, Btk Cry2AA and Btt
856 Cry3Aa, while structure-based alignment – performed using SSM [90] – was blurred by the
857 varying size of secondary structure elements in the three domains of the various toxins. Hence,
858 for Supplementary Fig. 1, 3, the alignment of Bti Cry11Aa, Cry11Ba, Cry2AA and Cry3Aa was
859 performed using strap [91] which takes into account both sequence and structural information.
860 Specifically, the online version of the program was used (http://www.bioinformatics.org/strap/)
861 [92].
862
863 **Structure prediction using AlphaFold2 and RosettaFold:** RosettaFold [50] predictions were
864 obtained by submitting the sequence to the Rosetta structure-prediction server
865 (https://robetta.bakerlab.org). AlphaFold2 [51] predictions were obtained by use of the
866 Colaboratory service from Google Research
867 (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_a
868 dvanced.ipynb). The mmseq2 method [93,94] was employed for the multiple-sequence
869 alignement instead of the slower jackhmmer method [95,96] used in [51]. Structural alignments were
870 performed using the *align* tool in PyMOL [86]. Molecular replacements trials were carried out with
871 Phaser [69]. Using the best five RosettaFold models, all characterized by an overall rmsd to the
872 final structure superior to 4 Å, no molecular replacement solution could be found, due to
873 inaccurate prediction of domain II β_{pin} and α_h-β_h regions, resulting in clashes. The best
874 alphafold2 model was yet successful at predicting the domain II structure, which enabled
875 successful phasing by molecular replacement, yielding a model characterized by R_{free} and R_{work}
876 values of 0.322 and 0.292, respectively. This model was further use as a starting model for
877 automatic model building and refinement using the buccaneer pipeline in CCP4, resulting in a
878 model characterized by R_{free} and R_{work} values of 0.245 and 0.215, respectively, after only five
879 automatic cycles of iterative model-building, refinement and density modification using
880 bucanneer [52] and refmac5 [53] in the CCP4 suite [74].
881
882 **Crystal solubilization assays.** The solubility of crystals of Cry11Aa WT and of the mutants
883 F17Y, Y272Q, Y349F, Y449F, D507N-D514N and E583Q was measured at different pH values
884 as previously described [12]. Briefly, crystal suspensions were centrifuged at 11,000 g for 2 min
885 and resuspended in 18 different buffers with pH ranging from 8.6 to 14.2. After 1h incubation
886 in each buffer, crystals were centrifuged and the supernatant was collected. The concentration

887 of soluble toxin in the supernatant was quantified using a Nanodrop 2000 (Thermo Fisher
888 Scientist) by measuring the OD at 280 nm and by using the molar extinction coefficient and
889 toxin size (102,930 M^{-1} cm^{-1} and 72.349 kDa, respectively, as calculated with the ProtParam
890 tool of ExPASy (https://www.expasy.org) using the Cry11Aa protein sequence available under
891 accession number "P21256 [https://www.uniprot.org/uniprot/ P21256]"). Solubility was
892 measured in triplicate for each toxin (Cry11Aa WT and mutants) and each pH. Data are
893 normalized and represented as percentage of solubilization by dividing the concentration
894 measured at a given pH by the concentration at the highest pH measured. Solubility of
895 Cry11Aa WT and its different mutants was compared by calculating SP_{50} (pH leading to
896 solubilization of 50% of crystals) as previously described [12], by fitting the data using a logistic
897 regression model for binomial distribution using a script modified from [97]. Differences in SP_{50}
898 between mutants were considered significant when 95% confidence intervals (CI), calculated
899 using a Pearson's chi square goodness-of-fit test, did not overlap [98]. All statistics were
900 conducted using the software R 3.5.2 [99].

901 For the Cry11Ba, the crystal suspensions were centrifuged at 13,300 g for 3 min and ultrapure
902 H2O was removed from crystals. They were then resuspended in one of 18 buffers ranging
903 from pH 7 to 14. These crystals were incubated for 1 hr, afterwards the solution was
904 centrifuged at 13,300 g and the supernatant was separated from the crystal pellet. The
905 concentration of the supernatant was then quantified by a ThermoFisher Nanodrop One
906 (Thermo) by measuring the OD for 280 nm and utilizing the molar extinction coefficient and
907 toxin size (114600 $M^{-1}.cm^{-1}$ and 81344.18 Da respectively) that were calculated with Expasy
908 ProtParam using the Cry11Ba sequence available at Uniprot.org under accession number
909 Q45730. Solubility was measured in triplicate for the toxin at each pH measured. This was
910 then further tested by conducting a turbidity assay by resuspending the crystal pellet in 150 µL
911 ultrapure H2O and placed in a 96-well plate to be read on an NEPHELOstar Plus (BMG
912 Labtech) nephelometer. These counts were normalized by subtracting the background signal
913 and conducted in triplicate.
914

915 **Proteomic characterization**

916 For SDS-PAGE experiments, samples heated to 95 °C were migrated on 12 % SDS-PAGE
917 gels (1 h, 140 V) after addition of Laemmle buffer devoid of DTT. After staining by overnight
918 incubation in Instant*Blue* (Sigma Aldrich, France), gels were washed twice in ultrapure water
919 and migration results were digitalized using a ChemiDoc XRS+ imaging system controlled by
920 Image Lab software version 6.0.0 (BioRad, France).
921

922 **MALDI TOF mass spectrometry**

923 MALDI TOF mass spectra on Cry11Aa were acquired on an Autoflex mass spectrometer
924 (Bruker Daltonics, Bremen, Germany) operated in linear positive ion mode. External mass
925 calibration of the instrument, for the *m/z* range of interest, was carried out using as calibrants
926 the monomeric (66.4 kDa) and dimeric (132.8 kDa) ions of BSA (reference 7030, Sigma
927 Aldrich). Just before analysis, crystals of Cry11Aa were firstly dissolved in acetonitrile/water
928 mixture (70:30, *v/v*). For samples under reducing condition, DTT was added at a final
929 concentration of 10 mM. The obtained solutions were therefore directly mixed in variable ratios
930 (1:5, 1:10, 1:20, *v/v*) with sinapinic acid matrix (20 mg/mL solution in
931 water/acetonitrile/trifluoroacetic acid, 70:30:0.1, *v/v/v*, Sigma Aldrich) to obtain the best signal-
932 to-noise ratio for MALDI mass spectra. 1 to 2 µL of these mixtures were then deposited on the
933 target and allowed to air dry (at room temperature and pressure). Mass spectra were acquired
934 in the 10 to 160 kDa *m/z* range and data processed with Flexanalysis software (v.3.0, Bruker
935 Daltonics).
936 MALDI TOF mass spectra on Cry11Ba were collected at the USC Mass Spectrometry Core
937 Facility, Los Angeles, CA, USA. Purified Cry11Ba protein was dissolved in water (~ 5 mg/mL)
938 and heated at 70 °C for 10 min to facilitate dissolution. One microliter of protein solution was
939 spotted on a 384 Big Anchor MALDI target and let dry at room temperature. Crystallized protein
940 was washed on-target twice with MQ water, on top of which 0.5 µL of 2,6
941 dihydroxyacetophenone (DHAP) solution (30 mg/ml in 50% acetonitrile:0.1% formic acid) was
942 spotted and let dry at room temperature. Crystallized sample was then analyzed using Bruker
943 Rapiflex® MALDI-TOF MS equipped with a Smartbeam 3D, 10 kHz, 355 nm Nd:YAG laser.
944 The laser parameters were optimized as follows: scan range = 26 µm; number of shots per
945 sample = 1000; laser frequency = 5000 Hz. The mass spectrometer was calibrated for high-
946 mass range using Protein A and Trypsinogen standards under Linear Mode. Data were
947 analyzed using FlexAnalysis software and plotted using Graphpad Prism.
948
949 **In-gel digestion and peptide mass fingerprinting of Cry11Aa using MALDI.**
950 Selected bands were in-gel digested with trypsin as previously described[100]. MALDI mass
951 spectra of the tryptic peptides were recorded on an Autoflex mass spectrometer (Bruker
952 Daltonics, Bremen, Germany) in the reflectron positive ion mode. Before analysis samples
953 were desalted and concentrated on RP-C18 tips (Millipore) and eluted directly with 2 µl of α-
954 cyano-4-hydroxy cinnamic acid matrix (10 mg/ml in water/acetonitrile/trifluoroacetic acid:
955 50/50/0.1, *v/v/v*) on the target.
956
957 **In-gel digestion and peptide mass fingerprinting of Cry11Ba using GeLC-MS/MS.**
958 Gel Liquid Chromatography tandem mass spectrometry mass spectra collected on Cry11Ba
959 were acquired on a ThermoFisher Q-Exactive Plus (UCLA Molecular Instrumentation Center,

960 Los Angeles, CA, USA). Before analysis, the Cry11Ba crystals were diluted at a 1:5 dilution
961 with ultrapure H2O and 4x SDS Loading Buffer Dye. These samples were then boiled for 3
962 min at 98°C and were loaded on a 4-12% Bis-Tris SDS-PAGE gel. Protein embedded in gel
963 bands were extracted and digested with 200 ng trypsin at 37°C overnight. The digested
964 products were extracted from the gel bands in 50% acetonitrile/49.9% H2O/ 0.1%
965 trifluoroacetic acid (TFA) and desalted with C18 StageTips prior to analysis by tandem mass
966 spectrometry. Peptides were injected on an EASY-Spray HPLC column (25 cm x 75 µm ID,
967 PepMap RSLC C18, 2 µm, ThermoScientific). Tandem mass spectra were acquired in a data-
968 dependent manner with a quadrupole orbitrap mass spectrometer (Q-Exactive Plus Thermo
969 Fisher Scientific) interfaced to a nanoelectrospray ionization source. The raw MS/MS data
970 were converted into MGF format by Thermo Proteome Discoverer (VER. 1.4, Thermo
971 Scientific). The MGF files were then analyzed by a MASCOT sequence database search.
972
973 **Native mass spectrometry.** Crystals of Cry11Aa were centrifuged for 5 minutes at 5000 g
974 during the buffer wash and washed twice with ammonium acetate buffer (pH adjusted to 6.4
975 with acetic acid). Pelleted crystals were then dissolved in ammonium acetate buffer (pH
976 adjusted to 11.5 using ammonium hydroxide). Gold-coated capillary emitters were prepared
977 as previously described and used to load the protein sample [101]. The sample was analyzed on
978 a Synapt G1 mass spectrometer (Waters Corporation). The instrument was tuned to preserve
979 non-covalent interactions. Briefly, the capillary voltage was set to 1.60 kV, the sampling cone
980 voltage was 20 V, the extraction cone voltage was 5 V, the source temperature was 80 °C, the
981 trap transfer collision energy was 10V, and the trap collision energy (CE) was set at 30 V. For
982 MS/MS characterization, a particular charge state was isolated in the quadrupole and the
983 complex was dissociated by application of 200V of CE. The data collected were deconvoluted
984 and analyzed using UniDec [102].
985
986 **Heat stability and aggregation propensity.** The thermal unfolding of Cry11Aa WT and
987 mutants was measured by following changes as a function of temperature (15 – 95 °C) in
988 tryptophan fluorescence leading to an increase of the F350/F330 ratio. Scattering was also
989 monitored to address aggregation propensity of Cry11Aa WT and of the mutants F17Y,
990 Y272Q, Y349F, Y449F, D507N-D514N and E583Q (Supplementary Fig. 12). All the
991 measurements were performed on a Prometheus NT.48 (Nanotemper) following
992 manufacturer's instructions.
993
994 **Data availability**
995 Structures and structure factor amplitudes have been deposited in the PDB databank under
996 accession codes 7QX4, 7QYD, 7QX5, 7QX6, 7QX7 ("7QX4"; [10.2210/pdb7QX4/pdb], etc).

Raw image files are deposited in cxi.db accession number 190 (https://www.cxidb.org/id-190.html). The source data for Figure 3 and for Supplementary Figs. 5, 6, 7, 8, 10, 11 and 12 are provided as a Source Data file.

1002 **References**

1003 1. Bravo, A., Likitvivatanavong, S., Gill, S. S. & Soberon, M. Bacillus thuringiensis: A
1004 story of a successful bioinsecticide. *Insect Biochem. Mol. Biol.* **41**, 423–431 (2011).
1005 2. Vachon, V., Laprade, R. & Schwartz, J. L. Current models of the mode of action of
1006 Bacillus thuringiensis insecticidal crystal proteins: a critical review. *J. Invertebr. Pathol.* **111**,
1007 1–12 (2012).
1008 3. Delecluse, A., Rosso, M. L. & Ragni, A. Cloning and expression of a novel toxin gene
1009 from Bacillus thuringiensis subsp. jegathesan encoding a highly mosquitocidal protein. *Appl*
1010 *Env. Microbiol* **61**, 4230–5 (1995).
1011 4. Federici, B. A., Park, H.-W. & Sakano, Y. Insecticidal Protein Crystals of Bacillus
1012 thuringiensis. in *Inclusions in Prokaryotes* (ed. Shively, J. M.) 195–236 (Springer, 2006).
1013 doi:10.1007/7171_008/Published.
1014 5. Boonserm, P., Mo, M., Angsuthanasombat, C. & Lescar, J. Structure of the functional
1015 form of the mosquito larvicidal Cry4Aa toxin from Bacillus thuringiensis at a 2.8-Angstrom
1016 resolution. *J. Bacteriol.* **188**, 3391–3401 (2006).
1017 6. Boonserm, P., Davis, P., Ellar, D. J. & Li, J. Crystal structure of the mosquito-
1018 larvicidal toxin Cry4Ba and its biological implications. *J. Mol. Biol.* **348**, 363–382 (2005).
1019 7. Cohen, S. *et al.* Cyt1Aa Toxin: Crystal Structure Reveals Implications for Its
1020 Membrane-Perforating Function. *J. Mol. Biol.* **413**, 804–814 (2011).
1021 8. Gutierrez, P., Alzate, O. & Orduz, S. A theoretical model of the tridimensional
1022 structure of Bacillus thuringiensis subsp. medellin Cry 11Bb toxin deduced by homology
1023 modelling. *Mem Inst Oswaldo Cruz* **96**, 357–64 (2001).
1024 9. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**,
1025 73–77 (2011).
1026 10. Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2**, 246–
1027 255 (2015).
1028 11. Boutet, S., Fromme, P. & Hunter, M. S. *X-ray Free Electron Lasers - A Revolution in*
1029 *Structural Biology*. vol. 1 (Springer International Publishing, 2018).
1030 12. Tetreau, G. *et al.* Serial femtosecond crystallography on in vivo-grown crystals drives
1031 elucidation of mosquitocidal Cyt1Aa bioactivation cascade. *Nat. Commun.* **11**, 1153 (2020).
1032 13. Li, J. D., Carroll, J. & Ellar, D. J. Crystal structure of insecticidal delta-endotoxin from
1033 Bacillus thuringiensis at 2.5 A resolution. *Nature* **353**, 815–821 (1991).
1034 14. Morse, R. J., Yamamoto, T. & Stroud, R. M. Structure of Cry2Aa Suggests an
1035 Unexpected Receptor Binding Epitope. *Structure* **9**, 409–417 (2001).
1036 15. Grochulski, P. *et al.* Bacillus thuringiensis CryIA(a) Insecticidal Toxin: Crystal
1037 Structure and Channel Formation. *J. Mol. Biol.* **254**, 447–464 (1995).
1038 16. Engilberge, S. *et al.* Crystallophore: a versatile lanthanide complex for protein
1039 crystallography combining nucleating effects, phasing properties, and luminescence. *Chem.*
1040 *Sci.* **8**, 5909–5917 (2017).
1041 17. Engilberge, S. *et al.* Protein crystal structure determination with the crystallophore, a
1042 nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722–731 (2019).
1043 18. Xu, C., Wang, B. C., Yu, Z. & Sun, M. Structural insights into Bacillus thuringiensis
1044 Cry, Cyt and parasporin toxins. *Toxins* **6**, 2732–70 (2014).
1045 19. Promponas, V. J. *et al.* CAST: an iterative algorithm for the complexity analysis of
1046 sequence tracts. *Bioinformatics* **16**, 915–922 (2000).
1047 20. Wootton, J. C. & Federhen, S. Statistics of local complexity in amino acid sequences
1048 and sequence databases. *Comput. Chem.* **17**, 149–163 (1993).
1049 21. Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light
1050 Source. *J. Synchrotron Radiat.* **22**, 514–519 (2015).
1051 22. Sierra, R. G. *et al.* Concentric-flow electrokinetic injector enables serial
1052 crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).
1053 23. DePonte, D. P. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet
1054 streams. *J. Phys. -Appl. Phys.* **41**, 7 (2008).
1055 24. Leaver-Fay, A. *et al.* Rosetta3: An Object-Oriented Software Suite for the Simulation

1056 and Design of Macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).
1057 25. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its
1058 supplement TrEMBL in 1999. *Nucleic Acids Res.* **27**, 49–54 (1999).
1059 26. Pardo-Lopez, L., Soberon, M. & Bravo, A. Bacillus thuringiensis insecticidal three-
1060 domain Cry toxins: mode of action, insect resistance and consequences for crop protection.
1061 *FEMS Microbiol. Rev.* **37**, 3–22 (2013).
1062 27. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for
1063 macromolecular crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 143–151 (2018).
1064 28. Colletier, J.-P. *et al.* De novo phasing with X-ray laser reveals mosquito larvicide
1065 BinAB structure. *Nature* **539**, 43–47 (2016).
1066 29. Oestergaard, J., Ehlers, R. U., Martinez-Ramirez, A. C. & Real, M. D. Binding of
1067 Cyt1Aa and Cry11Aa toxins of Bacillus thuringiensis serovar israelensis to brush border
1068 membrane vesicles of Tipula paludosa (Diptera : Nematocera) and subsequent pore
1069 formation. *Appl. Environ. Microbiol.* **73**, 3623–3629 (2007).
1070 30. Yamagiwa, M., Sakagawa, K. & Sakai, H. Functional Analysis of Two Processed
1071 Fragments of Bacillus thuringiensis Cry11A Toxin. *Biosci. Biotechnol. Biochem.* **68**, 523–528
1072 (2014).
1073 31. Munoz-Garay, C. *et al.* Oligomerization of Cry11Aa from Bacillus thuringiensis Has an
1074 Important Role in Toxicity against Aedes aegypti. *Appl. Environ. Microbiol.* **75**, 7548–7550
1075 (2009).
1076 32. Fernandez, L. E. *et al.* Cry11Aa toxin from Bacillus thuringiensis binds its receptor in
1077 Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514
1078 (2005).
1079 33. Carmona, D. *et al.* Dominant Negative Phenotype of Bacillus thuringiensis Cry1Ab,
1080 Cry11Aa and Cry4Ba Mutants Suggest Hetero-Oligomer Formation among Different Cry
1081 Toxins. *PLoS ONE* **6**, e19952 (2011).
1082 34. Du, J. P., Knowles, B. H., Li, J. & Ellar, D. J. Biochemical characterization of Bacillus
1083 thuringiensis cytolytic toxins in association with a phospholipid bilayer. *Biochem. J.* **338**, 185–
1084 193 (1999).
1085 35. Ntountoumi, C. *et al.* Low complexity regions in the proteins of prokaryotes perform
1086 important functional roles and are highly conserved. *Nucleic Acids Res.* **47**, 9998–10009
1087 (2019).
1088 36. Henrich, B. *et al.* The adaptive gain integrating pixel detector AGIPD a detector for
1089 the European XFEL. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect.*
1090 *Assoc. Equip.* **633**, S11–S14 (2011).
1091 37. Likitvivatanavong, S. *et al.* Multiple Receptors as Targets of Cry Toxins in
1092 Mosquitoes. *J. Agric. Food Chem.* **59**, 2829–2838 (2011).
1093 38. Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A. & Soberon, M. A GPI-
1094 anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in Aedes
1095 aegypti larvae. *Biochem. J.* **394**, 77–84 (2006).
1096 39. Bourgouin, C., Delecluse, A. & Rapoport, G. Specificity and Synergism of Bacillus
1097 thuringiensis israelensis Toxins on Mosquito Larvae. in *Bacterial Protein Toxins* (ed.
1098 Fehrenbach) 225–226 (Gustav Fisher, 1988).
1099 40. Lopez-Diaz, J. A., Emiliano Canton, P., Gill, S. S., Soberon, M. & Bravo, A.
1100 Oligomerization is a key step in Cyt1Aa membrane insertion and toxicity but not necessary to
1101 synergize Cry11Aa toxicity in Aedes aegypti larvae. *Environ. Microbiol.* **15**, 3030–3039
1102 (2013).
1103 41. Pérez, C. *et al.* Bacillus thuringiensis subsp. israelensis Cyt1Aa synergizes Cry11Aa
1104 toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. U. S. A.* **102**,
1105 18303–18308 (2005).
1106 42. Sawaya, M. R. *et al.* Protein crystal structure obtained at 2.9 Å resolution from
1107 injecting bacterial cells into an X-ray free-electron laser beam. *Proc. Natl. Acad. Sci.* **111**,
1108 12769–12774 (2014).
1109 43. Galitsky, N. *et al.* Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of
1110 Bacillus thuringiensis. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1101–1109 (2001).

1111 44. Adalat, R., Saleem, F., Crickmore, N., Naz, S. & Shakoori, A. R. In Vivo
1112 Crystallization of Three-Domain Cry Toxins. *Toxins Basel* **9**, (2017).
1113 45. Hughes, M. P. *et al.* Atomic structures of low-complexity protein segments reveal
1114 kinked β sheets that assemble networks. *Science* **359**, 698–701 (2018).
1115 46. Salinas, N., Colletier, J. P., Moshe, A. & Landau, M. Extreme amyloid polymorphism
1116 in Staphylococcus aureus virulent PSMα peptides. *Nat Commun* **9**, 3512 (2018).
1117 47. Sun, Y. *et al.* Construction and characterization of the interdomain chimeras using
1118 Cry11Aa and Cry11Ba from Bacillus thuringiensis and identification of a possible novel toxic
1119 chimera. *Biotechnol. Lett.* **36**, 105–111 (2013).
1120 48. Fernandez, L. E. *et al.* Cry11Aa toxin from Bacillus thuringiensis binds its receptor in
1121 Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514
1122 (2005).
1123 49. Schonherr, R., Rudolph, J. M. & Redecke, L. Protein crystallization in living cells. *Biol*
1124 *Chem* **399**, 751–772 (2018).
1125 50. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a
1126 three-track neural network. *Science* **373**, 871–876 (2021).
1127 51. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*
1128 **596**, 583–589 (2021).
1129 52. Cowtan, K. Completion of autobuilt protein models using a database of protein
1130 fragments. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 328–335 (2012).
1131 53. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal
1132 structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
1133 54. Banneville, A.-S. *et al.* Structural and functional characterization of DdrC, a novel
1134 DNA damage-induced nucleoid associated protein involved in DNA compaction | bioRxiv.
1135 https://www.biorxiv.org/content/10.1101/2021.10.27.466113v1?rss=1.
1136 55. Park, H. W., Delecluse, A. & Federici, B. A. Construction and characterization of a
1137 recombinant Bacillus thuringiensis subsp israelensis strain that produces Cry11B. *J.*
1138 *Invertebr. Pathol.* **78**, 37–44 (2001).
1139 56. Wu, D. & Federici, B. A. Improved production of the insecticidal CryIVD protein in
1140 Bacillus-thuringiensis using CryLA(c) promoters to express the gene for an associated 20-
1141 KDa protein. *Appl. Microbiol. Biotechnol.* **42**, 697–702 (1995).
1142 57. Kirmitzoglou, I. & Promponas, V. J. LCR-eXXXplorer: a web platform to search,
1143 visualize and share data for low complexity regions in protein sequences. *Bioinformatics* **31**,
1144 2208–2210 (2015).
1145 58. Girish, V. & Vijayalakshmi, A. Affordable image analysis using NIH Image/ImageJ.
1146 *Indian J Cancer* **41**, 47 (2004).
1147 59. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis.
1148 *Cent. Eur. J. Phys.* **10**, 181–188 (2012).
1149 60. Chen, S. W. & Pellequer, J.-L. DeStripe: frequency-based algorithm for removing
1150 stripe noises from AFM images. *BMC Struct. Biol.* **11**, 7 (2011).
1151 61. Sierra, R. G. *et al.* Concentric-flow electrokinetic injector enables serial
1152 crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).
1153 62. Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light
1154 Source. *J. Synchrotron Radiat.* **22**, 514–519 (2015).
1155 63. Waterman, D. G. *et al.* Diffraction-geometry refinement in the DIALS framework. *Acta*
1156 *Crystallogr. Sect. Struct. Biol.* **72**, 558–575 (2016).
1157 64. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based
1158 methods for data processing. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1274–1282 (2013).
1159 65. Hattne, J. *et al.* Accurate macromolecular structures using minimal measurements
1160 from X-ray free-electron lasers. *Nat. Methods* **11**, 545–548 (2014).
1161 66. Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package.
1162 *Acta Crystallogr. Sect. Struct. Biol.* **74**, 85–97 (2018).
1163 67. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular
1164 Replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
1165 68. Leaver-Fay, A. *et al.* ROSETTA3: an object-oriented software suite for the simulation

1166 and design of macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).
1167 69. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
1168 (2007).
1169 70. Coquelle, N. *et al.* Raster-scanning serial protein crystallography using micro- and
1170 nano-focused synchrotron beams. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1184–1196
1171 (2015).
1172 71. Foucar, L. *et al.* CASS—CFEL-ASG software suite. *Comput. Phys. Commun.* **183**,
1173 2207–2213 (2012).
1174 72. White, T. A. *et al.* Crystallographic data processing for free-electron laser sources.
1175 *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1231–1240 (2013).
1176 73. Skubák, P. & Pannu, N. S. Automatic protein structure solution from weak X-ray data.
1177 *Nat. Commun.* **4**, 1–6 (2013).
1178 74. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta*
1179 *Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
1180 75. Gevorkov, Y. *et al.* XGANDALF - extended gradient descent algorithm for lattice
1181 finding. *Acta Crystallogr. Sect. Found. Adv.* **75**, 694–704 (2019).
1182 76. Duisenberg, A. J. M. Indexing in single-crystal diffractometry with an obstinate list of
1183 reflections. *J. Appl. Crystallogr.* **25**, 92–96 (1992).
1184 77. Ginn, H. M. *et al.* ıt TakeTwo: an indexing algorithm suited to still images with known
1185 crystal parameters. *Acta Crystallogr. Sect. D* **72**, 956–965 (2016).
1186 78. Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr.*
1187 *D Biol. Crystallogr.* **62**, 48–57 (2006).
1188 79. White, T. A. *et al.* ıt CrystFEL: a software suite for snapshot serial crystallography. *J.*
1189 *Appl. Crystallogr.* **45**, 335–341 (2012).
1190 80. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for
1191 macromolecular crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 143–151 (2018).
1192 81. Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density
1193 modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 61–
1194 69 (2008).
1195 82. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*
1196 *Crystallogr Biol Crystallogr* **60**, 2126–32 (2004).
1197 83. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with
1198 phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
1199 84. Uervirojnangkoorn, M. *et al.* Enabling X-ray free electron laser crystallography for
1200 challenging biological systems from a limited number of crystals. *eLife* **4**, (2015).
1201 85. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood
1202 in BUSTER-TNT. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2210–2221 (2004).
1203 86. *The PyMOL Molecular Graphics System*. (Schrödinger, LLC.).
1204 87. Bond, C. S. & Schüttelkopf, A. W. ALINE: a WYSIWYG protein-sequence alignment
1205 editor for publication-quality alignments. *Acta Crystallogr Biol Crystallogr* **65**, 510–2 (2009).
1206 88. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline
1207 state. *J. Mol. Biol.* **372**, 774–797 (2007).
1208 89. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019.
1209 *Nucleic Acids Res.* **47**, W636–W641 (2019).
1210 90. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast
1211 protein structure alignment in three dimensions. *Acta Crystallogr Biol Crystallogr* **60**, 2256–
1212 68 (2004).
1213 91. Gille, C. & Frömmel, C. STRAP: editor for STRuctural Alignments of Proteins.
1214 *Bioinformatics* **17**, 377–8 (2001).
1215 92. Gille, C., Fähling, M., Weyand, B., Wieland, T. & Gille, A. Alignment-Annotator web
1216 server: rendering and annotating sequence alignments. *Nucleic Acids Res.* **42**, W3–W6
1217 (2014).
1218 93. Mirdita, M., Steinegger, M., Breitwieser, F., Söding, J. & Levy Karin, E. Fast and
1219 sensitive taxonomic assignment to metagenomic contigs. *Bioinformatics* **37**, 3029–3031
1220 (2021).

1221 94. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching
1222 for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
1223 95. Prakash, A., Jeffryes, M., Bateman, A. & Finn, R. D. The HMMER Web Server for
1224 Protein Sequence Similarity Search. *Curr. Protoc. Bioinforma.* **60**, 3.15.1-3.15.23 (2017).
1225 96. Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–
1226 W204 (2018).
1227 97. Savi, M. K., Mangamana, E. T., Deguenon, J. M., Hounmenou, C. G. & Kakaï, R. G.
1228 Determination of Lethal Concentrations Using an R Software Function Integrating the Abbott
1229 Correction. *J. Agric. Sci. Technol. A* **7**, (2017).
1230 98. Wheeler, M. W., Park, R. M. & Bailer, A. J. Comparing median lethal concentration
1231 values using confidence interval overlap or ratio tests. *Env. Toxicol Chem* **25**, 1441–4 (2006).
1232 99. R. *R: a language and environment for statistical computing.* (R Foundation for
1233 Statistical Computing, 2011).
1234 100. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. In-gel digestion of
1235 proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis.
1236 *Anal. Biochem.* **203**, 173–179 (1992).
1237 101. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry
1238 of intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).
1239 102. Marty, M. T. *et al.* Bayesian Deconvolution of Mass and Ion Mobility Spectra: From
1240 Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
1241

Acknowledgements. We thank Aline LeRoy for help during experiments with the Nanotemper apparatus and Dr. Guy Schoehn for his support. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). We thank the LCLS for beamtime allocation under proposals cxi0416, L091, P127, P125 and P141, and the EuXFEL for beamtime allocation under proposals P2156 and P2545. This work was supported by the Agence Nationale de la Recherche (grants ANR-17-CE11-0018-01 and ANR-2018-CE11-0005-02 to J.-P.C.), the CNRS (PEPS SASLELX grant to MW) and used the atomic force microscopy (AFM) platform at the IBS and the electron microscopy (EM) platform of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). The EM facility is supported by the Auvergne-Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER and the GIS-Infrastructures en Biologie Sante et Agronomie (IBiSA). Use of the LCLS at the Stanford Linear Accelerator Center (SLAC) National Laboratory, is supported by the US Department of Energy, Office of Science, and Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. The CXI instrument was funded by the Linac Coherent Light Source Ultrafast Science Instruments project, itself funded by the DOE Office of Basic Energy Sciences. Parts of the sample injector used at LCLS for this research were funded by the National Institute of Health, P41GM103393, formerly P41RR001209. This research was also supported by NIH grant GM117126. Data processing was supported by National Institutes of Health grant GM117126 to N.K.S. N.A.S.'s contributions reported in this publication were supported by the Ruth L. Kirschstein National Research Service Award of the National Institutes of Health under award number GM007185.

Author contributions

G.T., A.-S.B, D.B., E.L., L.D., H.-W.P., B.F. designed and constructed and transformed WT and mutants plasmids; G.T., E.A.A., A.-S.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.A.B. 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.-S.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. and J.-L.P. performed the statistical analysis of solubilization data and G.T. and J.-L.P. performed the statistical analyses of AFM data; A.-S.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.-S.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 TbX-o4; G.T. derivatized Cry11Aa crystals for injection at LCLS; G.T., M.R.S., E.A.A., A.-S.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.-S.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.-S.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.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.-S.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.

Competing interests

The authors declare no competing interests.

Tables

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

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

1307 **Figures captions :**

1308

1309 **Fig. 1. Crystals and overall fold of Cry11 toxins. a-b,** scanning (left; SEM) and transmission
1310 (middle, right; TEM) electron micrographs of gold plated and negatively-stained Cry11Aa (**a**)
1311 and Cry11Ba (**b**) crystals, respectively. The right panels show a close-up view of the crystal
1312 surface. **c,** Cry11Aa crystal structure, depicted as cartoon. Domain I is shown in blue; domain
1313 II is shown in orange except for the $\alpha_h\beta_h$-handle and βpin which are shown in purple and red,
1314 respectively; domain III is shown in pink. **d**, Topology diagram of a Cry11Aa dimer. Domain I
1315 is shown in green, except for central helix α5, which is shown in blue: domain II is shown in
1316 magenta, except for the $\alpha_h\beta_h$-handle, which is shown in purple; and domain III is shown in
1317 cyan, respectively. The two monomers in a dimer assemble via the β_{pin}, resulting in the
1318 formation of a large β-sheet.

1319

1320 **Fig. 2. Monomer interactions in Cry11Aa and Cry11Ba. a,** Cry11Aa crystal packing,
1321 coloured according to sequence (from blue to red) indicating the domain-based assembly; and
1322 coloured according to tetramer assembly (see panel (**b**)). The highlighted areas indicate the
1323 regions shown in (**b**) (full line) and (**c**) (dashed line). **b,** Cry11Aa tetramer with zoom on each
1324 of the three interfaces identified by PISA (interface #1, #3 and #6), with the involved residues
1325 depicted as spheres. For interfaces with hydrogen and/or salt bridges (see g), an additional
1326 (right) image shows only those residues that make up these interactions. **c,** Cry11Aa crystal
1327 assembly by interactions between neighbouring tetramers, formed by interface #2, #4 and #5,
1328 visualized as in b. **d,** Cry11Ba crystal packing, coloured as in (**a**). **e,** Cry11Ba tetramer with
1329 zoom on the interfaces as in (**b**). **f,** Cry11Ba crystal assembly, visualized as in (**c**). As
1330 compared to Cry11Aa, Cry11Ba crystals contain an additional interface #7 between an A-B
1331 pair from two neighbouring tetramers. **g,** interface statistics as identified by PISA for Cry11Aa
1332 (blue) and Cry11Ba (red).

1333

1334 **Fig. 3. Point-mutations of Cry11Aa affect the shape, size and pH-sensitivity of *in vivo*-**
1335 **grown nanocrystals. a, C**rystals from mutants exhibit similar sigmoidal patterns of crystal
1336 solubilization as a function of pH, except F17Y and E583Q that are more and less sensitive to
1337 pH, respectively. Error bars indicate the standard error of the measurements. **b,** Cry11Aa WT
1338 and mutants exhibit similar heat stability. As expected, toxins are more stable (+ 17.5 ± 0.3°C)
1339 in their crystalline than soluble form, irrespective of the mutation. **c,** Visualization of a
1340 representative crystal for Cry11Aa WT (black) and mutants F17Y (red), Y272Q (brown), Y349F
1341 (purple), Y449F (blue), D507N-D514N (orange) and E583Q (green) by SEM (scale bar = 500
1342 nm). **d,** Crystals of Y449F, F17Y and E583Q imaged by AFM were all smaller in length (L),
1343 width (W), thickness (T) and volume than WT 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.

1347 **Figures**

1348

1349

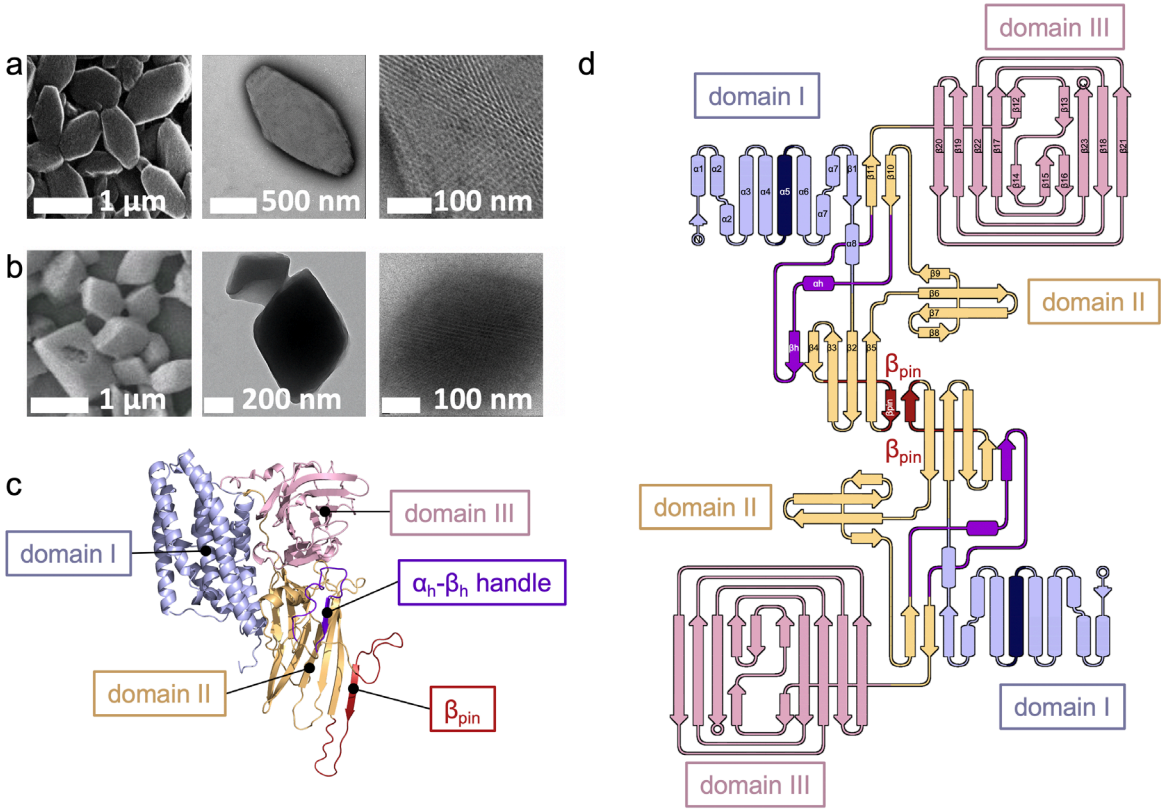

Fig. 1. Crystals and overall fold of Cry11 toxins. a-b, scanning (left; SEM) and transmission (middle, right; TEM) electron micrographs of gold plated and negatively-stained Cry11Aa (**a**) and Cry11Ba (**b**) crystals, respectively. The right panels shows a close-up view of the crystal surface. **c,** Cry11Aa crystal structure, depicted as cartoon. Domain I is shown in 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. Domain I is shown in blue, except for central helix α5, which is shown in dark-blue: domain II is shown in orange, except for the β_{pin} and $\alpha_h\beta_h$-handle, which are shown in red and purple, respectively; and domain III is shown in pink. The two monomers in a dimer assemble via the β_{pin}, resulting in the formation of a large β-sheet.

1350

1351

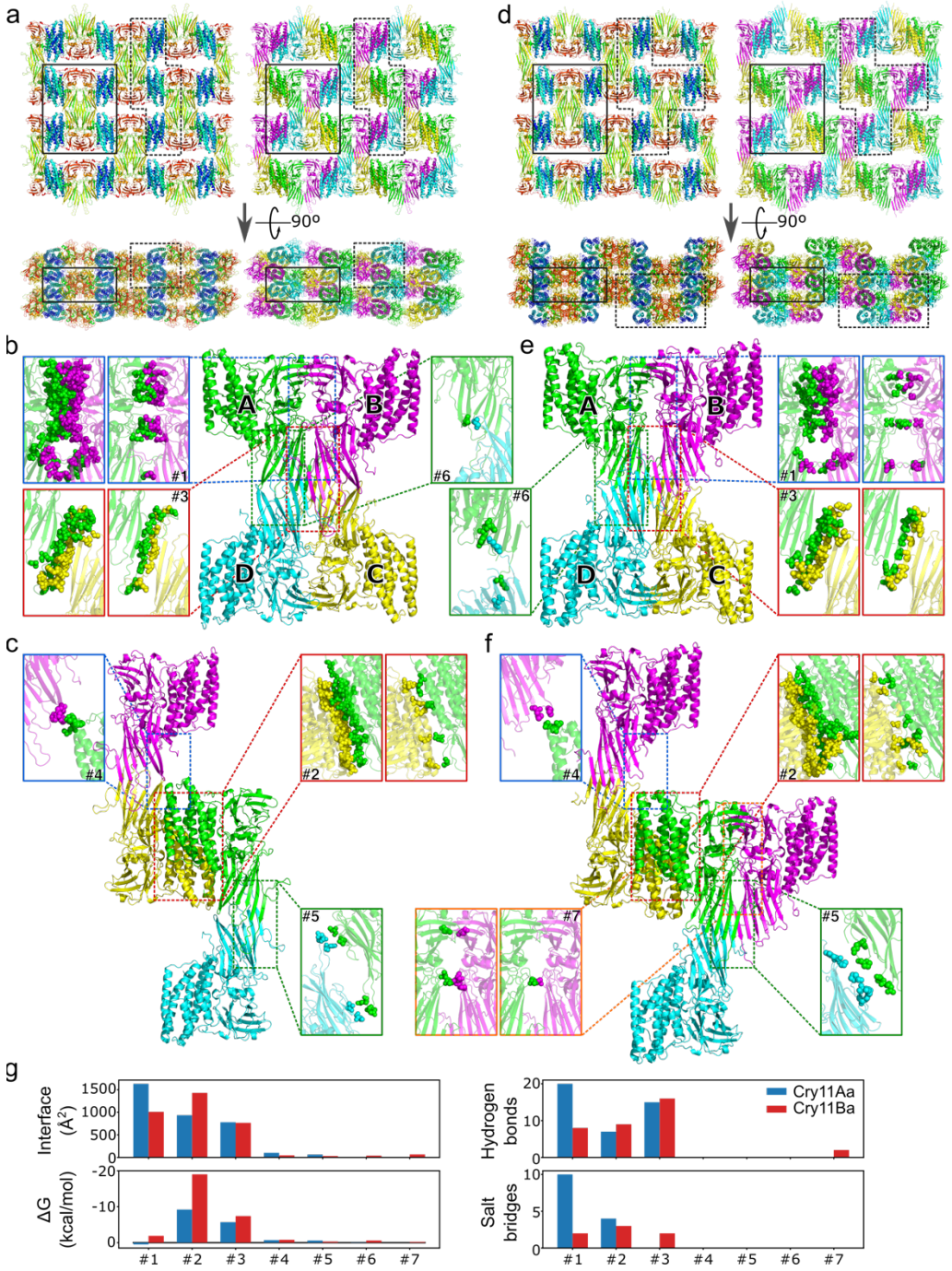

Fig. 2. Monomer interactions in Cry11Aa and Cry11Ba. a, Cry11Aa crystal packing, coloured according to sequence (from blue to red) indicating the domain-based assembly; and coloured according to tetramer assembly (see panel (**b**)). The highlighted areas indicate the regions shown in (**b**) (full line) 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. For interfaces with hydrogen and/or salt bridges (see g), an additional (right) image shows only those residues that make up these interactions. **c,** Cry11Aa crystal assembly by interactions between neighbouring tetramers, formed by interface #2, #4 and #5, visualized as in b. **d,** Cry11Ba crystal packing, coloured 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 neighbouring tetramers. **g,** interface statistics as identified by PISA for Cry11Aa (blue) and Cry11Ba (red).

1352

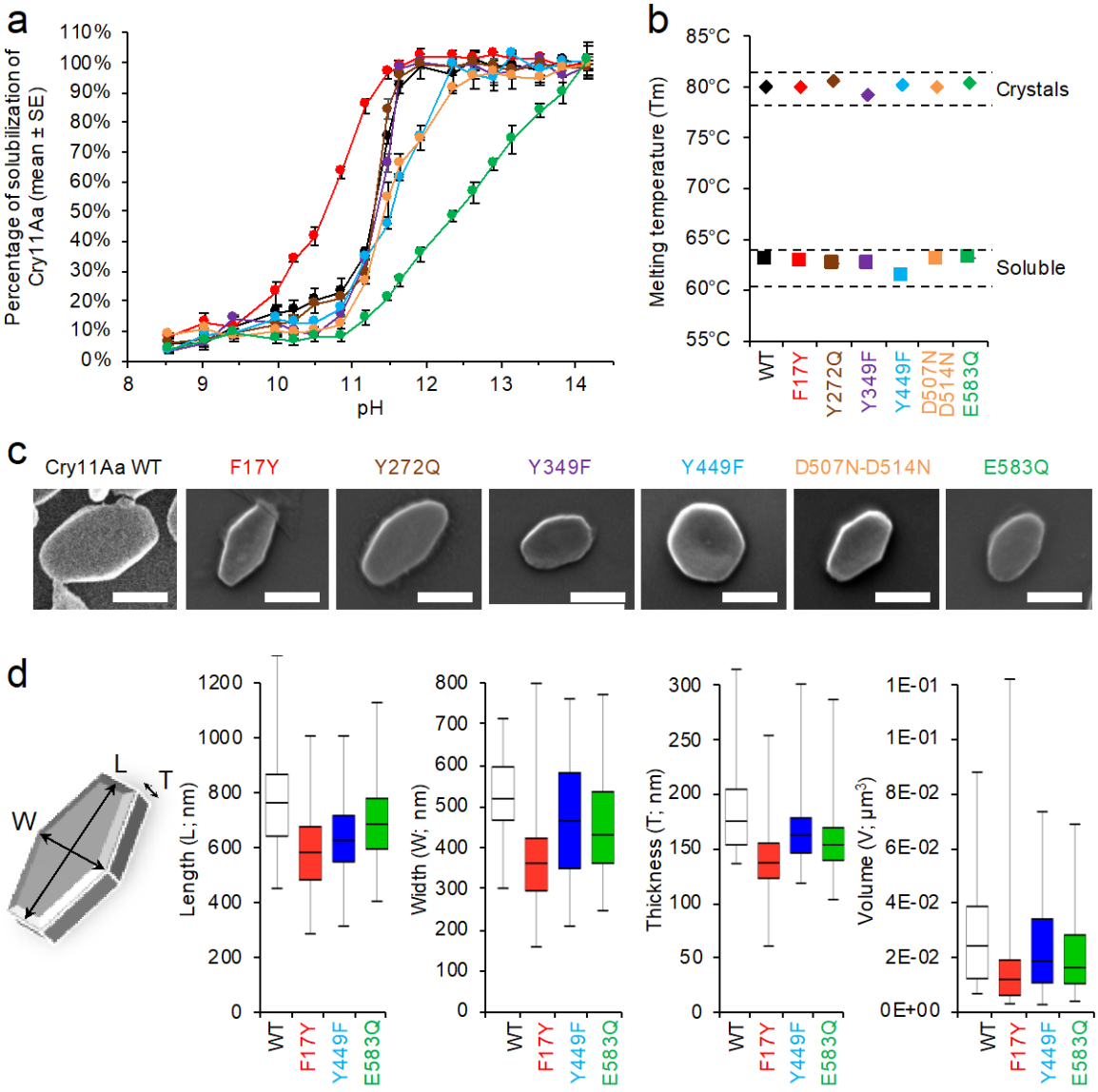

Fig. 3. Point-mutations of Cry11Aa affect the shape, size and pH-sensitivity of *in vivo*-grown nanocrystals. a, Crystals from mutants exhibit similar sigmoidal patterns of crystal solubilization as a function of pH, except F17Y and E583Q that are more and less sensitive to pH, respectively. Error bars indicate the standard error of the measurements. **b,** Cry11Aa WT and mutants exhibit similar heat stability. As expected, toxins are more stable (+ 17.5 ± 0.3°C) in their crystalline than soluble form, irrespective of the mutation. **c,** Visualization of a representative crystal for Cry11Aa WT (black) and mutants F17Y (red), Y272Q (brown), Y349F (purple), Y449F (blue), D507N-D514N (orange) and E583Q (green) by SEM (scale bar = 500 nm). **d,** Crystals of Y449F, F17Y and E583Q imaged by AFM were all smaller in length (L), width (W), thickness (T) and volume than WT 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.

1353
1354
1355

1356 **Tables**

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

	Cry11Aa pH 7	Cry11Aa-TBXO4 pH7	Cry11Ba pH 6.5	Cry11Ba pH 10.4
PDB ID	**7QX4**		**7QYD**	
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 ± 0.19 155.69 ± 0.80 171.14 ± 0.54	57.64 ± 0.15 156.29 ± 0.73 170.75 ± 0.40	168.18 ± 0.19 158.45 ± 0.26 57.51 ± 0.08	167.50 ± 0.29 157.99 ±0.47 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 (2.66 – 2.60)	33.51 – 2.55 (2.61 – 2.55)	44.73 – 2.20 (2.24 – 2.20)	35.72 – 2.55 (2.59 – 2.55)
No. observations	8253629 (365007)	14069217 (640046)	3996555 (31710)	3747530 (52399)
No. unique reflections	24198 (1583)	48634 (3297)	78897 (3815)	50646 (2488)
$\langle I/\sigma(I)\rangle$	9.50 (1.16)	11.23 (1.62)	3.87 (0.50)	3.65 (0.84)
R_{split} (%)	10.73 (95.40)	7.97 (70.58)	18.0 (111.8)	24.50 (97.10)
$CC_{1/2}$	1.00 (0.38)	1.00 (0.68)	0. 99 (0. 12)	0.984 (0.082)
Completeness (%)	99.9 (100.0)	100.0 (100.0)	97.9 (80.9)	99.4 (100.0)
Multiplicity	341.09 (230.58)	289.29 (194.13)	50.66 (8.3)	71.34 (21.01)
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		44.77 – 2.20	35.72 – 2.55
No. reflections	24196		70858	50657
R_{work}/R_{free}*	17.2 / 24.1		18.9 / 23.8	23.8 / 19.2
No. atoms				
Protein	5080		10083	9961
Water	261		623	98
B-factors (Å^2)				
Main chain	50.47		41.82 / 41.06§	41.02 / 42.58
Side chain	51.44		46.11 / 45.55	43.26 / 44.59
Water	46.17		41.36	40.04
R.m.s.d.				
Bonds lengths (Å)	0.004		0.008	0.009
Bonds angles (°)	0.633		1.324	1.590

* R_{free} is calculated using 5 and 10% % of random reflections excluded from refinement.

§ Average B-factor for chain A / chain B

1357

1358

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 ± 0.35 155.39 ± 1.49 171.66 ± 0.64	57.73 ± 0.24 155.55 ± 1.21 171.52 ± 0.57	57.76 ± 0.24 155.51 ± 0.98 171.51 ± 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 (3.40 – 3.48)	23.78 – 3.10 (3.10 – 3.17)	23.50 – 3.30 (3.30 – 3.38)
No. observations	2908715 (141787)	20279640 (1092683)	3210163 (154933)
No. unique reflections	10990 (707)	14447 (950)	12014 (787)
<I/σ (I)>	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)
$CC_{1/2}$	0.96 (0.21)	1.00 (0.60)	0.99 (0.31)
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}*	21.2 / 25.1	22.4 / 25.2	21.5 / 25.4
No. atoms			
Protein	4970	4965	4970
Water	5	13	6
B-factors (Å^2)			
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 (º)	0.448	0.441	0.489

1359 * R_{free} is calculated using 5% of random reflections excluded from refinement.