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], Natalie Schibrowsky[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 seven to thirty-seven times more
78 toxic than Cry11Aa against major mosquito vector species belonging to the genera *Aedes*,
79 *Anopheles*, and *Culex* [3], and in some bacterial hosts appears to form slightly larger crystals.
80 Cry11Ba's structure is also unknown, although it has been used in recombinant strains of *Bti*
81 to improve mosquitocidal 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 could possibly yield structural insights
84 for increasing the efficacy of these proteins for mosquito control. Structure determination of
85 Cry11Aa and Cry11Ba protoxins from natural nanocrystals requires cutting-edge technology.
86 Conventional crystallography is limited to projects in which crystals are sufficiently large to
87 mount and oscillate individually in a synchrotron X-ray beam. In the past, crystals of activated
88 Cry4Aa [5], Cry4Ba [6] and Cyt1Aa [7] attained sufficient size by growing these *in vitro* from toxins
89 dissolved from natural nanocrystals and activating the toxins enzymatically. However,
90 Cry11Aa and Cry11Ba do not recrystallize *in vitro* from dissolved nanocrystals [8]. Moreover,
91 enzymatic activation is unwanted since our goal is to understand the pH-controlled mechanism
92 of natural crystal dissolution. To observe the protoxin state in natural nanocrystals produced
93 in bacterial cells, we applied serial femtosecond crystallography (SFX) at X-ray free electron
94 lasers (XFEL) [9–11]. In the SFX experiment, high brilliance XFEL beam pulses, each lasting only
95 ~10-50 fs, intercept a series of nanocrystals, one pulse-per-crystal, eliciting the strongest
96 possible diffraction signal from each tiny crystal before it vaporizes, and producing a series of
97 diffraction snapshots, later assembled into a full data set. Feasibility of this strategy had been
98 demonstrated by the recent elucidation of the full bioactivation cascade of Cyt1Aa [12].

99 Our success in determining the structures of Cry11Aa and Cry11Ba protoxins highlights
100 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.

127 **Results (4407 words)**

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

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

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

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

164 gadolinium, gold, platinum, and mercury salts, but they failed to produce interpretable
165 isomorphous or anomalous difference Patterson peaks. Finally, a recently introduced caged-
166 terbium compound [16,17], Tb-Xo4, produced a successful derivative of Cry11Aa (after a 30h
167 soak at 10 mM concentration), and phases were determined by the single wavelength
168 anomalous dispersion (SAD) method at 2.6 Å resolution (using anomalous signal up to 3.5 Å).
169 Two Tb-Xo4 molecules were identified bound to the single Cry11Aa molecule in the
170 asymmetric unit (isomorphous peaks at 23 and 9 σ, and anomalous peaks at 33 and 8.1 σ,
171 respectively; Supplementary Fig. 2a). The success of Tb-Xo4 can be partly ascribed to the
172 dramatically high anomalous dispersion signal (*i.e.* f' and f'') of terbium, but likely also stems
173 from stronger binding of TbX04 to the protein owing to presence of an organic cage; indeed, f'
174 and f'' of Gd and Tb are similar at the X-ray energy used for data collection (9 keV). Regardless,
175 phases were of sufficient quality to reveal all Cry11Aa residues from N13 to the C-terminal
176 K643.
177 The Cry11Ba structure was thereafter phased successfully by MR using the Cry11Aa
178 structure as a search model. *A posteriori*, we discovered that two of the heavy atom
179 compounds that we used for soaking actually did bind Cry11Ba (Supplementary Fig. 2b-c).
180 Difference Fourier maps revealed 7-8 σ peaks indicating Pt bound near Met 19 and 200, and
181 Gd bound near Asp83 and Asp427 (Supplementary Fig. 2b). Surprisingly, however, there were
182 no peaks in the anomalous difference Fourier maps. We speculate that if we had achieved
183 higher heavy-atom occupancy and/or higher multiplicity in our measurements, the anomalous
184 signal would have been strong enough to detect and perhaps used for phasing. Our MR-phase
185 2.3 Å resolution map reveals two Cry11Ba molecules in the asymmetric unit. All residues are
186 visible except for the N-terminus (residues M1-N16), two loops (residues G330-E340, and
187 D352-I358) and the C-terminal extension (residues T654-K724). The lack of order in this
188 extension is not surprising given the low complexity of its sequence.
189
190 *Cry11 domain organization is similar to* δ-endotoxins, but exhibits some *non-canonical features*
191 Cry11Aa and Cry11Ba structures maintain the three-domain organization characteristic
192 of Cry δ-endotoxins [13,26] (Fig. 1c, Supplementary Fig. 3). Domain I is implicated to form a pore
193 in the target membrane. Like other Cry δ-endotoxins it forms a seven-α-helix bundle; at the
194 center of the bundle is α5 (residues 146-170), surrounded by the remaining six helices. Domain
195 II is implicated to recognize mosquito-specific receptors. It forms a β-prism composed of three-
196 β-sheets, wherein the first two β-sheets (β4-β3-β2-β5 and β8-β7-β6-β9) each adopts a Greek-
197 key topology while the third β-sheet is three-stranded (β1-β10-β11). Domain III is implicated to
198 oligomerize. It forms a β-sandwich of two antiparallel five-stranded β-sheets (viz. $\beta20-\beta19-$
199 $\beta22-\beta17-\beta4-{}^{\beta12}/_{\beta14}$ and $\beta15-{}^{\beta13}/_{\beta16}-\beta23-\beta18-\beta21$) forming a jelly-roll

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

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

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

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

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

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

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

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

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

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

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

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

461 **Discussion (2219 words)**

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

478

479 In vivo *crystallization pathway of Cry11 toxins*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

900 For the Cry11Ba, the crystal suspensions were centrifuged at 13,300 g for 3 min and ultrapure
901 H2O was removed from crystals. They were then resuspended in one of 18 buffers ranging
902 from pH 7 to 14. These crystals were incubated for 1 hr, afterwards the solution was
903 centrifuged at 13,300 g and the supernatant was separated from the crystal pellet. The
904 concentration of the supernatant was then quantified by a ThermoFisher Nanodrop One
905 (Thermo) by measuring the OD for 280 nm and utilizing the molar extinction coefficient and
906 toxin size (114600 $M^{-1}.cm^{-1}$ and 81344.18 Da respectively) that were calculated with Expasy
907 ProtParam using the Cry11Ba sequence available at Uniprot.org under accession number
908 Q45730. Solubility was measured in triplicate for the toxin at each pH measured. This was
909 then further tested by conducting a turbidity assay by resuspending the crystal pellet in 150 µL
910 ultrapure H2O and placed in a 96-well plate to be read on an NEPHELOstar Plus (BMG
911 Labtech) nephelometer. These counts were normalized by subtracting the background signal
912 and conducted in triplicate.
913
914 **Proteomic characterization**
915 For SDS-PAGE experiments, samples heated to 95 °C were migrated on 12 % SDS-PAGE
916 gels (1 h, 140 V) after addition of Laemmle buffer devoid of DTT. After staining by overnight
917 incubation in Instant*Blue* (Sigma Aldrich, France), gels were washed twice in ultrapure water
918 and migration results were digitalized using a ChemiDoc XRS+ imaging system controlled by
919 Image Lab software version 6.0.0 (BioRad, France).
920
921 **MALDI TOF mass spectrometry**

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

959 Los Angeles, CA, USA). Before analysis, the Cry11Ba crystals were diluted at a 1:5 dilution
960 with ultrapure H2O and 4x SDS Loading Buffer Dye. These samples were then boiled for 3
961 min at 98°C and were loaded on a 4-12% Bis-Tris SDS-PAGE gel. Protein embedded in gel
962 bands were extracted and digested with 200 ng trypsin at 37°C overnight. The digested
963 products were extracted from the gel bands in 50% acetonitrile/49.9% H2O/ 0.1%
964 trifluoroacetic acid (TFA) and desalted with C18 StageTips prior to analysis by tandem mass
965 spectrometry. Peptides were injected on an EASY-Spray HPLC column (25 cm x 75 µm ID,
966 PepMap RSLC C18, 2 µm, ThermoScientific). Tandem mass spectra were acquired in a data-
967 dependent manner with a quadrupole orbitrap mass spectrometer (Q-Exactive Plus Thermo
968 Fisher Scientific) interfaced to a nanoelectrospray ionization source. The raw MS/MS data
969 were converted into MGF format by Thermo Proteome Discoverer (VER. 1.4, Thermo
970 Scientific). The MGF files were then analyzed by a MASCOT sequence database search.
971
972 **Native mass spectrometry.** Crystals of Cry11Aa were centrifuged for 5 minutes at 5000 g
973 during the buffer wash and washed twice with ammonium acetate buffer (pH adjusted to 6.4
974 with acetic acid). Pelleted crystals were then dissolved in ammonium acetate buffer (pH
975 adjusted to 11.5 using ammonium hydroxide). Gold-coated capillary emitters were prepared
976 as previously described and used to load the protein sample [101]. The sample was analyzed on
977 a Synapt G1 mass spectrometer (Waters Corporation). The instrument was tuned to preserve
978 non-covalent interactions. Briefly, the capillary voltage was set to 1.60 kV, the sampling cone
979 voltage was 20 V, the extraction cone voltage was 5 V, the source temperature was 80 °C, the
980 trap transfer collision energy was 10V, and the trap collision energy (CE) was set at 30 V. For
981 MS/MS characterization, a particular charge state was isolated in the quadrupole and the
982 complex was dissociated by application of 200V of CE. The data collected were deconvoluted
983 and analyzed using UniDec [102].
984
985 **Heat stability and aggregation propensity.** The thermal unfolding of Cry11Aa WT and
986 mutants was measured by following changes as a function of temperature (15 – 95 °C) in
987 tryptophan fluorescence leading to an increase of the F350/F330 ratio. Scattering was also
988 monitored to address aggregation propensity of Cry11Aa WT and of the mutants F17Y,
989 Y272Q, Y349F, Y449F, D507N-D514N and E583Q (Supplementary Fig. 12). All the
990 measurements were performed on a Prometheus NT.48 (Nanotemper) following
991 manufacturer's instructions.
992
993 **Data availability**
994 Structures and structure factor amplitudes have been deposited in the PDB databank under
995 accession codes 7QX4 (Cry11Aa WT, pH 7.0; 10.2210/pdb7QX4/pdb), 7QX5 (Cry11Aa

Y449F, pH 7.0; 10.2210/pdb7QX5/pdb), 7QX6 (Cry11Aa E583Q, pH 7.0; 10.2210/pdb7QX6/pdb), 7QX7 (Cry11Aa F17Y, pH 7.0; 10.2210/pdb7QX7/pdb), 7QYD (Cry11Ba WT, pH 6.5; 10.2210/pdb7QYD/pdb), 7R1E (Cry11Ba WT, pH 10.4; 10.2210/pdb7R1E/pdb). 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 in a combined Source Data file.

1004 **References**

1005 1. Bravo, A., Likitvivatanavong, S., Gill, S. S. & Soberon, M. Bacillus thuringiensis: A
1006 story of a successful bioinsecticide. *Insect Biochem. Mol. Biol.* **41**, 423–431 (2011).
1007 2. Vachon, V., Laprade, R. & Schwartz, J. L. Current models of the mode of action of
1008 Bacillus thuringiensis insecticidal crystal proteins: a critical review. *J. Invertebr. Pathol.* **111**,
1009 1–12 (2012).
1010 3. Delecluse, A., Rosso, M. L. & Ragni, A. Cloning and expression of a novel toxin gene
1011 from Bacillus thuringiensis subsp. jegathesan encoding a highly mosquitocidal protein. *Appl*
1012 *Env. Microbiol* **61**, 4230–5 (1995).
1013 4. Federici, B. A., Park, H.-W. & Sakano, Y. Insecticidal Protein Crystals of Bacillus
1014 thuringiensis. in *Inclusions in Prokaryotes* (ed. Shively, J. M.) 195–236 (Springer, 2006).
1015 doi:10.1007/7171_008/Published.
1016 5. Boonserm, P., Mo, M., Angsuthanasombat, C. & Lescar, J. Structure of the functional
1017 form of the mosquito larvicidal Cry4Aa toxin from Bacillus thuringiensis at a 2.8-Angstrom
1018 resolution. *J. Bacteriol.* **188**, 3391–3401 (2006).
1019 6. Boonserm, P., Davis, P., Ellar, D. J. & Li, J. Crystal structure of the mosquito-
1020 larvicidal toxin Cry4Ba and its biological implications. *J. Mol. Biol.* **348**, 363–382 (2005).
1021 7. Cohen, S. *et al.* Cyt1Aa Toxin: Crystal Structure Reveals Implications for Its
1022 Membrane-Perforating Function. *J. Mol. Biol.* **413**, 804–814 (2011).
1023 8. Gutierrez, P., Alzate, O. & Orduz, S. A theoretical model of the tridimensional
1024 structure of Bacillus thuringiensis subsp. medellin Cry 11Bb toxin deduced by homology
1025 modelling. *Mem Inst Oswaldo Cruz* **96**, 357–64 (2001).
1026 9. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**,
1027 73–77 (2011).
1028 10. Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2**, 246–
1029 255 (2015).
1030 11. Boutet, S., Fromme, P. & Hunter, M. S. *X-ray Free Electron Lasers - A Revolution in*
1031 *Structural Biology*. vol. 1 (Springer International Publishing, 2018).
1032 12. Tetreau, G. *et al.* Serial femtosecond crystallography on in vivo-grown crystals drives
1033 elucidation of mosquitocidal Cyt1Aa bioactivation cascade. *Nat. Commun.* **11**, 1153 (2020).
1034 13. Li, J. D., Carroll, J. & Ellar, D. J. Crystal structure of insecticidal delta-endotoxin from
1035 Bacillus thuringiensis at 2.5 A resolution. *Nature* **353**, 815–821 (1991).
1036 14. Morse, R. J., Yamamoto, T. & Stroud, R. M. Structure of Cry2Aa Suggests an
1037 Unexpected Receptor Binding Epitope. *Structure* **9**, 409–417 (2001).
1038 15. Grochulski, P. *et al.* Bacillus thuringiensis CryIA(a) Insecticidal Toxin: Crystal
1039 Structure and Channel Formation. *J. Mol. Biol.* **254**, 447–464 (1995).
1040 16. Engilberge, S. *et al.* Crystallophore: a versatile lanthanide complex for protein
1041 crystallography combining nucleating effects, phasing properties, and luminescence. *Chem.*
1042 *Sci.* **8**, 5909–5917 (2017).
1043 17. Engilberge, S. *et al.* Protein crystal structure determination with the crystallophore, a
1044 nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722–731 (2019).
1045 18. Xu, C., Wang, B. C., Yu, Z. & Sun, M. Structural insights into Bacillus thuringiensis
1046 Cry, Cyt and parasporin toxins. *Toxins* **6**, 2732–70 (2014).
1047 19. Promponas, V. J. *et al.* CAST: an iterative algorithm for the complexity analysis of
1048 sequence tracts. *Bioinformatics* **16**, 915–922 (2000).
1049 20. Wootton, J. C. & Federhen, S. Statistics of local complexity in amino acid sequences
1050 and sequence databases. *Comput. Chem.* **17**, 149–163 (1993).
1051 21. Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light
1052 Source. *J. Synchrotron Radiat.* **22**, 514–519 (2015).
1053 22. Sierra, R. G. *et al.* Concentric-flow electrokinetic injector enables serial
1054 crystallography of ribosome and photosystem II. *Nat. Methods* **13**, 59–62 (2016).
1055 23. DePonte, D. P. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet
1056 streams. *J. Phys. -Appl. Phys.* **41**, 7 (2008).
1057 24. Leaver-Fay, A. *et al.* Rosetta3: An Object-Oriented Software Suite for the Simulation

1058 and Design of Macromolecules. *Methods Enzymol.* **487**, 545–574 (2011).
1059 25. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence data bank and its
1060 supplement TrEMBL in 1999. *Nucleic Acids Res.* **27**, 49–54 (1999).
1061 26. Pardo-Lopez, L., Soberon, M. & Bravo, A. Bacillus thuringiensis insecticidal three-
1062 domain Cry toxins: mode of action, insect resistance and consequences for crop protection.
1063 *FEMS Microbiol. Rev.* **37**, 3–22 (2013).
1064 27. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for
1065 macromolecular crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 143–151 (2018).
1066 28. Colletier, J.-P. *et al.* De novo phasing with X-ray laser reveals mosquito larvicide
1067 BinAB structure. *Nature* **539**, 43–47 (2016).
1068 29. Oestergaard, J., Ehlers, R. U., Martinez-Ramirez, A. C. & Real, M. D. Binding of
1069 Cyt1Aa and Cry11Aa toxins of Bacillus thuringiensis serovar israelensis to brush border
1070 membrane vesicles of Tipula paludosa (Diptera : Nematocera) and subsequent pore
1071 formation. *Appl. Environ. Microbiol.* **73**, 3623–3629 (2007).
1072 30. Yamagiwa, M., Sakagawa, K. & Sakai, H. Functional Analysis of Two Processed
1073 Fragments of Bacillus thuringiensis Cry11A Toxin. *Biosci. Biotechnol. Biochem.* **68**, 523–528
1074 (2014).
1075 31. Munoz-Garay, C. *et al.* Oligomerization of Cry11Aa from Bacillus thuringiensis Has an
1076 Important Role in Toxicity against Aedes aegypti. *Appl. Environ. Microbiol.* **75**, 7548–7550
1077 (2009).
1078 32. Fernandez, L. E. *et al.* Cry11Aa toxin from Bacillus thuringiensis binds its receptor in
1079 Aedes aegypti mosquito larvae through loop alpha-8 of domain II. *Febs Lett.* **579**, 3508–3514
1080 (2005).
1081 33. Carmona, D. *et al.* Dominant Negative Phenotype of Bacillus thuringiensis Cry1Ab,
1082 Cry11Aa and Cry4Ba Mutants Suggest Hetero-Oligomer Formation among Different Cry
1083 Toxins. *PLoS ONE* **6**, e19952 (2011).
1084 34. Du, J. P., Knowles, B. H., Li, J. & Ellar, D. J. Biochemical characterization of Bacillus
1085 thuringiensis cytolytic toxins in association with a phospholipid bilayer. *Biochem. J.* **338**, 185–
1086 193 (1999).
1087 35. Ntountoumi, C. *et al.* Low complexity regions in the proteins of prokaryotes perform
1088 important functional roles and are highly conserved. *Nucleic Acids Res.* **47**, 9998–10009
1089 (2019).
1090 36. Henrich, B. *et al.* The adaptive gain integrating pixel detector AGIPD a detector for
1091 the European XFEL. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect.*
1092 *Assoc. Equip.* **633**, S11–S14 (2011).
1093 37. Likitvivatanavong, S. *et al.* Multiple Receptors as Targets of Cry Toxins in
1094 Mosquitoes. *J. Agric. Food Chem.* **59**, 2829–2838 (2011).
1095 38. Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A. & Soberon, M. A GPI-
1096 anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in Aedes
1097 aegypti larvae. *Biochem. J.* **394**, 77–84 (2006).
1098 39. Bourgouin, C., Delecluse, A. & Rapoport, G. Specificity and Synergism of Bacillus
1099 thuringiensis israelensis Toxins on Mosquito Larvae. in *Bacterial Protein Toxins* (ed.
1100 Fehrenbach) 225–226 (Gustav Fisher, 1988).
1101 40. Lopez-Diaz, J. A., Emiliano Canton, P., Gill, S. S., Soberon, M. & Bravo, A.
1102 Oligomerization is a key step in Cyt1Aa membrane insertion and toxicity but not necessary to
1103 synergize Cry11Aa toxicity in Aedes aegypti larvae. *Environ. Microbiol.* **15**, 3030–3039
1104 (2013).
1105 41. Pérez, C. *et al.* Bacillus thuringiensis subsp. israelensis Cyt1Aa synergizes Cry11Aa
1106 toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. U. S. A.* **102**,
1107 18303–18308 (2005).
1108 42. Sawaya, M. R. *et al.* Protein crystal structure obtained at 2.9 Å resolution from
1109 injecting bacterial cells into an X-ray free-electron laser beam. *Proc. Natl. Acad. Sci.* **111**,
1110 12769–12774 (2014).
1111 43. Galitsky, N. *et al.* Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of
1112 Bacillus thuringiensis. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1101–1109 (2001).

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

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

1223 94. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching
1224 for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
1225 95. Prakash, A., Jeffryes, M., Bateman, A. & Finn, R. D. The HMMER Web Server for
1226 Protein Sequence Similarity Search. *Curr. Protoc. Bioinforma.* **60**, 3.15.1-3.15.23 (2017).
1227 96. Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–
1228 W204 (2018).
1229 97. Savi, M. K., Mangamana, E. T., Deguenon, J. M., Hounmenou, C. G. & Kakaï, R. G.
1230 Determination of Lethal Concentrations Using an R Software Function Integrating the Abbott
1231 Correction. *J. Agric. Sci. Technol. A* **7**, (2017).
1232 98. Wheeler, M. W., Park, R. M. & Bailer, A. J. Comparing median lethal concentration
1233 values using confidence interval overlap or ratio tests. *Env. Toxicol Chem* **25**, 1441–4 (2006).
1234 99. R. *R: a language and environment for statistical computing.* (R Foundation for
1235 Statistical Computing, 2011).
1236 100. Rosenfeld, J., Capdevielle, J., Guillemot, J. C. & Ferrara, P. In-gel digestion of
1237 proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis.
1238 *Anal. Biochem.* **203**, 173–179 (1992).
1239 101. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry
1240 of intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).
1241 102. Marty, M. T. *et al.* Bayesian Deconvolution of Mass and Ion Mobility Spectra: From
1242 Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
1243

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.

1309 **Figures captions :**

1310

1311 **Fig. 1. Crystals and overall fold of Cry11 toxins. a-b,** scanning (left; SEM) and transmission
1312 (middle, right; TEM) electron micrographs of gold plated and negatively-stained Cry11Aa (**a**)
1313 and Cry11Ba (**b**) crystals, respectively. The right panels show a close-up view of the crystal
1314 surface. **c,** Cry11Aa crystal structure, depicted as a cartoon. Domain I is shown in blue; domain
1315 II is shown in orange except for the $\alpha_h\beta_h$-handle and βpin which are shown in purple and red,
1316 respectively; domain III is shown in pink. **d**, Topology diagram of a Cry11Aa dimer. Domain I
1317 is shown in green, except for central helix α5, which is shown in blue: domain II is shown in
1318 magenta, except for the $\alpha_h\beta_h$-handle, which is shown in purple; and domain III is shown in
1319 cyan, respectively. The two monomers in a dimer assemble via the β_{pin}, resulting in the
1320 formation of a large β-sheet.

1321

1322 **Fig. 2. Monomer interactions in Cry11Aa and Cry11Ba. a,** Cry11Aa crystal packing,
1323 coloured according to sequence (from blue to red) indicating the domain-based assembly; and
1324 coloured according to tetramer assembly (see panel (**b**)). The highlighted areas indicate the
1325 regions shown in (**b**) (full line) and (**c**) (dashed line). **b,** Cry11Aa tetramer with zoom on each
1326 of the three interfaces identified by PISA (interface #1, #3 and #6), with the involved residues
1327 depicted as spheres. For interfaces with hydrogen and/or salt bridges (see g), an additional
1328 (right) image shows only those residues that make up these interactions. **c,** Cry11Aa crystal
1329 assembly by interactions between neighbouring tetramers, formed by interface #2, #4 and #5,
1330 visualized as in b. **d,** Cry11Ba crystal packing, coloured as in (**a**). **e,** Cry11Ba tetramer with
1331 zoom on the interfaces as in (**b**). **f,** Cry11Ba crystal assembly, visualized as in (**c**). As
1332 compared to Cry11Aa, Cry11Ba crystals contain an additional interface #7 between an A-B
1333 pair from two neighbouring tetramers. **g,** interface statistics as identified by PISA for Cry11Aa
1334 (blue) and Cry11Ba (red).

1335

1336 **Fig. 3. Point-mutations of Cry11Aa affect the shape, size and pH-sensitivity of *in vivo*-**
1337 **grown nanocrystals. a, C**rystals from mutants exhibit similar sigmoidal patterns of crystal
1338 solubilization as a function of pH, except F17Y and E583Q that are more and less sensitive to
1339 pH, respectively. Error bars indicate the standard error of the measurements. **b,** Cry11Aa WT
1340 and mutants exhibit similar heat stability. As expected, toxins are more stable (+ 17.5 ± 0.3°C)
1341 in their crystalline than soluble form, irrespective of the mutation. **c,** Visualization of a
1342 representative crystal for Cry11Aa WT (black) and mutants F17Y (red), Y272Q (brown), Y349F
1343 (purple), Y449F (blue), D507N-D514N (orange) and E583Q (green) by SEM (scale bar = 500
1344 nm). **d,** Crystals of Y449F, F17Y and E583Q imaged by AFM were all smaller in length (L),
1345 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.

Figures

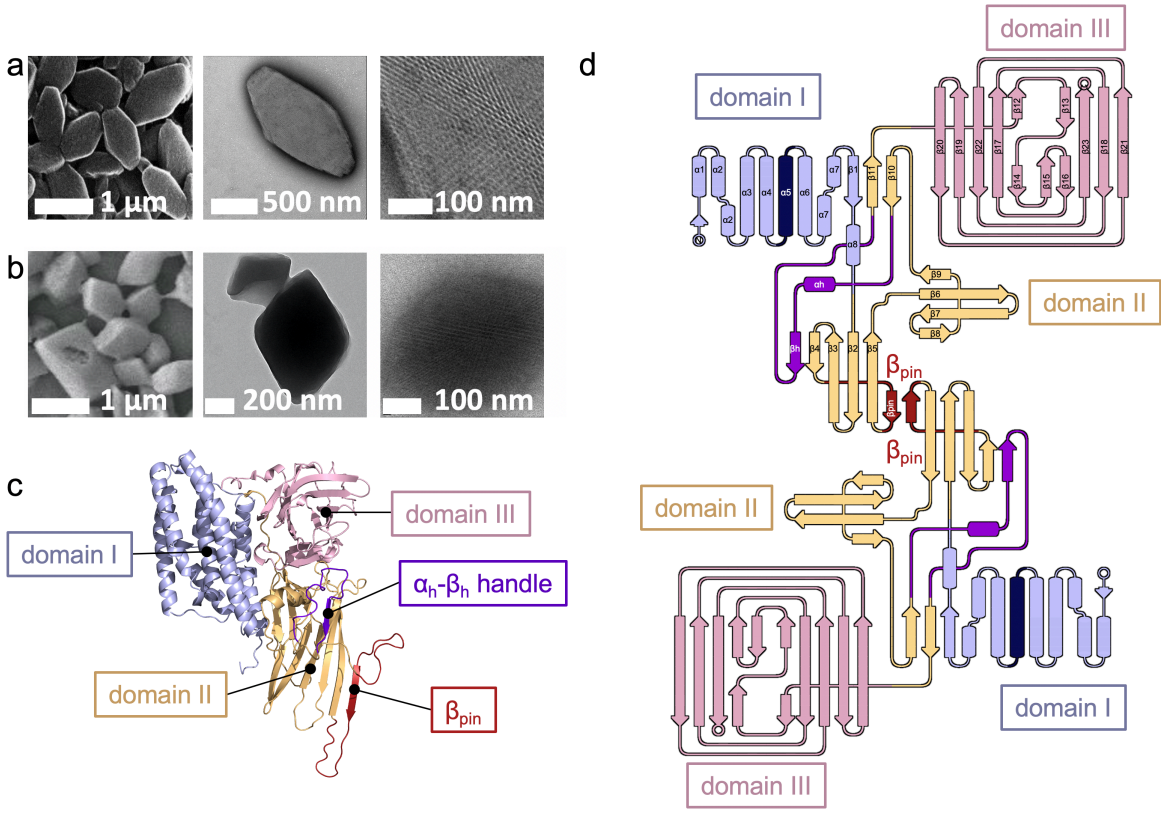

Fig. 1. Crystals and overall fold of Cry11 toxins. a-b, scanning (left; SEM) and transmission (middle, right; TEM) electron micrographs of gold plated and negatively-stained Cry11Aa (**a**) and Cry11Ba (**b**) crystals, respectively. The right panels show a close-up view of the crystal surface. **c,** Cry11Aa crystal structure, depicted as a 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.

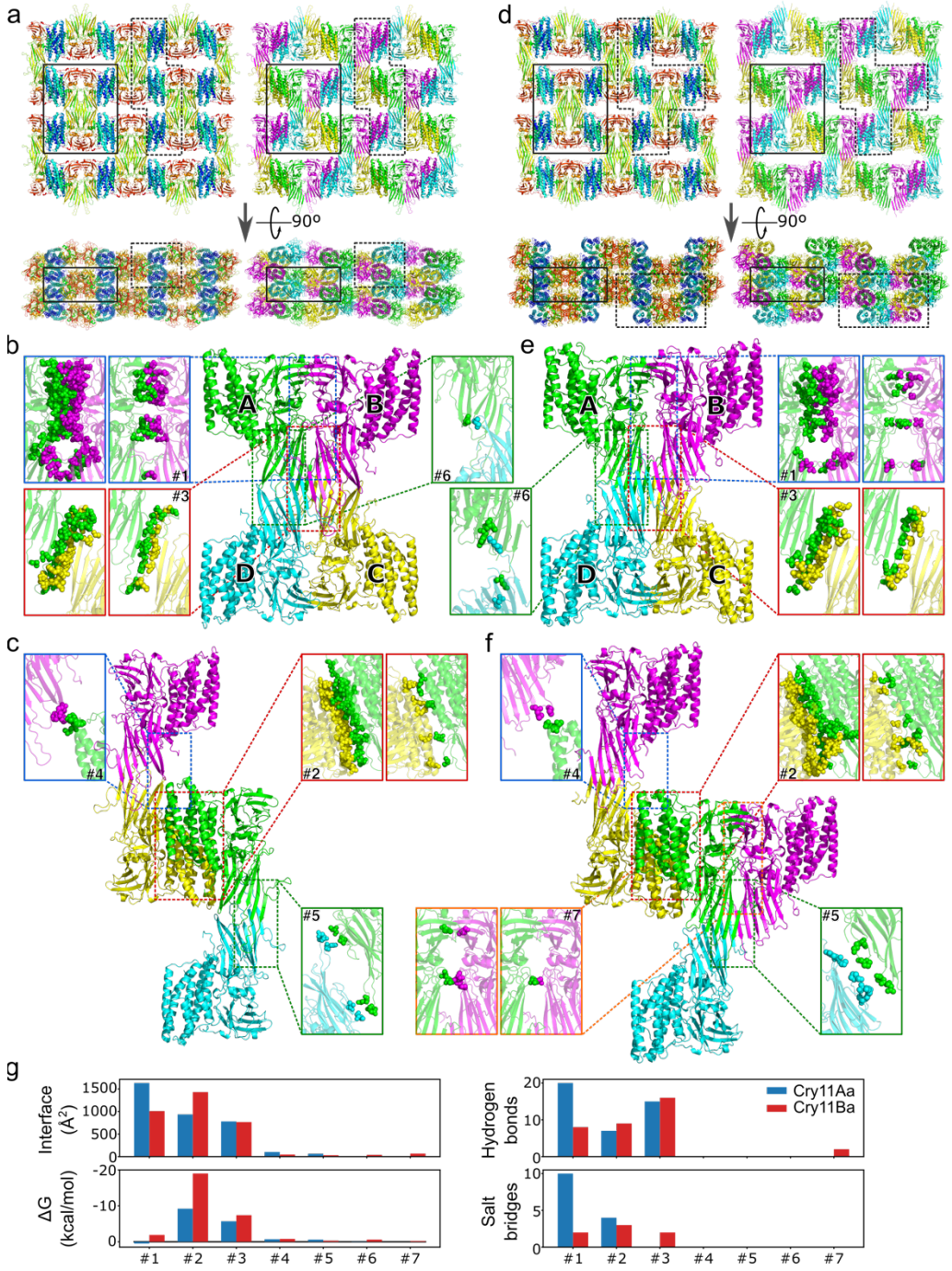

Fig. 2. Monomer interactions in Cry11Aa and Cry11Ba. a, Cry11Aa crystal packing, coloured according to sequence (from blue to red) indicating the domain-based assembly; and coloured according to tetramer assembly (see panel (**b**)). The highlighted areas indicate the regions shown in (**b**) (full line) 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).

1354

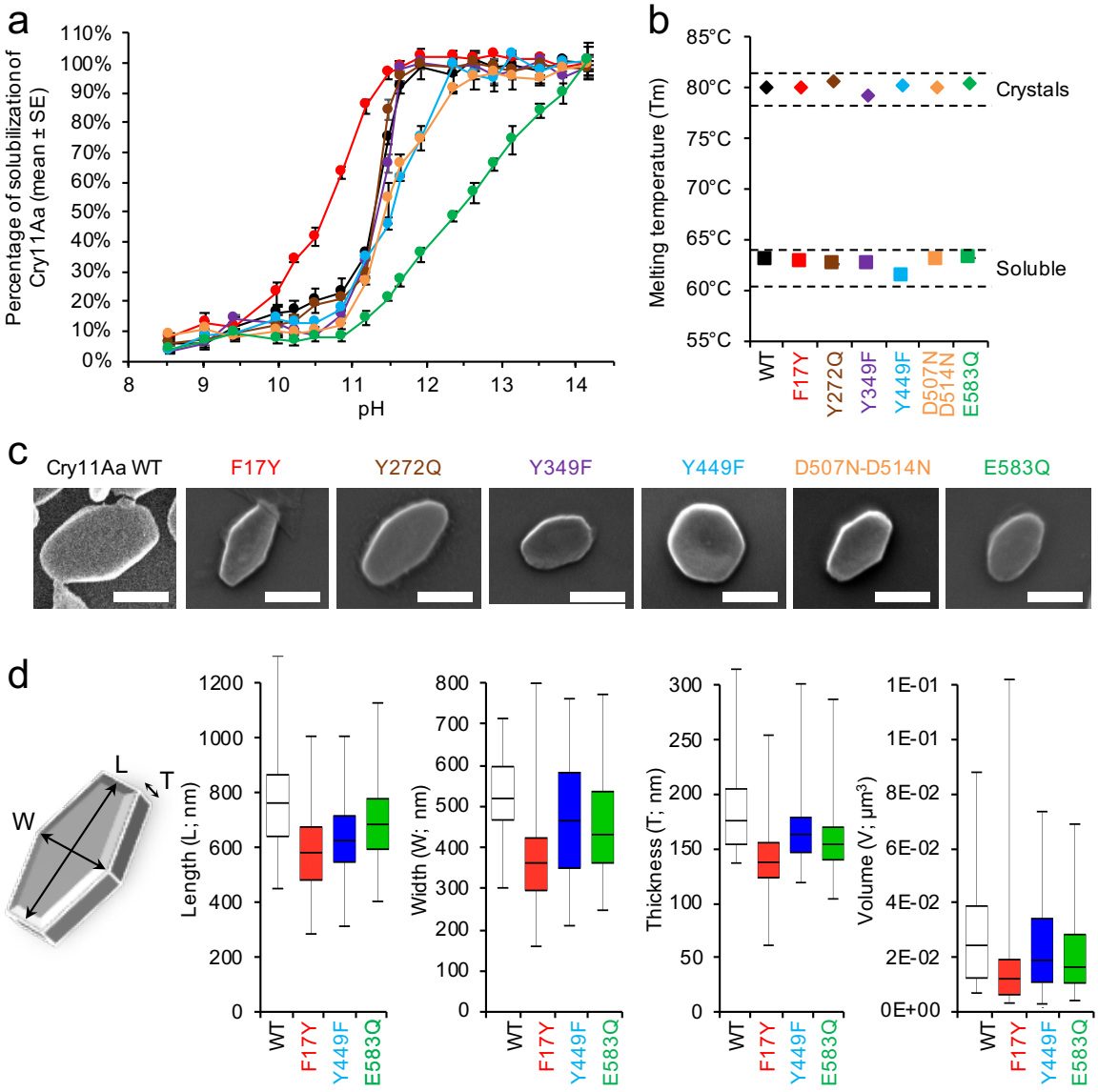

Fig. 3. Point-mutations of Cry11Aa affect the shape, size and pH-sensitivity of *in vivo*-grown nanocrystals. a, Crystals from mutants exhibit similar sigmoidal patterns of crystal solubilization as a function of pH, except F17Y and E583Q that are more and less sensitive to pH, respectively. Error bars 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.

1355
1356
1357

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**	**7R1E**
Data collection				
Space group	*I* 2 2 2	*I* 2 2 2	*P* 2_1 2_1 2	*P* 2_1 2_1 2
Cell dimensions (Å)	57.64 ± 0.19 155.69 ± 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)
<I/σ (I)>	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 (º)	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

1359

1360

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

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

1361