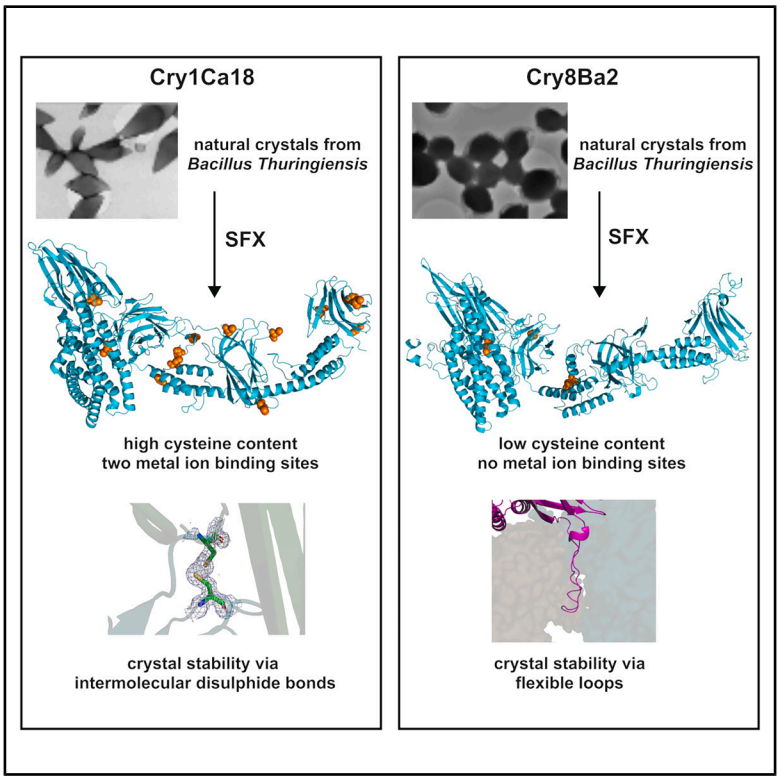
Structure

Article

The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*

Graphical abstract



Highlights

- *In vivo* Cry nanocrystals show distinct packing by XFEL crystallography
- Cry1Ca18 packing uses disulfide bonds; Cry8Ba2 relies on noncovalent contacts
- Cry1Ca18 has conserved Zn/Ca binding sites and a rare non-proline cis peptide
- Alternative crystal-packing drive environmental stability and *in vivo* toxin release

Authors

Hannah L. Best, Lainey J. Williamson, Adam B. Cutts, ..., Pierre J. Rizkallah, Colin Berry, Dominik Oberthür

Correspondence

berry@cardiff.ac.uk (C.B.), dominik.oberthuer@desy.de (D.O.)

In brief

Best et al. reveal how *Bacillus thuringiensis* toxins naturally pack into stable protein crystals. Using ultrafast X-ray methods, they show that different toxins use distinct molecular strategies to control crystal stability and dissolution, informing our understanding of how these biopesticides remain stable in the environment yet activate inside insects.

Best et al., 2026, Structure *34*, 1–14
May 7, 2026
https://doi.org/10.1016/j.str.2026.01.014

CellPress

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

Structure

CellPress
OPEN ACCESS

Article

The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*

Hannah L. Best,[1,2,15] Lainey J. Williamson,[1,15] Adam B. Cutts,[1] Marina Galchenkova,[3] Oleksandr Yefanov,[3] Nicole Bryce-Sharron,[4] Emily A. Heath,[1] Raphael de Wijn,[5] Robin Schubert,[5] Anna Munke,[3] Alessandra Henkel,[3] Bjarne Klopprogge,[3,6] T. Emilie S. Scheer,[3] Viviane Kremling,[3] Salah Awel,[3] Gisel Pena,[3] Juraj Knoska,[3] Anusha Keloth,[3] Julia Maracke,[3] Romain Letrun,[5] Egor Sobolev,[5] Johan Bielecki,[5] Diogo Melo,[5] Sravya Kantamneni,[5] Katerina Doerner,[5] Marco Kloos,[5] Joachim Schulz,[5] P. Lourdu Xavier,[3,7] Marius Lauffer,[8,9,10] Maite Villanueva,[11] Primitivo Caballero,[12] Helen Waller-Evans,[1,2] Emyr Lloyd-Evans,[1,2] Charlotte Uetrecht,[9,10] Richard Bean,[5] Henry N. Chapman,[3,6,13] Neil Crickmore,[4] Pierre J. Rizkallah,[14] Colin Berry,[1,16,*] and Dominik Oberthür[3,6,*]

[1]School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK
[2]Medicines Discovery Institute, Cardiff University, Cardiff CF10 3AX, UK
[3]Center for Free Electron Laser Science CFEL, Deutsches Elektronen-Synchrotron DESY, 22607 Hamburg, Germany
[4]School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK
[5]European XFEL GmbH, 22869 Schenefeld, Germany
[6]Centre for Ultrafast Imaging, Universität Hamburg, 22761 Hamburg, Germany
[7]Max-Planck Institute for the Structure and Dynamics of Matter, 22761 Hamburg, Germany
[8]Leibniz Institute of Virology, 20251 Hamburg, Germany
[9]CSSB Centre for Structural Systems Biology, Deutsches Elektronen-Synchrotron DESY & Leibniz Institute of Virology (LIV) & University of Lübeck, 22607 Hamburg, Germany
[10]Institute of Chemistry and Metabolomics, University of Lübeck, 23562 Lübeck, Germany
[11]Bioinsectis SL, Plaze de cein, 5, Nave A14, Poligono de Mocholi, 31110 Nain, Navarra, Spain
[12]Institute for Multidisciplinary Research in Applied Biology, Avda. De Pamplona 123, 31192 Mutilva Baja, Spain
[13]Department of Physics, Universität Hamburg, 22761 Hamburg, Germany
[14]School of Medicine, Cardiff University, Cardiff CF10 3AX, UK
[15]These authors contributed equally
[16]Lead contact
*Correspondence: berry@cardiff.ac.uk (C.B.), dominik.oberthuer@desy.de (D.O.)
https://doi.org/10.1016/j.str.2026.01.014

SUMMARY

Bacillus thuringiensis (Bt) strains naturally produce pesticidal proteins as nanocrystalline inclusions that are extraordinarily stable in aqueous environments, but which dissolve selectively at specific pH conditions. These proteins have been used in agriculture for >50 years and are critical to global food security. The majority of previously determined Bt Cry protein structures lack the extended C-terminal "crystallization domain," which is thought to stabilize crystal packing and control selective solubility in insect targets, often via manipulation of disulfide bridges. It has also recently been shown to influence toxicity and target specificity. Here, we use serial femtosecond crystallography (SFX) to determine high-resolution full-length native structures of Cry1Ca18 (1.65 Å) and Cry8Ba2 (2.27 Å) in their natural nanocrystalline state. Differences in cysteine content (19 versus 4 residues) reveal distinct *in vivo* crystal-stabilization strategies. Understanding Bt toxin domain architecture and natural crystal formation is essential for improving biopesticide design and advancing agricultural genetic engineering.

INTRODUCTION

Bacillus thuringiensis (Bt) is an entomopathogenic bacterium, strains of which can produce a variety of protein toxins during sporulation and sequester them in the form of crystal inclusions. Bt-produced insecticidal proteins demonstrate specificity to different insect orders including, *Coleoptera*, *Hymenoptera*, *Diptera*, and *Lepidoptera*, while demonstrating human, veterinary, and environmental safety.[1–3] This has allowed the substitution of broad-spectrum chemical insecticides with specific Bt toxins in insect control programs for over 50 years.[4,5] The exact mechanisms behind Bt protein(s) toxicity remain unclear, and increasing understanding is critical to counteract emerging field resistance and for the development of new Bt proteins.

Bt crystal proteins have recently been recategorized and named according to the structural class to which they belong.[6]

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

The largest family of these proteins is the Cry family, a group with three structural domains in the activated proteins. Cry proteins are produced as protoxins either of approximately 65 or 130 kDa, with the latter exhibiting an extended C-terminus, often referred to as the "crystallization domain" (although in fact this region consists of 4 structural domains, DIV–DVII). The mode of action of Cry proteins, which has largely been studied in lepidopteran insects, begins with dissolution of the crystalline inclusions in the alkaline midgut environment. Solubilized proteins are subsequently cleaved by midgut proteinases to produce the active form. For the 130 kDa long-form Cry proteins, activation includes the proteolytic cleavage of the extended C-terminus, to produce a 60–70 kDa proteinase-stable toxin core. The activated protein then binds to gut-cell target receptors, undergoes oligomerization leading to pore-formation, cell lysis, and larval death.[7,8] The precise role of the C-terminal "crystallization domain" is unclear, with many earlier studies suggesting it is dispensable for toxicity, yet important for crystal formation, stability, and selective solubility in the appropriate region of the insect gut.[9,10] However, an increasing number of studies show the C-terminus may be able to play an active role in the mode of action,[11–14] with toxin activation shown, in different studies, both to enhance and reduce the species selectivity.[15–18]

Deducing the *in vivo* structure is essential for understanding the structure function relationship—especially for the extended C-terminal domain. Determining the structure of naturally produced Bt proteins *in vivo* requires the use of world-leading technology, due to the small size of the nanocrystals (approx. 500 nm long), and the difficulty associated with recrystallizing solubilized protoxins preventing their use in conventional crystallography.[19,20] Previously, two full-length 3D-Cry structures have been identified. First Cry1Ac, where 14 out of a total 16 cysteine residues were mutated to alleviate aggregation during *in vitro* recrystallization, resulting in a non-native structure (Cry1Ac-Δ14C, 2.78 Å).[21] More recently, the first *in vivo* native low-resolution (3.60 Å) structure of Cry1Aa was determined, allowing characterization of packing and disulfide bond formation in the natural crystal.[22]

At present, there are no published structures for any Cry1C or Cry8B proteins in either a full-length or activated form. Both proteins represent important tools for insect control, with Cry1C proteins demonstrating activity against a range of lepidopteran pests as well as some Diptera, and Cry8B proteins demonstrating activity against several coleopteran species (*Cotinis* spp, *Cyclocephala borealis*, and *Anthonomus grandis*).[3] Cry1C proteins are actively used in pest control to protect fruiting vegetables, as part of the approved Bt aizawai strain (ABTS-1857). To probe the structural differences behind this target specificity and gain insight into the role of the C-terminus in natural crystals, we applied serial femtosecond crystallography (SFX) at the European X-ray Free-Electron Laser (EuXFEL) facility to natural crystals, containing either Cry1Ca18 (cysteine rich, 19 in total), or Cry8Ba2 (low cysteine content, 4 in total). We also varied the pH to mimic the alkaline environment found in target insect species to investigate early structural events leading to dissolution. We present two novel structures of Cry1Ca18 at pH 7 (1.80 Å, PDB: 9H99) and pH 9 (1.65 Å, PDB: 9QXQ), and two novel and distinct Cry8Ba2 structures from the same crystal sample at pH 7 (2.27 Å, PDB: 9H9A and 9H9B). The Cry8Ba2 structures include two regions that are not resolved in any of the previously derived full-length structures or in the Cry1Ca18 structure presented here. This work has enabled investigation of the network of molecular interactions present in long-form Cry proteins, and how native crystal packing can support stability and function.

RESULTS

Structures of Cry1Ca18 and Cry8Ba2 show a conserved domain architecture

Diffraction data (Table 1) were collected from Cry8Ba2 and Cry1Ca18 nanocrystals that had been previously visualized by transmission electron microscopy (Figure S1). In common with other Bt long-form 3D-Cry protoxins, the structures of both Cry1Ca18 and Cry8Ba2 contain the characteristic architecture, comprising seven distinct domains (DI–DVII), with domain 1–3 (DI–DIII) forming the toxin core and domains 4–7 (DIV–DVII) the extended C-terminal half that is removed upon protoxin activation (Figures 1A–1D). Interestingly, within the Cry8Ba2 dataset, two distinct populations of crystals were identified and could be resolved separately. The first crystal form was found to exhibit an extended (Ext) *c*-axis (the axis which runs down the length of the unit cell, the smallest repeating portion of the crystal lattice) and is hence referred to as Cry8Ba2-Ext (PDB: 9H9A). The second crystal form was found to exhibit a more compact (Com) *c*-axis and is hence referred to as Cry8Ba2-Com (PDB: 9H9B; Figure 1C).

Notably, the first approx. 30 residues of DI were not observed in any of the Cry8Ba2 or Cry1Ca18 structures and are also absent from the previously published Cry1Ac structures, indicating that this region may be inherently disordered. Domain I is associated with pore formation and is generally a seven α-helix bundle. In both Cry1Ca18 and Cry8Ba2, α2 is split into α2a and α2b, a structural characteristic present in the majority of resolved Cry proteins, including the closely related Cry1Ac-Δ14C (PDB: 4W8J) and activated Cry8Ea1 (PDB: 3EB7) structures. Cry8Ba2 DI also reveals the presence of an extended N-terminal region that does not appear in the chymotrypsin activated Cry8Ea1 structure, likely due to it being cleaved during proteolytic processing. The placement of the Cry8Ba2 extended N-terminal region represents one of the key differences between the two Cry8Ba2 crystal forms (Figure 1D). In Cry8Ba2-Com, the compact-cell model, the N-terminal region is tucked into the rest of the molecule, while in Cry8Ba2-Ext, the extended-cell model, the N-terminal region is projected out by a hinge opening motion around region Ala50 to Ser60. These differences result in an extension on average of ∼8 Å per Cry8Ba2-Ext monomer, or ∼33 Å in a unit cell *c*-axis.

Core toxin DII[23] and DIII[24,25] are known to influence Cry toxin specificity, in particular, several exposed loops in DII have been shown to have important roles in target selectivity.[23,26] Since no other Cry1C or Cry8B proteins are currently available, we compared the DII loops of our structures with those of closely related structures Cry1Ac-Δ14C (PDB: 4W8J) and Cry8Ea1 (PDB: 3EB7), respectively (Figure S2). The sequence and length of these loops was distinct in these comparisons, which would be consistent with the potential to bind different targets. This may be a factor underlying observations that Cry1Ac or Cry1Ca resistant insects may not show cross-resistance.[27]

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

CellPress
OPEN ACCESS

Table 1. Data collection and refinement statistics for Cry8Ba2 and Cry1Ca18

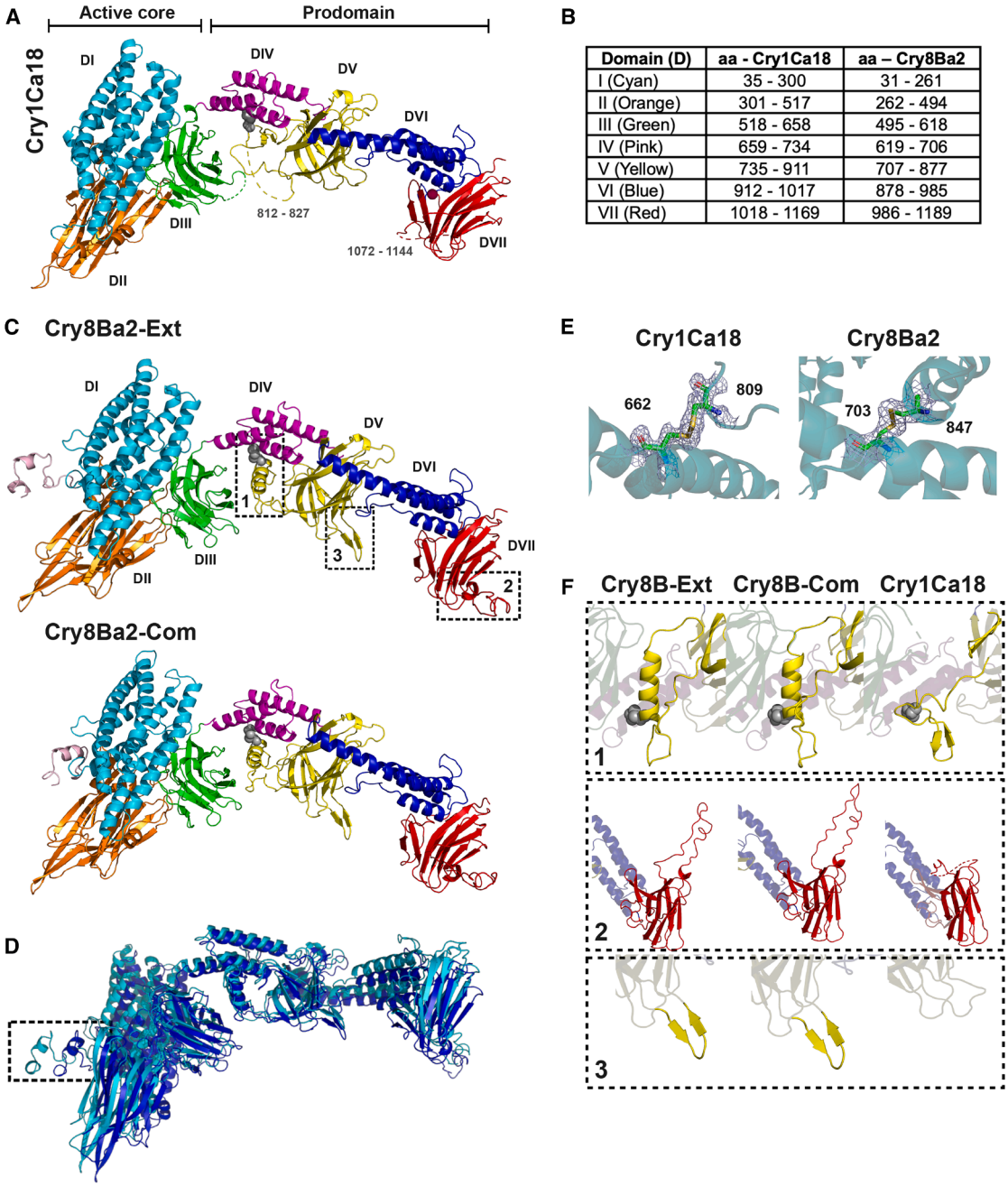
Data collection				
PDB ID	9H9B	9H9A	9H99	9QXQ
Beamline	SPB/SFX at European XFEL	SPB/SFX at European XFEL	SPB/SFX at European XFEL	SPB/SFX at European XFEL
X-ray energy (keV)	9.3	9.3	9.3	9.3
Wavelength (Å)	1.33	1.33	1.33	1.33
Images collected	9,977,877	9,977,877	1,875,129	3,225,045
Number of hits	73,131	73,131	147,768	802,849
Number of indexed	33,552	30,123	111,631	458,966
Crystal data (numbers in brackets refer to outer resolution shell)				
Crystal form	**Cry8Ba2-Com**	**Cry8Ba2-Ext**	**Cry1Ca18 pH7**	**Cry1Ca18 pH9**
a, *b*, *c* (Å)	93.32, 93.32, 275.51	92.23, 92.23, 308.23	88.16, 88.16, 266.43	88.8, 88.8, 267.82
	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Space group	P 4_12_12	P 4_12_12	P 4_12_12	P 4_12_12
Resolution (Å)	28.11–2.27	28.17–2.27	36.19–1.80	26.91–1.65
Outer shell	2.27–2.32	2.27–2.32	1.82–1.80	1.69–1.65
R-split (%)	13.72 (212.47)	14.45 (196.70)	7.00 (40.03)	4.88 (203.8)
CC*	0.996 (0.519)	0.996 (0.499)	0.999 (0.94)	0.999 (0.55)
$I/\sigma(I)$	6.41 (0.65)	6.11 (0.69)	13.75 (2.77)	23.21 (0.44)
Completeness (%)	99.96 (99.38)	99.97 (99.59)	99.98 (100.0)	99.42 (92.51)
Multiplicity	242.89 (11.0)	210.01 (11.6)	207.97 (59.6)	1379.3 (93.5)
Unique reflections	57,348 (3,705)	62,717 (4,081)	98,288 (6,425)	122,231(8,223)
Wilson B-factor (Å^2)	37.7	37.0	19.0	17.8
Refinement statistics				
Refined atoms	9,360	9,353	9,318	9,526
Protein atoms	9,185	9,176	8,507	8,809
Non-protein atoms	0	0	2	2
Water molecules	175	177	813	718
R-work reflections	57,279	59,478	93,379	128,618
R-free reflections	2,896	3,164	4,841	6,387
R-work/R-free (%)	18.0/22.6	18.0/22.4	18.9/22.9	17.8/20.6
rms deviations (numbers in brackets)				
Bond lengths (Å)	0.0060 (0.012)	0.0063 (0.012)	0.010 (0.012)	0.010 (0.012)
Bond angles (°)	1.577 (1.827)	1.643 (1.827)	1.743 (1.824)	1.745 (1.826)
Coordinate error (Å)[a]	0.201	0.189	0.097	0.089
Mean *B* value (Å^2)	56.0	55.2	18.835	25.5
Ramachandran Statistics (PDB Validation)				
Favored/allowed/Outliers	1,068/50/16	1,084/40/9	1,024/24/1	1,056/22/2
%	94/4/1	96/4/1	98/2/0	98/2/0

[a]Coordinate estimated standard uncertainty in (Å), calculated based on maximum likelihood statistics.

As observed with Cry1Ac-Δ14C and Cry1Ac structures, the extended C-terminal domain, or "crystallization" domain of Cry8Ba2 and Cry1Ca18 is composed of four domains (DIV–DVII). Comparing the Cry1Ca18 structure to the previously published full-length Cry1 structures, shows a higher degree of structural conservation in the crystallization prodomain compared to the active core (Table S1). For example, Cry1Ca18 and Cry1Ac display an all-atom RMSD of 2.7 Å, while alignment of the prodomains (DIV–DVII) produced an all-atom RMSD of 1.3 Å, and alignment of the activated core (DI–DIII) produced an all-atom RMSD of 3.4 Å. Comparing between subfamilies (Cry1 vs. Cry8), shows no consistent pattern of higher structural conservation between prodomains (e.g., Cry8Ba2 and Cry1Ac share higher structural similarity within the core domains; Table S1). Individual Bt strains sometimes produce multiple variants of Cry1 proteins and conservation of DIV–DVII among Cry1 family proteins may facilitate the formation of heterologous Cry1 co-crystals in such strains.

Although the number of cysteine residues is considerably different between the Cry1Ca18 and Cry8Ba2 sequences (see section below), there is one conserved intramolecular disulfide bond present between DIV and DV in the structures

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014



Domain (D)	aa - Cry1Ca18	aa – Cry8Ba2
I (Cyan)	35 - 300	31 - 261
II (Orange)	301 - 517	262 - 494
III (Green)	518 - 658	495 - 618
IV (Pink)	659 - 734	619 - 706
V (Yellow)	735 - 911	707 - 877
VI (Blue)	912 - 1017	878 - 985
VII (Red)	1018 - 1169	986 - 1189

Figure 1. Comparison of Cry1Ca18 and Cry8Ba2 domain architecture

(A) Cry1Ca18 and (C) Cry8Ba2 share seven conserved domains (DI–DVII), as detailed in table.

(B) DI–DIII are present in the active core toxin, and DIV–DVII form the prodomain.

The boxed regions in (C) highlight three regions with notable structural differences between Cry1Ca18 and Cry8Ba2, which are highlighted in (F).

(D) The two Cry8Ba2 variants (Ext = extended, Com = compact).

(E) Cry1Ca18 and Cry8Ba2 share one conserved intramolecular disulfide bond, linking DIV and DV. Electron density difference maps (Fo – Fc, 0.6 e/Å^3) visualize the disulfide bond.

(F) Notable structural differences include (F1) the resolution of a helix in DV (aa 864–859) in Cry8Ba2, which is not resolved in Cry1Ca18, (F2) the resolution of a flexible loop in DVII of Cry8Ba2 (aa 1107–1128), which is not present in Cry1Ca18 and (F3) the presence of a hairpin in DV of Cry8Ba2 (aa 773–780), which is not present in Cry1Ca18.

(Figure 1E). One key difference between the two Cry variants was identified in Cry8Ba2 DV, where residues 773-REIDTETY-780 form a short β-hairpin, with no equivalent seen in the Cry1Ca18 or the Cry1Ac-Δ14C structure (Figure 1F3). The Cry8Ba structures have also revealed prodomain sections that are missing in the electron density maps of both Cry1Ca18, and the

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

CellPress
OPEN ACCESS

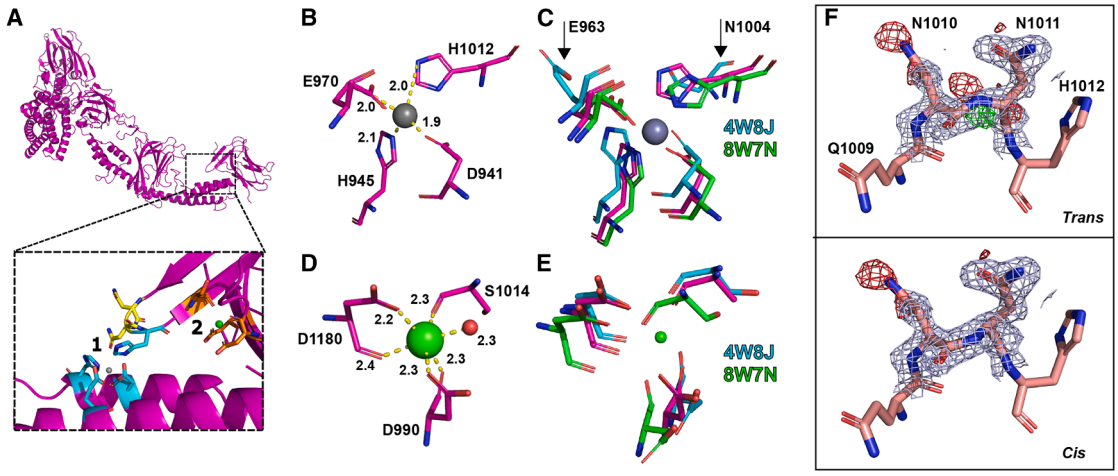


Figure 2. Putative metal binding sites conserved in Cry1 structures

(A) Cry1Ca18 contains two putative metal binding sites, present between domain VI and domain VII (site 1, cyan) and within DVII (site 2, orange).

(B) Site one represents a putative zinc binding site, coordinated by four surrounding amino acids (D941, H945, E970, and H1012). Zinc is represented by a gray sphere, with the coordination distances (Å) to the surrounding amino acids shown.

(C) Comparison of site one with other Cry1 structures. The putative zinc site is conserved in the previously elucidated Cry1Aa (PDB: 8W7N, green) but not in Cry1Ac-Δ14C (PDB: 4W8J, cyan) where the E963 side chain is pointed away from the site.

(D) Site two represents a putative calcium binding site (calcium represented by a green sphere), coordinated by three surrounding amino acids (D990, S1014, and D1180) and one water molecule (red sphere).

(E) Comparison of site two with other Cry1 structures. The putative calcium site is conserved in the previously published full-length Cry1Aa and Cry1Ac structures.

(F) Directly preceding H1012 of the putative zinc site is a non-proline cis peptide bond (A, yellow). Electron density difference maps (Fo – Fc, 0.6 e/Å^3) visualize the improved fit of the cis conformation compared to the trans conformation (red = negative electron density, green = positive electron density).

previously published Cry1Ac-Δ14C and Cry1Ac. In one of these regions, residues 864–859 in DV form a single helix that interrupts an extended loop region (Figure 1F1). In a second region, weak density was present between Cry8Ba2 residues 1107–1128 in DVII but a model could be constructed nonetheless as a loop that projects out perpendicular to the structure (Figure 1F2).

Cry1Ca18 has two putative metal ion binding sites whereas Cry8Ba2 has none

Further analysis of the electron density map indicated two putative metal ion binding sites in Cry1Ca18, as indicated by regions of positive electron density between coordinated side chains (Figure 2A). No equivalent putative metal ion binding sites were observed in the Cry8Ba2 structures, and Cry8Ba2 does not share the conserved putative binding residues present in Cry1Ac, Cry1Aa, and Cry1Ca. The first site is between DVI and DVII, occurring between D941, H945, E970, and H1012, where protein side chain atoms are within 2.0 Å from the identified metal site (Figure 2B). An equivalent pocket was previously identified, and noted by the depositors, in the full-length Cry1Aa structure (PDB: 8W7N; D932, H936, E961, and H1003). However, in the mutant Cry1Ac-Δ14C model (PDB: 4W8J), while the conserved residues are present (D934, H938, E963, and N1004), E963 adopts a conformation taking the side chain away from potential binding distance to the putative metal site (Figure 2C). Also, N1004 is located on a beta-hairpin folded to take the side chain further away from the would-be metal site, rendering it unoccupied.

The second Cry1Ca18 putative ion binding site was observed within DVII, coordinated by D990, S1014, and D1180 (Figure 2D). Five protein atoms and a water make up the coordination set around the site, with contact distances ranging from 2.2 Å to 2.4 Å. An equivalent putative binding site was noted by the depositing authors of Cry1Ac-Δ14C structure (PDB: 4W8J; D983, S1007, and D1169; Figure 2E). The binding site was not noted by the Cry1Aa (PDB: 8W7N) depositors, although the structure shares the conserved residues (Cry1Aa–D981, S1005, and D1171) and there is unoccupied positive density in the electron density maps, indicating the presence of an ion and conservation of this metal site between Cry1Ac, Cry1Aa, and Cry1Ca18. To hypothesize what ions are bound in these locations, we used "CheckMyMetal" (CMM)—a metal-binding site validation server that implements a systematic inspection of the entire binding environment to rank metals in order of confidence.[28] CMM indicated the most probable ions associated with these two sites are zinc and calcium, respectively (Tables S2 and S3). Absorption spectra showed the sample does contain Zn and Ca (Figure S3), supporting the interpretation of the structure determination. Other elements observed in the spectra could still be present in the sample bulk, but are not bound to the structure in an ordered or discernable manner.

Initial structural refinement had regions of poor density fit around residues proceeding H1012 of the putative zinc binding site. Refinement with a *cis*-peptide bond between residues N1010 and N1011 reduced surrounding negative electron density (Figures 2A and 2F). A non-proline cis peptide is a rare occurrence in protein structure (approx. 0.03%[29]). Our high-quality dataset and structural refinement gives us confidence this is a genuine occurrence. Non-proline cis peptides are most often observed in carbohydrate binding, where evidence indicates they can be stabilized by metal coordination sites.[30] The metal

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

binding sites, and the cis peptide may be a strategy used by the *Bacillus* to stabilize the toxin fold for a biological function that is, at present, unknown. It is notable that a pocket of similar dimensions to the putative Zn metal binding site of Cry1Ca18 is present in the two Cry8Ba2 structures but the no *cis*-peptide bonds are present and the distinct residues lining the Cry8Aa2 pocket provide no sidechains in orientations able to chelate a metal ion.

In vivo crystal packing shows different stabilization strategies between Cry1Ca18 and Cry8Ba2

It is suggested that forming a natural crystal ensures both protein stability in the open environment and a controlled release of virulence factors in the target insect. Structure elucidation from natural crystals allows us an insight into the nature of *in vivo* packing, and how this may facilitate dissolution under the specific conditions of the insect midgut lumen.[9] To provide some context to crystal packing, we ascertained the pH solubility of both proteins across a range of pH points (pH 3, 5, 7, 8, 8.5, 9, 9.5, 10, and 11) and observed Cry1Ca18 to solubilize at pH 11, and Cry8Ba2 to solubilize at pH 9, 9.5, 10, and 11 (Figure S4).

Cry1Ca18 *in vivo* crystal packing utilizes intermolecular disulfide bonds

The crystal lattice is held together via a network of disulfide bridges, salt bridges, and hydrogen bonds (Table S4). Each monomer has nine interfaces with other monomers totalling 8,648.8 Å^2 with a solvation free energy gain of −62.7 kcal/mol—as predicted by PISA.[31] The largest interface with the highest affinity, and largest number of hydrogen bonds (48) and salt bridges (18), is that between a dimer (3,995 Å^2)—where the crystallization domain of one monomer cups the toxin core of the other monomer, supporting the hypothesis that the C-terminal regions are involved in crystallization and stability (Figure 3A). Cry1Ca18 contains 19 cysteine residues, 13 of which are visible in the structure (Figure 3B). In addition to the one intramolecular disulfide bond between C662 and C809 (between DIV and DV; Figures 1D and 3B), there are two intermolecular disulfide bonds between prodomain regions (Figure 3C), at interface 2 (C803 of one monomer with C1070 of another monomer, linking domains V and VII), and interface 3 (C1032 of one monomer with C829 of another monomer, linking domains VII and V). Of the seven remaining visible cysteine residues, C844, C997, and C1052 are accessible but have no obvious option for an interaction, C206, C430, and C668 are shielded from other cysteines by the surrounding monomer. C737 has a significant amount of additional density extending from the sulfur atom, indicative of a disulfide linkage. Given its location, it is plausible it could interact with C1136 in the unresolved disordered region of DVI. Indeed, Cry1Ca18 AlphaFold modeling positions C1136 in an unstructured region with the sulfur atom 2.6 Å away from the sulfur atom of C737 (Figure S5).

The crystal lattice is formed by successive molecules lined up head-to-tail rectilinearly, parallel to the 4-fold screw (4_1) axis in two strands. Each strand is formed by the ''crystallization domains,'' DIV to DVII, of each molecule lining up along the 4-fold axis, while the toxin core is offset away from the 4_1 axes. The two strands run in opposite directions by 2-fold screw symmetry (2_1) normal to the 4-fold screw axis (Figure 3D). The core points sideways from the main strand and interdigitates with the neighboring strand, making the lattice mechanically robust.

Controlled dissolution of crystals under the conditions of the target insect midgut is hypothesized to ensure a potent and targeted toxin delivery. To understand the process of crystal dissolution at a biologically relevant pH more clearly, we looked at the crystal structure under alkaline conditions—representative of midgut conditions in some lepidopteran and dipteran target insects (typically pH 8–12, but can vary[32]). A pH solubilization range indicated Cry1Ca18 solubilized at pH 11 (Figure S4). We tried mixing Cry1Ca18 crystals with a pH 10 or pH 11 buffer, but solubilization occurred too quickly and blocked the injection nozzle. Therefore, native crystals of Cry1Ca18 were mixed with a pH 9 buffer in the path to the beamline for approximately 40 s prior to diffraction data collection. Comparison of the overall structures at pH 7 (PDB: 9H99) and pH 9 (PDB: 9QXQ) shows an increased unit cell volume (approx. 2%), with a highly similar structure (all-atom RMSD of 1.031 Å), with the major difference being a reduction in total crystal interface area of 295 Å^2 and a loss of 10 hydrogen bonds (Table S5). There was no alteration in disulfide bonds. These changes may reflect some of the earliest structural changes observed on increasing pH pre-solubilization, and natural mutation of these residues might considerably affect crystal integrity and, therefore, potency in the insect gut.

In vivo crystal packing in Cry8Ba2 is mediated by non-covalent interactions

Both Cry8Ba2 crystal forms share a generally conserved conformation with Cry1Ca18 (Figure 4A). PISA interface analysis identified 10 interfaces in the Cry8Ba2-Com crystal (Table S6) and 8 interfaces in the Cry8Ba2-Ext crystal (Table S7). Total interface areas are 8,288.1 Å^2 in Cry8Ba2-Com and 7,584.3 Å^2 in Cry8Ba2-Ext, with the biggest interface—as with Cry1Ca18—present at the head-to-tail dimer interface (Figure 4B). From the indexing rates (Com:30123 and Ext:33552), we see the two populations are present in an almost 50/50 split. The biological relevance of two distinct populations of Cry8Ba2 crystals will require further investigation but given the weaker dimer interface and lower number of crystal contacts, we hypothesize that Cry8Ba2-Ext may represent the more readily solubilized and activated crystal form out of the two.

As expected, the Cry8Ba2 crystal lattice is held together by a network of salt bridges and hydrogen bonds. Unlike Cry1Ca18 and Cry1Aa (PDB: 8W7N), which have 19 and 17 cysteine residues, respectively, Cry8Ba2 only contains four cysteine residues. Aside from the conserved disulfide bond (Figure 1D), unpaired C246 in DI is structurally conserved with Cry1Ca18 C206, while unpaired C560 in DIII is not. Furthermore, there are no intermolecular disulfide bonds within the Cry8Ba2 crystal lattice (Figure 4C). In addition to the lack of intermolecular disulfide bridges, compared to Cry1Ca18 there are also significantly fewer salt bridges (21 in Cry8-Ext, 34 in Cry8-Com, and 44 in Cry1Ca18), and a substantial reduction in the total interface protein affinity (Δ^iG in Cry8Ba −19 kcal/mol and in Cry1Ca18 −62 kcal/mol).

As shown in Figure 1F2, Cry8Ba2 model building revealed an extended loop (residues 1105–1129) in DVII. When visualized in context of the crystal lattice, this region appears to act as a cross-strand peg—potentially stabilizing the packing of neighboring molecules (Figure 4D). The weak electron density around this peg may suggest a degree of flexibility. The loop would be easier to insert into the neighboring strand if it is flexible, while

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

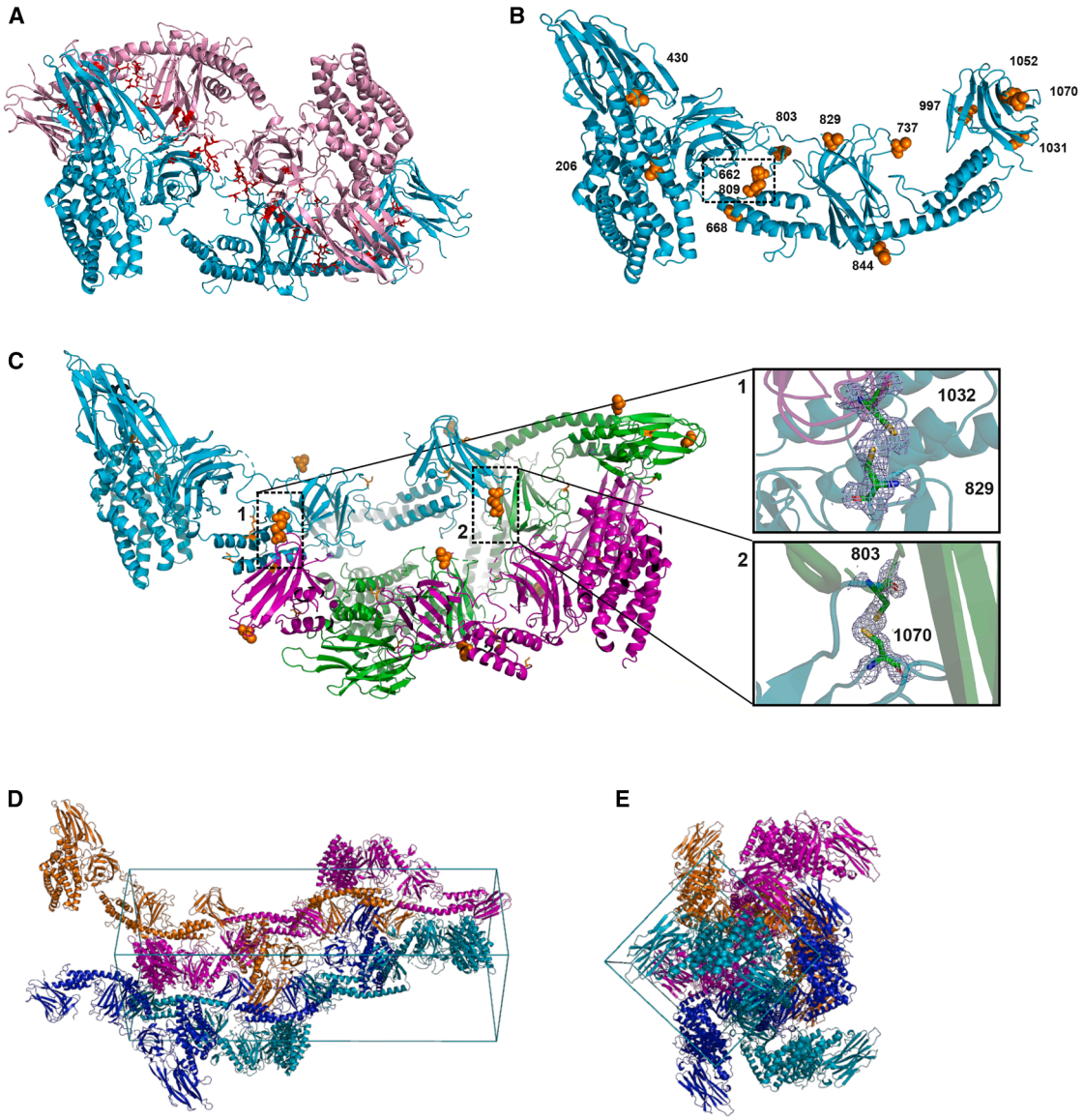


Figure 3. Cry1Ca18 crystal packing is stabilized by intramolecular and intermolecular disulfide bonds

(A) Cry1Ca18 packs as a head-to-tail dimer, stabilized by a network of hydrogen bonds and salt bridges (red).

(B) Thirteen cysteine residues are visible in the electron density map (orange spheres), boxed region shows intramolecular disulfide bond.

(C) Two of the 13 cysteine residues are occupied in intermolecular disulfide bonds between DIV and DVII (C1) and DVII and DV and (C2), electron density difference maps (Fo – Fc, 0.5 e/Å^3).

(D) 2D-projection looking down the diagonal between the a/b axis. The lattice is formed by successive molecules lined up head-to-tail rectilinearly. Each strand is formed by the prodomains of each molecule lining up along the 4-fold axis, while the toxin core is offset away from the 4_1 axes. The two strands run in opposite directions to one another and are visualized as colored monomers: strand 1 (alternating orange and pink) and strand 2 (alternating cyan and dark blue).

(E) 2D-projection looking down the *c*-axis.

a rigid loop may cause the docking to fail. Shape complementarity (sc) analysis suggested a reasonably complementary shape, with a sc value of 0.41 in Cry8Ba2-Com and 0.49 in Cry8Ba2-Ext (well packed protein-protein interfaces range approx. = 0.55–0.75). PISA analysis indicated multiple hydrogen bonds and salt bridges between the loop and neighboring monomers in both the Cry8Ba2-Com and Cry8Ba2-Ext variants (Table S8). Collectively, this suggests that this flexible interface may aid in Cry8Ba2 crystal packing and stability.

As indicated above in Figure 4E, structural characterization of the Cry8Ba2 crystals show pH sensitive intermolecular interactions, including salt bridges and hydrogen bonds. Several of these interactions include titratable residues, such as histidine, lysine, aspartic acid, glutamic acid, and tyrosine. To compare with Cry1Ca18, we attempted to look at the effect of an alkaline pH 9 buffer on the crystal lattice.[33] However, under the same mixing conditions, we were not able to resolve the structure. This is consistent with Cry8Ba2 solubilizing at pH $\geq$ 9

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

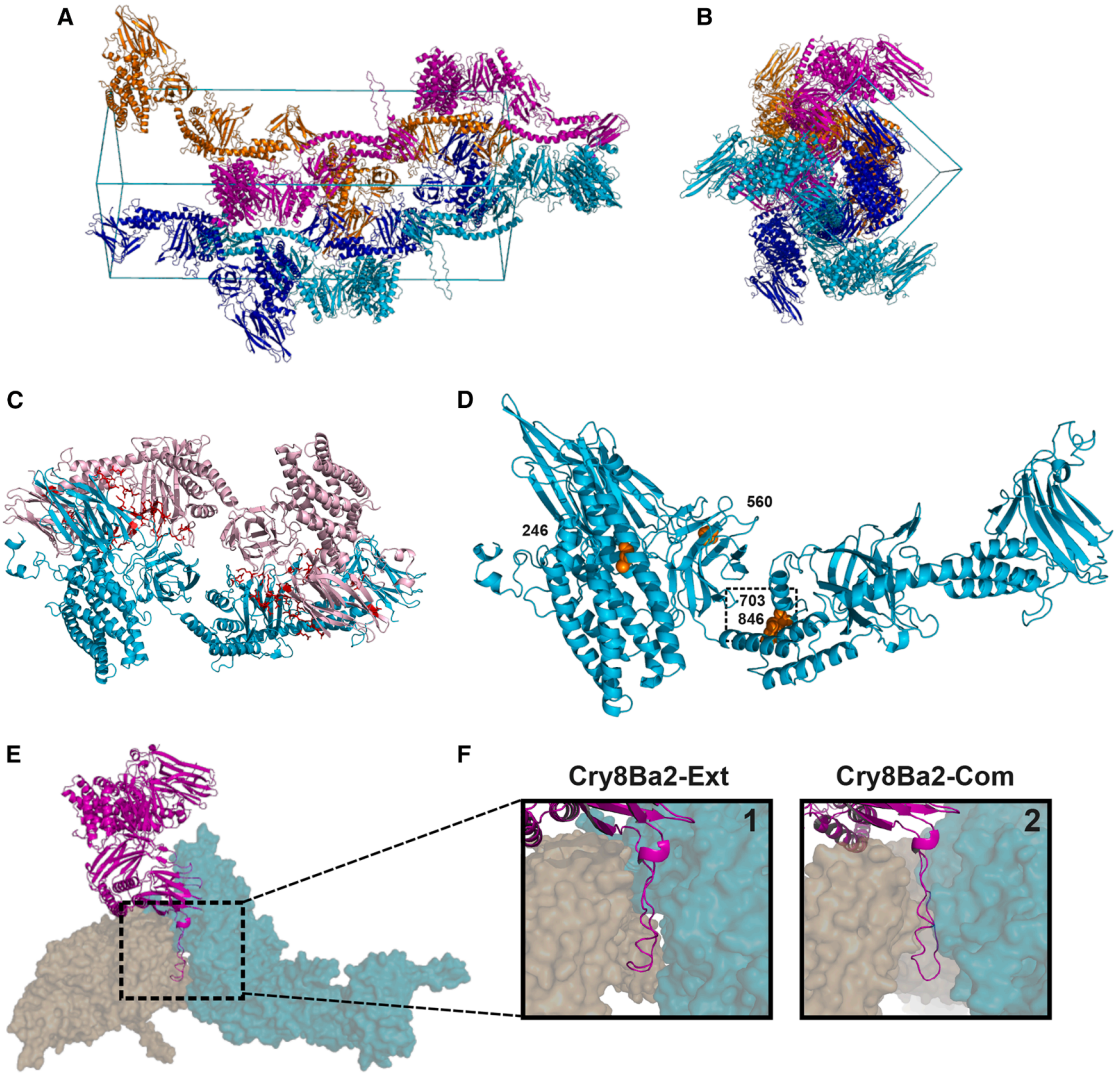


Figure 4. Cry8Ba2 crystal packing is not stabilized through disulfide bonds

The Cry8Ba2-Ext variant is depicted unless stated otherwise.

(A) 2D-projection looking down the diagonal between the *a*/*b* axes. Cry8Ba2 shares a conserved packing formation with Cry1Ca18 (see Figures 3D and 3E).

(B) 2D-projection looking down the *c*-axis.

(C) Cry8Ba2 packs as a head-to-tail dimer, stabilized by a network of hydrogen bonds and salt bridges (red).

(D) Four cysteine residues are visible in the electron density map (orange spheres).

(E) An interface between three monomers in the Cry8Ba2-Ext structure shows the DVII putative loop (residues 1105–1129) stabilization of crystal packing.

(F) A closeup of the putative stabilization loop interface shown in (E), in both the Cry8Ba2-Ext interface (1), and Cry8Ba2-Com (2) interface.

(Figure S4), reduced total interface protein affinity, reduced protein affinity ($\Delta^{i}G$), fewer disulfide bonds, and fewer salt bridges compared to Cry1Ca18. A summary of the similarities and differences of Cry1Ca18 and Cry8Ba2 is presented in Table S9.

Cysteine residues show limited conservation between Cry families

Given the distinct methods that may stabilize the natural crystals in this study: the structural peg feature in the Cry8Ba2 lattice or the presence of intermolecular disulfide bonds in the Cry1Ca18 lattice, we analyzed the distribution of cysteine residues across the Cry family (Data S1). Cysteine residues are found in almost all Cry proteins, except for Cry70 proteins that contain none. For other Cry variants, numbers can range up to 32 cysteines seen in Cry32Ka1. Predictably, long-form Cry proteins (those of ~120–140 kDa containing a "crystallization domain" such as Cry1Ca18) have more cysteines than short-form proteins such as Cry2 and Cry3—the exception to this are some proteins in the Cry8 group, including Cry8Ba2.

A common feature of cysteine localization is pairs of residues in close succession, e.g., CxC, CxxC, and CxxxC. Such motifs are reminiscent of previously proposed metal-binding sites in metallochaperones[34] and in the case of Cry proteins, these may be involved in the coordination of metal ions to aid the

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

packing of the crystal superstructure during assembly or may have a role in the function of these regions in target insects.

Using AlphaFold predicted models, we looked for patterns in cysteine residue location and found that, within Cry subfamilies, cysteine residues are typically well structurally conserved. A subset of these residues is shown in Table S10, which lists the degree of conservation across long-form (>1,100 residues) protoxins and indicates known/predicted linkages in Cry1Aa (PDB: 8W7N), Cry1Ca18, and Cry8Aa2. The highest conservation is seen for the residue equivalent to Cry1Ca18 C809 (92%), which is most often paired to form a disulfide bond with C622 (85% conserved)—as we demonstrate in our Cry1Ca18 and Cry8Ba2 structures. Cry subfamily members lacking these residues (and, therefore, the disulfide bridge) include Cry5A, Cry5B, Cry12A, Cry14A, Cry21A, Cry48A, and Cry67A. We note that most of these families contain nematocidal proteins (while Cry48A and Cry67A proteins have not been assayed against nematode targets) although whether this observation is significant, remains to be tested.

When considering the position of cysteine residues in AlphaFold models of Cry structures, cysteines within the "crystallization domain" are predominantly surface accessible/exposed. This positioning could allow for intermolecular disulfides to form, acting as a stabilizing feature in the crystal. Modeling and sequence alignments highlight several regions where cysteine occurrence tends to be higher; (region 1) cysteine residues are frequently found in the N-terminal ~27 residue pro-region (two in Cry1Ca18), which is disordered in our structures. (region 2) A subset of Cry proteins have sequences extending beyond the C-termini seen in Cry1C and Cry8B, which are predicted to fold to form an extra domain (e.g., residues M1274 to V1522 in Cry32Ka1), and these extensions may be cysteine-rich (8 in the above region of Cry31Ka1) (region 3). The loop extending from DV that may contain the well-conserved residue equivalent to Cry1Ca18 C809 (see above), is variable in length. This loop often appears to be a cysteine-rich region (five in Cry1Ca18, V785–H835 loop; one in Cry8Ba2, V825–E868 loop; ten in Cry32Ka1, V877–V980 loop). Previous work has demonstrated that mutants in Cry7Ba1 residues C697, C834, C840, and C854, in the equivalent region, all result in a reduction in the pH necessary for crystal solubilization.[15] C697 and C840 are predicted to form an intramolecular disulfide bond equivalent to the C662-C809 bond in Cry1Ca18. (region 4) The region in DVII that contributes the peg feature in Cry8Ba2 (F1099–I1133) with no cysteine residues, in contrast, displays a number of cysteines in Cry1Ca18 (region from F1067–V1153), including C1136, which forms an intramolecular disulfide with C737, and C1070, which forms an intermolecular disulfide with C803 in a neighboring monomer. The latter two regions would appear to be important across the Cry family for stabilizing crystals while the roles of the cysteine in the other regions has not been investigated.

Some Cry proteins, such as the Cry10, Cry30, and Cry41 subfamilies, are expressed as split-form proteins, with the 3-domain toxin core encoded by one coding sequence (CDS) followed by a second CDS encoding a protein with high sequence identity to domains IV–VII. Occurrence of cysteine residues in these cases ranges from 10 residues in the Cry10Aa3 CDSs (2 cysteines in toxin and 8 cysteines in the DIV–DVII CDS), 16 in Cry30Ba1 (4 cysteines in toxin and 12 cysteines in the DIV–DVII CDS), to 21 residues in Cry41Aa1 (7 cysteines in toxin and 14 cysteines in the DIV–DVII CDS). This indicates a range in the number of cysteines similar to many long-form Cry proteins. It has been shown that DIV–DVII CDS proteins are present in crystals of Cry41Aa1[35] although the nature of packing in these crystals is currently not known.

DISCUSSION

Cry proteins are environmentally friendly microbial insecticides that have been effectively utilized in agriculture for over 50 years. Understanding these proteins' structure and mechanisms of action is key for the continued identification and development of Bt biopesticides. Here, we publish structures of Cry1C and Cry8B variants. These are non-mutant structures of long-form Cry proteins in their natural nanocrystalline form, as grown *in vivo* by Bt. Cry1Ca and Cry8Ba family members demonstrate activity against lepidopteran and dipteran (Cry1Ca), and coleopteran (Cry8Ba) pests. Although as many specificity experiments use strains which contain a mixture of Cry proteins, the only target specifically reported for Cry1Ca18 is *Aedes aegypti*[36] (although Cry1Ca15, with which it shares 100% identity, is active against *Spodoptera exigua*[37]), and there is no reported specificity data for Cry8Ba2. As proteins with potential for insect control, the work presented here provides structural insights for future protein engineering for enhancing insecticidal efficacy of these and other Cry proteins.

The intrinsic properties of Bt crystals that assist with toxicity, namely the stability in aqueous environments followed by selective solubility in the insect midgut, are controlled by the protein structure and crystallization properties. In long-form Cry proteins, the C-terminal "crystallization domain" represents approximately half the size of the protein.[38] They are so-called as previous work has indicated that a C-terminal domain was essential for crystallization of the associated Cry protein in Bt.[39] Conserved space group and unit cell parameters between the derived Cry1Ca18 and Cry8Ba2 structures presented here (along with previously published Cry1Ac-Δ14C, Cry1Ac, and Bt crystals studied using powder X-ray diffraction) hints at an evolutionary conservation of this packing to stabilize crystallization in proteins with a "crystallization domain."[19] It is interesting to note that the structures of natural crystals of short-form Cry proteins adopt $C222_1$ (Cry3Aa),[40] I222 (Cry11Aa), or $P2_12_12$ (Cry11Ba).[41] Furthermore, previous analysis of the toxin portions (DI–DIII) of Cry proteins has indicated that although sequence identity between Bt Cry toxins is quite low at primary rank (e.g., 36% between Cry1Ac7 and Cry8Ea1), the structural similarity is reasonably high (e.g., 2.31 Å all-atom RMSD between Cry1Aa7 PDB: 4ARX and Cry8Ea PDB: 3EB7).[42] It is possible that structural conservation between prodomains is a mechanism to enable the packaging of diverse Cry toxin cores into a single co-crystal to ensure synchronous co-exposure of an insect target to multiple Cry variants.

Specifically, cysteine rich amino acid sequences that are largely conserved across certain subfamilies (e.g., Cry1) have been proposed as elements that stabilize the crystal, via the formation of intermolecular disulfide cross-links, and facilitate solubilization in the alkaline midgut via disulfide bond dissociation.[43] Altering

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

solubility properties through mutation of cysteine bonds is a potential route for insecticidal toxin engineering, with a lack of solubility shown to contribute to a lack of toxicity.[15–17] For example, previous work has shown a single cysteine substitution in Cry7Ba1 allowed for crystal dissolution at a lower pH and conferred toxicity against *Plutella xylostella* larvae.[15] However, making a protein more soluble does not always confer greater toxicity, demonstrated by previous work showing the activation of Cry8Ha1 disables its toxicity against *Holotrichia oblita* adults, reduces toxicity against *Holotrichia parallela* adults, but gains toxicity against *Colaphellus bowringi* larvae.[18]

We have considered the possibility that the number of disulfide bonds may be related to the pH of the host gut and affect target specificity, since such bonds are known to be more labile under alkaline conditions and many nematodes and beetles have acidic guts, but the overall correlation is not strong. Indeed, Cry1Ca18 was only fully solubilized *in vitro* at pH 11: in target insect guts, pH may act in concert with additional factors, e.g., proteolysis, surfactants, reducing conditions, to facilitate crystal solubilization. Although six cysteine residues were not observable in the Cry1Ca18 map, we report two intermolecular disulfide bonds. These intermolecular links are not present between the dimer, suggesting that the dissociation of disulfide bonds in the midgut would first release the protein in a dimeric form, consistent with the observation that intermolecular disulfide bonds are less stable than intramolecular ones.[44] This may also provide an explanation for the fact that at solubilizing pH, Cry1Ca18 clogged injection nozzles (possibly through dissociation of the lattice with disulfide cross-links still intact), whereas Cry8Ba2 caused no blockage, consistent with the lack of intermolecular disulfide bonds. Most cysteine residues are positioned within flexible loop regions, making it plausible that more than one version of the disulfide network may exist, dependent on the crystal environment. Indeed, analysis of the Cry family showed cysteine residues are common in disordered regions—hinting at their requirement for contact points in crystal packing. In contrast to the idea that disulfide bridges are a necessity for crystal formation and stability, previous work demonstrated that Cry1Ac-Δ14C can crystalize in Bt, and these crystals are still toxic to target pests. However, the same study also made a Cry1Ac-Δ16C mutant, including mutation of the two most N-terminal cysteines (amino acids 10 and 15), which are present in the disordered region and not visible in the structure. This resulted in the crystals dissolving at the end of sporulation—indicating either/both of these additional two residues are critical to crystal integrity.[21] Furthermore, the long-term stability of the Cry1Ac-Δ14C crystals in the environment was not tested. The fact that the absence of intermolecular disulfide bridges can be compatible with Cry crystal formation, is further demonstrated by Cry8Ba2, which has four cysteines, of which, the only two present in the "crystallization domain" form the conserved intramolecular disulfide bond between DIV and DV. This suggests that other characteristics of the Cry8Ba2 "crystallization domain" perform similar roles in crystal packing. The presence of the structural "peg" in DVII offers a different solution to the packing stability conundrum.

Another notable feature of the Cry1Ca18 structure was the presence of two putative metal binding sites. The putative zinc site was conserved in the previously deduced full-length Cry1Aa (PDB: 8W7N) structure, and the putative calcium site was conserved in the Cry1Ac-Δ14C (PDB: 48WJ). Metal ion binding sites can perform a variety of functions, including protein folding and assembly, structural stability, and the regulation of unfolding and conformational changes.[45–47] The existence of shared metal ion binding sites in the Cry1 group members in the C-terminal "crystallization domain" hints at a conserved role in crystal packing, stability, or controlled conformation change upon dissolution. These putative conserved sites are not present in the Cry8Ba2 structure and may help explain differences in pH solubility and target selectivity.

Several studies have indicated that metal ions alter Cry protein production, for example, Mg^{2+} and Ca^{2+} has been shown to increase Cry4 and Cry11 production by Bt *israelensis* HD500,[48] and Cu^{2+} can increase crystal protein production of Cry1Ac in BtX023.[49] Many studies have shown metal ions also affect activity. For example, ion chelators reduce Cry41Aa-induced HepG2 cytotoxicity, and Cry1Aa, Cry1Ac, and Cry1Ea pore formation in insect BBMVs,[50,51] while increased Ca^{2+} enhances the toxicity of Cry1C in *Spodoptera frugiperda* cells and Cry31Aa1 in HeLa cells.[52,53] The mechanisms behind the role(s) of metal ions in Cry toxin activity may be diverse. For example, Cry41Aa resistance in HepG2 cells, in the presence of EGTA, appears to occur through the chelation of plasma membrane-associated cations preventing a stable interaction of Cry41Aa with the plasma membrane.[50] This alteration of cation concentrations could affect the structure of the plasma-membrane bound receptors, as indicated in previous studies showing EDTA addition causes a large conformational change in a cadherin Cry1 receptor,[54] and removal of Ca^{2+} mediates the proteolytic cleavage of the Cry1A BT-R1 cadherin receptor.[55]

In close proximity to both Cry1Ca18 putative metal sites, and next to residue H1012, which forms part of the putative zinc binding site, we show the presence of a non-proline cis peptide. Only a small proportion of peptide bonds occurs in the energetically less favored cis confirmation (approx. 5.2%), and even fewer between two non-proline residues (approx. 0.03%).[29,56] Although they are found more frequently in high resolution structures, suggesting they may be more abundant but undetected at lower resolution.[30] Non-proline cis bonds are found more frequently near ligand/binding sites and dimerization interfaces.[57–59] The N1010-cis-N1011 peptide bond is located in DVII, a lectin-like domain with structural homology to carbohydrate binding proteins.[21] Non-proline cis bonds have been noted as more frequent in carbohydrate binding proteins,[30] such as concanavalin A—where binding activity can be turned on/off through subsequent Zn^{2+} and Ca^{2+} binding inducing isomerization of an Ala-cis-Ala bond.[60] Legume lectins often contain a conserved *cis*-peptide bond where the formation is driven by divalent cation binding (often Ca^{2+} and Mn^{2+}), and functions to dispose side chain atoms for ligand recognition.[61–63] However, the biological function here may be other than carbohydrate recognition. As with the putative metal binding sites, the cis peptide bond is not present in either of the Cry8Ba2 conformations, highlighting differences between the structures that may affect *in vivo* activity.

Proteolytic activation of long-form Cry toxins involves the removal of the "crystallization domain" and a short region at the N-terminus of the proteins. The N-terminal regions of both Cry1Ca18 and Cry8Ba2 are disordered and not visible in the

structures; mirroring observations for other Cry protoxins.[40,41] Proteolytic processing of Cry1Ca removes the first 27 amino acids of the N-terminus (along with the C-terminal domains VI–VII).[64] Removal of the N-terminal peptide from Cry proteins is believed to be required to activate the toxin, and this has been shown using Cry41Aa in cell line models.[65] Deletion of the region encoding these 27 residues in Cry1Ca, results in no detectable expression in *E. coli*[64] and their deletion, along with the region encoding domains VI–VII, appeared to result in toxicity to the host cell, although the protein appears to have folded correctly.[26] In the Cry2Aa structure (1I5P), the N-terminal residues are visible and appear to interact with parts of DII and adjacent regions of DIII, and it was proposed that the N-terminus may mask an epitope in this part of the protein that is essential for the binding to its receptor.[66] The lack of defined density for the N-terminal section in our structures suggests that, within these natural crystals, there is no defined interaction equivalent to that in Cry2Aa with DII and DIII.

While a significant proportion of the literature suggests the C-terminus is dispensable in terms of toxicity, in addition to altering the protein solubility, several studies demonstrate the C-terminus can also play a functional role in receptor binding. The C-terminal region of Cry1Ab can provide additional binding sites for alkaline phosphatase and aminopeptidase receptors and provides higher binding affinity to the target midgut.[14] Furthermore, intact Cry1Ac protoxin can be cytotoxic to the *Choristoneura fumiferana* midgut cell line (CF-203)[11] and provides increased binding affinity and increased potency in several resistant strains of *Lepidoptera*.[12,13,67] Interestingly, most, if not all, Cry proteins currently used commercially in bioengineered plants, utilize only the core region (DI–DIII). With an increasing amount of research indicating that the extended C-domain can affect toxicity, the structures elucidated here will be key to future studies focused on understanding the structure-function relationship of the protoxin and may be beneficial in protein engineering to develop more effective insecticides and manage resistance evolution.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Prof. Colin Berry (berry@cardiff.ac.uk).

Materials availability

There are restrictions to the availability of strains due to commercial potential (Na210) because of the lack of an external centralized repository for its distribution and our need to maintain the stock. The sharing of strains will be subject to suitable materials transfer agreements and reasonable compensation by the requestor for processing and shipping.

Data and code availability

Datasets for the EuXFEL structures presented here are deposited in the Protein DataBank under accession numbers 9H9A, 9H9B, 9QXQ, and 9H99 and are publicly available as of the date of publication. The data measured at EuXFEL can be made available upon a reasonable request. The sequences for Cry1Ca18 and Cry8Aa2 have been deposited with the accession numbers GenBank: PQ691233 and GenBank: MZ355710. This paper does not report any original code, any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC, grant reference BB/S002774/1), a BBSRC-funded SouthWest Biosciences Doctoral Training Partnership (training grant reference BB/M009122/1) and an EPSRC Doctoral Training Program (Physics of Life grant, EP/W524682/1). We acknowledge the support of the Supercomputing Wales project, which *has been in part funded* by the European Regional Development Fund via the Welsh Government. The authors would like to thank Bio-insectis SL for the provision of strain Na210 for use in this study. P.L.X. thanks Joachim Herz Stiftung for a fellowship. We acknowledge European XFEL in Schenefeld, Germany, for provision of X-ray free-electron laser beamtime at the Scientific Instrument SPB/SFX (Single Particles, Clusters, and Biomolecules and Serial Femtosecond Crystallography) under proposal numbers 3244 and 2819 and would like to thank the staff for their assistance. This research was supported in part through the Maxwell computational resources operated at Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany. The authors would like to thank Diamond Light Source beamtime (proposal mx-36446) and the staff of beamline i23 for assistance with collecting absorption peaks.

AUTHOR CONTRIBUTIONS

Experimental design, C.B., H.N.C., N.C., and D.O.; sample production, H.L.B., L.J.W., E.A.H., A.B.C., N.B.-S., N.C., M.V., P.C., E.L.E., and C.B.; protein analyses were carried out by H.L.B., E.A.H., and A.B.C.; experimental lead at European XFEL, D.O.; sample delivery at European XFEL, M.K., J.V., G.P., S.A., J.K., and J.B.; European XFEL data collection was executed by, A.M., A.R., K.D., K.L., D.O., R.J.B., R.d.W., C.K., H.K., G.M., M.S., A.P.M., K.L.P., and R.S.; experimental settings at European XFEL were coordinated by R.J.B. and A.P.M.; data processing was carried out by L.G., O.Y., and M.G.; TEM imaging was carried out by P.L.X. and R.S.; data were analyzed by D.O., P.J.R., C.B., H.L.B., A.B.C., and L.J.W. The manuscript was written by H.L.B., L.J.W., C.B., and D.O. with R.J.B., O.Y., M.G., A.B.C., N.C., and input from all the authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Purification and sample preparation of Cry8Ba2 and Cry1Ca18 protein crystals
  - Transmission electron microscopy
  - XFEL data collection
  - Structure determination
  - X-ray fluorescence spectra of Cry1Ca18
  - Sequence and model analysis of Cry protein cysteine content
  - Cry protein pH digest analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2026.01.014.

Received: May 16, 2025
Revised: November 25, 2025
Accepted: January 28, 2026

REFERENCES

1. Xu, C., Wang, B.-C., Yu, Z., and Sun, M. (2014). Structural insights into Bacillus thuringiensis Cry, Cyt and parasporin toxins. Toxins *6*, 2732–2770.
2. Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). Bacillus thuringiensis toxins: an overview of their biocidal activity. Toxins *6*, 3296–3325.
3. Berry, C., Valby, V., Mishra, R., Bonning, B.C., Palma, L., and Crickmore, N. (2025). Specificity database for bacterial pesticidal proteins against invertebrate targets. J. Invertebr. Pathol. *211*, 108319. https://doi.org/10.1016/j.jip.2025.108319.
4. Gassmann, A.J., and Reisig, D.D. (2023). Management of insect pests with Bt crops in the United States. Annu. Rev. Entomol. *68*, 31–49.
5. Heckel, D.G. (2020). How do toxins from Bacillus thuringiensis kill insects? An evolutionary perspective. Arch. Insect Biochem. Physiol. *104*, e21673. https://doi.org/10.1002/arch.21673.
6. Crickmore, N., Berry, C., Panneerselvam, S., Mishra, R., Connor, T.R., and Bonning, B.C. (2021). A structure-based nomenclature for Bacillus thuringiensis and other bacteria-derived pesticidal proteins. J. Invertebr. Pathol. *186*, 107438. https://doi.org/10.1016/j.jip.2020.107438.
7. Bravo, A., Pacheco, S., Gómez, I., and Soberón, M. (2023). Mode of action of Bacillus thuringiensis Cry pesticidal proteins. Adv. Insect Physiol. *65*, 55–92. https://doi.org/10.1016/bs.aiip.2023.09.003.
8. Bravo, A., Gill, S.S., and Soberón, M. (2007). Mode of action of Bacillus thuringiensis Cry and Cyt toxins and their potential for insect control. Toxicon *49*, 423–435.
9. Tetreau, G., Andreeva, E.A., Banneville, A.-S., De Zitter, E., and Colletier, J.-P. (2021). How does Bacillus thuringiensis crystallize such a large diversity of toxins? Toxins *13*, 443.
10. Adalat, R., Saleem, F., Crickmore, N., Naz, S., and Shakoori, A.R. (2017). In vivo crystallization of three-domain Cry toxins. Toxins *9*, 80.
11. Li, X., Zhao, F., Qiu, X., Ren, X., Mo, X., Ding, X., Xia, L., and Sun, Y. (2018). The full-length Cry1Ac protoxin without proteolytic activation exhibits toxicity against insect cell line CF-203. J. Invertebr. Pathol. *152*, 25–29.
12. Anilkumar, K.J., Rodrigo-Simón, A., Ferré, J., Pusztai-Carey, M., Sivasupramaniam, S., and Moar, W.J. (2008). Production and characterization of Bacillus thuringiensis Cry1Ac-resistant cotton bollworm Helicoverpa zea (Boddie). Appl. Environ. Microbiol. *74*, 462–469.
13. Siqueira, H.A.A., Nickerson, K.W., Moellenbeck, D., and Siegfried, B.D. (2004). Activity of gut proteinases from Cry1Ab-selected colonies of the European corn borer, Ostrinia nubilalis (Lepidoptera: Crambidae). Pest Manag. Sci. *60*, 1189–1196.
14. Peña-Cardeña, A., Grande, R., Sánchez, J., Tabashnik, B.E., Bravo, A., Soberón, M., and Gómez, I. (2018). The C-terminal protoxin region of Bacillus thuringiensis Cry1Ab toxin has a functional role in binding to GPI-anchored receptors in the insect midgut. J. Biol. Chem. *293*, 20263–20272.
15. Peng, D., Wang, F., Li, N., Zhang, Z., Song, R., Zhu, Z., Ruan, L., and Sun, M. (2011). Single cysteine substitution in Bacillus thuringiensis Cry7Ba1 improves the crystal solubility and produces toxicity to Plutella xylostella larvae. Environ. Microbiol. *13*, 2820–2831. https://doi.org/10.1111/j.1462-2920.2011.02557.x.
16. Lambert, B., Höfte, H., Annys, K., Jansens, S., Soetaert, P., and Peferoen, M. (1992). Novel Bacillus thuringiensis insecticidal crystal protein with a silent activity against coleopteran larvae. Appl. Environ. Microbiol. *58*, 2536–2542. https://doi.org/10.1128/aem.58.8.2536-2542.1992.
17. Du, C., Martin, P.A., and Nickerson, K.W. (1994). Comparison of Disulfide Contents and Solubility at Alkaline pH of Insecticidal and Noninsecticidal *Bacillus thuringiensis* Protein Crystals. Appl. Environ Microbiol. *60*, 3847–3853. https://doi.org/10.1128/aem.60.10.3847-3853.1994.
18. Shu, C., Yan, G., Huang, S., Geng, Y., Soberón, M., Bravo, A., Geng, L., and Zhang, J. (2020). Characterization of Two Novel Bacillus thuringiensis Cry8 Toxins Reveal Differential Specificity of Protoxins or Activated Toxins against Chrysomeloidea Coleopteran Superfamily. Toxins *12*, 642.
19. Holmes, K.C., and Monro, R.E. (1965). Studies on the structure of parasporal inclusions from Bacillus thuringiensis. J. Mol. Biol. *14*, 572–581. https://doi.org/10.1016/S0022-2836(65)80205-4.
20. Lecadet, M.M., and Dedonder, R. (1967). Enzymatic hydrolysis of the crystals of Bacillus thuringiensis by the proteases of Pieris brassicae I. Preparation and fractionation of the lysates. J. Invertebr. Pathol. *9*, 310–321. https://doi.org/10.1016/0022-2011(67)90065-1.
21. Evdokimov, A.G., Moshiri, F., Sturman, E.J., Rydel, T.J., Zheng, M., Seale, J.W., and Franklin, S. (2014). Structure of the full-length insecticidal protein Cry1Ac reveals intriguing details of toxin packaging into in vivo formed crystals. Protein Sci. *23*, 1491–1497. https://doi.org/10.1002/pro.2536.
22. Tanaka, J., Abe, S., Hayakawa, T., Kojima, M., Yamashita, K., Hirata, K., and Ueno, T. (2023). Crystal structure of the in-cell Cry1Aa purified from Bacillus thuringiensis. Biochem. Biophys. Res. Commun. *685*, 149144.
23. Smedley, D.P., and Ellar, D.J. (1996). Mutagenesis of three surface-exposed loops of a Bacillus thuringiensis insecticidal toxin reveals residues important for toxicity, receptor recognition and possibly membrane insertion. Microbiology *142*, 1617–1624. https://doi.org/10.1099/13500872-142-7-1617.
24. Lee, M.K., Young, B.A., and Dean, D.H. (1995). Domain III exchanges of Bacillus thuringiensis CryIA toxins affect binding to different gypsy moth midgut receptors. Biochem. Biophys. Res. Commun. *216*, 306–312. https://doi.org/10.1006/bbrc.1995.2625.
25. Bosch, D., Schipper, B., van der Kleij, H., de Maagd, R.A., and Stiekema, W.J. (1994). Recombinant Bacillus thuringiensis Crystal Proteins with New Properties: Possibilities for Resistance Management. Biotechnology *12*, 915–918. https://doi.org/10.1038/nbt0994-915.
26. Kouskoura, T., Tickner, C., and Crickmore, N. (2001). Expression and crystallization of an N-terminally activated form of the Bacillus thuringiensis Cry1Ca toxin. Curr. Microbiol. *43*, 371–373. https://doi.org/10.1007/s002840010318.
27. Umina, P.A., Lord, A., Micic, S., and Edwards, O. (2017). Discovery and characterisation of field resistance to organophosphorus chemicals in a major mite pest, Halotydeus destructor. Pest Manag. Sci. *73*, 1719–1724. https://doi.org/10.1002/ps.4520.
28. Zheng, H., Chordia, M.D., Cooper, D.R., Chruszcz, M., Müller, P., Sheldrick, G.M., and Minor, W. (2014). Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. Nat. Protoc. *9*, 156–170.
29. Weiss, M.S., Jabs, A., and Hilgenfeld, R. (1998). Peptide bonds revisited. Nat. Struct. Biol. *5*, 676.
30. Jabs, A., Weiss, M.S., and Hilgenfeld, R. (1999). Non-proline cis peptide bonds in proteins. J. Mol. Biol. *286*, 291–304.
31. Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797. https://doi.org/10.1016/j.jmb.2007.05.022.
32. Dow, J.A. (1992). pH GRADIENTS IN LEPIDOPTERAN MIDGUT. J. Exp. Biol. *172*, 355–375. https://doi.org/10.1242/jeb.172.1.355.
33. Biggs, D.R., and McGregor, P.G. (1996). Gut pH and amylase and protease activity in larvae of the New Zealand grass grub (Costelytra zealandica; Coleoptera: Scarabaeidae) as a basis for selecting inhibitors. Insect Biochem. Mol. Biol. *26*, 69–75. https://doi.org/10.1016/0965-1748(95)00064-X.
34. Rosenzweig, A.C. (2002). Metallochaperones: Bind and Deliver. Chem. Biol. *9*, 673–677. https://doi.org/10.1016/S1074-5521(02)00156-4.
35. Krishnan, V., Domanska, B., Elhigazi, A., Afolabi, F., West, M.J., and Crickmore, N. (2017). The human cancer cell active toxin Cry41Aa from Bacillus thuringiensis acts like its insecticidal counterparts. Biochem. J. *474*, 1591–1602. https://doi.org/10.1042/bcj20170122.
36. Bryce-Sharron, N., Nasiri, M., Powell, T., West, M.J., and Crickmore, N. (2024). A Shared Receptor Suggests a Common Ancestry between an

Insecticidal Bacillus thuringiensis Cry Protein and an Anti-Cancer Parasporin. Biomolecules *14*, 795.

37. Baranek, J., Banaszak, M., Lorent, D., Kaznowski, A., and Konecka, E. (2021). Insecticidal activity of Bacillus thuringiensis Cry1, Cry2 and Vip3 toxin combinations in Spodoptera exigua control: highlights on synergism and data scoring. Entomol. Gen. *41*, 71–82.
38. Baum, J.A., and Malvar, T. (1995). Regulation of insecticidal crystal protein production in Bacillus thuringiensis. Mol. Microbiol. *18*, 1–12.
39. Aronson, A.I. (1993). The two faces of Bacillus thuringiensis: insecticidal proteins and post-exponential survival. Mol. Microbiol. *7*, 489–496.
40. Sawaya, M.R., Cascio, D., Gingery, M., Rodriguez, J., Goldschmidt, L., Colletier, J.P., Messerschmidt, M.M., Boutet, S., Koglin, J.E., Williams, G.J., et al. (2014). Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. Proc. Natl. Acad. Sci. USA *111*, 12769–12774. https://doi.org/10.1073/pnas.1413456111.
41. Tetreau, G., Sawaya, M.R., De Zitter, E., Andreeva, E.A., Banneville, A.-S., Schibrowsky, N.A., Coquelle, N., Brewster, A.S., Grünbein, M.L., Kovacs, G.N., et al. (2022). De novo determination of mosquitocidal Cry11Aa and Cry11Ba structures from naturally-occurring nanocrystals. Nat. Commun. *13*, 4376. https://doi.org/10.1038/s41467-022-31746-x.
42. Bittrich, S., Segura, J., Duarte, J.M., Burley, S.K., and Rose, Y. (2024). RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments. Bioinformatics *40*, btae370. https://doi.org/10.1093/bioinformatics/btae370.
43. Bietlot, H.P., Vishnubhatla, I., Carey, P.R., Pozsgay, M., and Kaplan, H. (1990). Characterization of the cysteine residues and disulphide linkages in the protein crystal of Bacillus thuringiensis. Biochem. J. *267*, 309–315.
44. Creighton, T.E. (1986). Disulfide bonds as probes of protein folding pathways. Methods Enzymol. *131*, 83–106. https://doi.org/10.1016/0076-6879(86)31036-x.
45. Permyakov, E.A. (2021). Metal Binding Proteins. Encyclopedia *1*, 261–292.
46. Palm-Espling, M.E., Niemiec, M.S., and Wittung-Stafshede, P. (2012). Role of metal in folding and stability of copper proteins in vitro. Biochim. Biophys. Acta *1823*, 1594–1603. https://doi.org/10.1016/j.bbamcr.2012.01.013.
47. McConnell, K.D., Fitzkee, N.C., and Emerson, J.P. (2022). Metal Ion Binding Induces Local Protein Unfolding and Destabilizes Human Carbonic Anhydrase II. Inorg. Chem. *61*, 1249–1253. https://doi.org/10.1021/acs.inorgchem.1c03271.
48. Ozkan, M., Dilek, F.B., Yetis, U., and Ozcengiz, G. (2003). Nutritional and cultural parameters influencing antidipteran delta-endotoxin production. Res. Microbiol. *154*, 49–53. https://doi.org/10.1016/s0923-2508(02)00006-2.
49. Liu, Z., Xie, J., Deng, Z., Wang, M., Dang, D., Luo, S., Wang, Y., Sun, Y., Xia, L., and Ding, X. (2020). Enhancing the insecticidal activity of new Bacillus thuringiensis X023 by copper ions. Microb. Cell Fact. *19*, 195. https://doi.org/10.1186/s12934-020-01452-8.
50. Domanska, B., Fortea, E., West, M.J., Schwartz, J.L., and Crickmore, N. (2019). The role of membrane-bound metal ions in toxicity of a human cancer cell-active pore-forming toxin Cry41Aa from Bacillus thuringiensis. Toxicon *167*, 123–133. https://doi.org/10.1016/j.toxicon.2019.06.003.
51. Kirouac, M., Vachon, V., Quievy, D., Schwartz, J.L., and Laprade, R. (2006). Protease inhibitors fail to prevent pore formation by the activated Bacillus thuringiensis toxin Cry1Aa in insect brush border membrane vesicles. Appl. Environ. Microbiol. *72*, 506–515. https://doi.org/10.1128/AEM.72.1.506-515.2006.
52. Monette, R., Potvin, L., Baines, D., Laprade, R., and Schwartz, J.L. (1997). Interaction between Calcium Ions and Bacillus thuringiensis Toxin Activity against Sf9 Cells (Spodoptera frugiperda, Lepidoptera). Appl. Environ. Microbiol. *63*, 440–447. https://doi.org/10.1128/aem.63.2.440-447.1997.
53. Katayama, H., Kusaka, Y., Yokota, H., Akao, T., Kojima, M., Nakamura, O., Mekada, E., and Mizuki, E. (2007). Parasporin-1, a novel cytotoxic protein from Bacillus thuringiensis, induces Ca2+ influx and a sustained elevation of the cytoplasmic Ca2+ concentration in toxin-sensitive cells. J. Biol. Chem. *282*, 7742–7752. https://doi.org/10.1074/jbc.M611382200.
54. Pokutta, S., Herrenknecht, K., Kemler, R., and Engel, J. (1994). Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. Eur. J. Biochem. *223*, 1019–1026. https://doi.org/10.1111/j.1432-1033.1994.tb19080.x.
55. Candas, M., Francis, B.R., Griko, N.B., Midboe, E.G., and Bulla, L.A., Jr. (2002). Proteolytic cleavage of the developmentally important cadherin BT-R1 in the midgut epithelium of Manduca sexta. Biochemistry *41*, 13717–13724. https://doi.org/10.1021/bi026323k.
56. Stewart, D.E., Sarkar, A., and Wampler, J.E. (1990). Occurrence and role of cis peptide bonds in protein structures. J. Mol. Biol. *214*, 253–260. https://doi.org/10.1016/0022-2836(90)90159-J.
57. Exarchos, K.P., Exarchos, T.P., Rigas, G., Papaloukas, C., and Fotiadis, D.I. (2011). Extraction of consensus protein patterns in regions containing non-proline cis peptide bonds and their functional assessment. BMC Bioinf. *12*, 142. https://doi.org/10.1186/1471-2105-12-142.
58. Weiss, M.S., Metzner, H.J., and Hilgenfeld, R. (1998). Two non-proline cis peptide bonds may be important for factor XIII function. FEBS Lett. *423*, 291–296. https://doi.org/10.1016/s0014-5793(98)00098-2.
59. Stoddard, B.L., and Pietrokovski, S. (1998). Breaking up is hard to do. Nat. Struct. Biol. *5*, 3–5. https://doi.org/10.1038/nsb0198-3.
60. Bouckaert, J., Poortmans, F., Wyns, L., and Loris, R. (1996). Sequential structural changes upon zinc and calcium binding to metal-free concanavalin A. J. Biol. Chem. *271*, 16144–16150. https://doi.org/10.1074/jbc.271.27.16144.
61. Katoch, R., and Tripathi, A. (2021). Research advances and prospects of legume lectins. J. Biosci. *46*, 104. https://doi.org/10.1007/s12038-021-00225-8.
62. Sharon, N., and Lis, H. (2002). How proteins bind carbohydrates: lessons from legume lectins. J. Agric. Food Chem. *50*, 6586–6591. https://doi.org/10.1021/jf020190s.
63. Yamamoto, K., Konami, Y., and Osawa, T. (2000). A chimeric lectin formed from Bauhinia purpurea lectin and Lens culinaris lectin recognizes a unique carbohydrate structure. J. Biochem. *127*, 129–135. https://doi.org/10.1093/oxfordjournals.jbchem.a022573.
64. Strizhov, N., Keller, M., Koncz-Kálmán, Z., Regev, A., Sneh, B., Schell, J., Koncz, C., and Zilberstein, A. (1996). Mapping of the entomocidal fragment of Spodoptera-specific Bacillus thuringiensis toxin CryIC. Mol. Gen. Genet. *253*, 11–19. https://doi.org/10.1007/s004380050290.
65. Souissi, W., Kaloki, A., Etherington, S., Domanska, B., West, M.J., and Crickmore, N. (2019). Differential proteolytic activation of the Bacillus thuringiensis Cry41Aa parasporin modulates its anticancer effect. Biochem. J. *476*, 3805–3816. https://doi.org/10.1042/bcj20190732.
66. Morse, R.J., Yamamoto, T., and Stroud, R.M. (2001). Structure of Cry2Aa Suggests an Unexpected Receptor Binding Epitope. Structure *9*, 409–417. https://doi.org/10.1016/S0969-2126(01)00601-3.
67. Tabashnik, B.E., Zhang, M., Fabrick, J.A., Wu, Y., Gao, M., Huang, F., Wei, J., Zhang, J., Yelich, A., Unnithan, G.C., et al. (2015). Dual mode of action of Bt proteins: protoxin efficacy against resistant insects. Sci. Rep. *5*, 15107.
68. Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K.D., and Karsch-Mizrachi, I. (2020). GenBank. Nucleic Acids Res. *48*, D84–D86. https://doi.org/10.1093/nar/gkz956.
69. White, T.A., Kirian, R.A., Martin, A.V., Aquila, A., Nass, K., Barty, A., and Chapman, H.N. (2012). CrystFEL: a software suite for snapshot serial crystallography. J. Appl. Crystallogr. *45*, 335–341. https://doi.org/10.1107/S0021889812002312.
70. Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C.B., Baslé, A., Beilsten-Edmands, J., Borges, R.J., Brown, D.G., Burgos-Mármol, J.J., Berrisford, J.M., et al. (2023). The CCP4 suite: integrative software for macromolecular crystallography. Acta Crystallogr. D Struct. Biol. *79*, 449–461. https://doi.org/10.1107/s2059798323003595.

71. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

72. Schrodinger, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8 (Schrodinger, LLC).

73. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340.

74. Williamson, L.J., Galchenkova, M., Best, H.L., Bean, R.J., Munke, A., Awel, S., Pena, G., Knoska, J., Schubert, R., Dörner, K., et al. (2023). Structure of the Lysinibacillus sphaericus Tpp49Aa1 pesticidal protein elucidated from natural crystals using MHz-SFX. Proc. Natl. Acad. Sci. USA *120*, e2203241120. https://doi.org/10.1073/pnas.2203241120.

75. Jones, G.W., Nielsen-Leroux, C., Yang, Y., Yuan, Z., Dumas, V.F., Monnerat, R.G., and Berry, C. (2007). A new Cry toxin with a unique two-component dependency from Bacillus sphaericus. FASEB J. *21*, 4112–4120. https://doi.org/10.1096/fj.07-8913com.

76. Silva-Filha, M.H., Nielsen-Leroux, C., and Charles, J.F. (1997). Binding kinetics of Bacillus sphaericus binary toxin to midgut brush-border membranes of Anopheles and Culex sp. mosquito larvae. Eur. J. Biochem. *247*, 754–761. https://doi.org/10.1111/j.1432-1033.1997.00754.x.

77. Han, H., Round, E., Schubert, R., Gül, Y., Makroczyová, J., Meza, D., Heuser, P., Aepfelbacher, M., Barák, I., Betzel, C., et al. (2021). The XBI BioLab for life science experiments at the European XFEL. J. Appl. Crystallogr. *54*, 7–21.

78. Yefanov, O., Oberthür, D., Bean, R., Wiedorn, M.O., Knoska, J., Pena, G., Awel, S., Gumprecht, L., Domaracky, M., Sarrou, I., et al. (2019). Evaluation of serial crystallographic structure determination within megahertz pulse trains. Struct. Dyn. *6*, 064702. https://doi.org/10.1063/1.5124387.

79. Wiedorn, M.O., Oberthür, D., Bean, R., Schubert, R., Werner, N., Abbey, B., Aepfelbacher, M., Adriano, L., Allahgholi, A., Al-Qudami, N., et al. (2018). Megahertz serial crystallography. Nat. Commun. *9*, 4025. https://doi.org/10.1038/s41467-018-06156-7.

80. Mancuso, A.P., Aquila, A., Batchelor, L., Bean, R.J., Bielecki, J., Borchers, G., Doerner, K., Giewekemeyer, K., Graceffa, R., Kelsey, O.D., et al. (2019). The single particles, clusters and biomolecules and serial femtosecond crystallography instrument of the European XFEL: initial installation. J. Synchrotron Radiat. *26*, 660–676.

81. Wiedorn, M.O., Awel, S., Morgan, A.J., Ayyer, K., Gevorkov, Y., Fleckenstein, H., Roth, N., Adriano, L., Bean, R., Beyerlein, K.R., et al. (2018). Rapid sample delivery for megahertz serial crystallography at X-ray FELs. IUCrJ *5*, 574–584.

82. Knoška, J., Adriano, L., Awel, S., Beyerlein, K.R., Yefanov, O., Oberthuer, D., Peña Murillo, G.E., Roth, N., Sarrou, I., Villanueva-Perez, P., et al. (2020). Ultracompact 3D microfluidics for time-resolved structural biology. Nat. Commun. *11*, 657. https://doi.org/10.1038/s41467-020-14434-6.

83. Oberthuer, D., Knoška, J., Wiedorn, M.O., Beyerlein, K.R., Bushnell, D.A., Kovaleva, E.G., Heymann, M., Gumprecht, L., Kirian, R.A., Barty, A., et al. (2017). Double-flow focused liquid injector for efficient serial femtosecond crystallography. Sci. Rep. *7*, 44628. https://doi.org/10.1038/srep44628.

84. Mariani, V., Morgan, A., Yoon, C.H., Lane, T.J., White, T.A., O'Grady, C., Kuhn, M., Aplin, S., Koglin, J., Barty, A., and Chapman, H.N. (2016). OnDA: online data analysis and feedback for serial X-ray imaging. J. Appl. Crystallogr. *49*, 1073–1080. https://doi.org/10.1107/S1600576716007469.

85. Gevorkov, Y., Yefanov, O., Barty, A., White, T.A., Mariani, V., Brehm, W., Tolstikova, A., Grigat, R.-R., and Chapman, H.N. (2019). XGANDALF–extended gradient descent algorithm for lattice finding. Acta Crystallogr. A Found. Adv. *75*, 694–704.

86. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674. https://doi.org/10.1107/s0021889807021206.

87. Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. *62*, 1002–1011. https://doi.org/10.1107/s0907444906022116.

88. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. *67*, 355–367. https://doi.org/10.1107/s0907444911001314.

89. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. https://doi.org/10.1107/s0907444904019158.

Please cite this article in press as: Best et al., The long and short of it: Distinct natural crystal packing strategies of Cry toxins from *Bacillus thuringiensis*, Structure (2026), https://doi.org/10.1016/j.str.2026.01.014

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Bacillus thuringiensis strain Na210 encoding Cry8Ba2	This paper	N/A
Bacillus thuringiensis strain 4D7 (pHT1Ca) expressing recombinant Cry1Ca18	This paper	N/A
Deposited data		
Cry8Ba2 – compact variant	This paper	PDB: 9H9B
Cry8Ba2 – extended variant	This paper	PDB: 9H9A
Cry1Ca18 – pH7	This paper	PDB: 9H99
Cry1Ca18 – pH9	This paper	PDB: 9QXQ
Cry1Ca18	GenBank[68]	Accession number: GenBank: MZ355710
Cry8Ba2	GenBank[68]	Accession number: GenBank: PQ691233
Cry sequences for cysteine comparison	Bacterial Pesticidal Protein Resource Center (BPPRC)[6]	https://www.bpprc-db.org/home/
Software and algorithms		
CrystFEL version 0.10.1	Centre for Free-Electron Laser Science (CFEL) & Deutsches Elektronen-Synchrontron (DESY)[69]	https://www.desy.de/~twhite/crystfel/
CCP4i2	Collaborative Computational Project No. 4[70]	https://www.ccp4.ac.uk/
AlphaFold3	AlphaFold Server[71]	https://alphafoldserver.com/welcome
CheckMyMetal	CheckMyMetal metal binding site validation server[28]	https://cmm.minorlab.org/
PyMOL 3.1	Schrödinger[72]	https://www.pymol.org/
MUSCLE algorithm	Unipro UGENE version 51[73]	https://ugene.net/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains used in this study include the *B. thuringiensis* strain Na210, encoding the *cry8Ba2* gene (accession number GenBank: MZ355710) and the 4D7 (pHT1Ca) strain expressing recombinant *cry1Ca18* gene (accession number: GenBank: PQ691233).

METHOD DETAILS

Purification and sample preparation of Cry8Ba2 and Cry1Ca18 protein crystals

Cry8Ba2 and Cry1Ca18 protein crystals were purified as previously described.[74] Briefly, the *B. thuringiensis* strain Na210, encoding the *cry8Ba2* gene (accession number GenBank: MZ355710) or the 4D7 (pHT1Ca) strain expressing recombinant *cry1Ca18* gene (accession number GenBank: PQ691233)[36] were grown in 400 mL Embrapa medium[75] containing 5 μg/mL erythromycin (for Cry1Ca18) at 30°C with shaking (200 rpm) until sporulation reached >90%, as judged by phase contrast microscopy. Sporulated cultures were harvested, and the natural crystal proteins were isolated using stepped sucrose gradients as previously described.[74,76] Crystals were washed with deionized water (ddH_2O) and filtered through a series of nylon mesh filters (Sysmex, Celltrics) with decreasing pore sizes, ranging from 100 μm down to 5 μm. The crystal suspension was then centrifuged at 200 × g for one minute, and the supernatant containing the nanocrystals underwent additional filtration and washing. The prepared crystal suspension was transferred to high-pressure sample reservoirs for injection into the XFEL beam.

Transmission electron microscopy

Purified crystal batches were characterised using transmission electron microscopy (JEM 2100-Plus, JEOL) in the XBI lab of European XFEL as previously described.[74,77] Briefly, Holey carbon copper grids (Quantifoil R1.2/1.3) were glow-discharged with a

GloQube Plus system (Quorum Technologies) immediately prior to sample preparation. A 2 μL aliquot of the crystal suspension was deposited on each grid, allowed to sit for roughly 30 seconds, and then blotted with Whatman No. 1 filter paper. For negative staining, the grids were first touched to a droplet of 2% (w/v) uranyl acetate and blotted without delay. They were then transferred to a fresh droplet of the same stain, incubated for about 20 seconds, blotted again, and left to dry on filter paper. Imaging was performed on a TEM operated at 200 kV, using an Emsis Xarosa camera in both imaging mode and selected-area electron diffraction mode.

XFEL data collection

Sample delivery and data collection were performed as described previously.[74] For data collection at pH 9.0, a 0.1 M sodium carbonate buffer (pH 9.0) was used. This buffer was also transferred to high-pressure sample reservoirs for injection and mixed with the Cry1Ca18 crystals via a microfluidic mixer upstream of the XFEL beam. The mixing occurred approximately 40 s before exposure to the X-rays. Megahertz serial femtosecond crystallography[78,79] diffraction data were collected at the SPB/SFX instrument[80] of the European XFEL facility (Hamburg, Germany), employing fast liquid-jet injection[81] with a 3D-printed[82] double-flow focusing nozzle (DFFN).[83] The machine was set to send 202 pulses per train (with 0.5 MHz repetition rate) to the beamline and diffraction produced by each pulse was captured with AGIPD. A photon energy of 9.3 keV with an average pulse energy of 4 mJ was delivered to the instrument, focused to a spot size of ~300 nm in diameter using the Nanoscale-focusing Kirkpatrick–Baez (KB) optics. Considering the beamline transmission, this provided an intensity of approximately $6x10^{12}$ photons/μm^2/pulse at the sample. The online crystal diffraction hit rate was monitored using the OnDA program,[84] and raw data processing largely followed the methods outlined by Wiedorn et al.[79]

Structure determination

Measured diffraction patterns were indexed using CrystFEL[69] using the indexing method XGandalf and the –multi option.[85] Merging and scaling were performed using the Partialator program from CrystFEL. For Cry1Ca18 at pH 9 CrystFEL Version 0.10.1 was used for indexing and version 0.11.1-175-gff3dc4b8 was used for post refinement, scaling and merging (xsphere, –push-res=3, 3 rounds). For Cry1Ca18 pH7 CrystFEL Version 0.10.1 was used for indexing, post refinement, scaling and merging (xsphere –push-res=1, 3 rounds), for both Cry8Ba2 datasets CrystFEL 0.9.1 was used (unity, –push-res=1, 3 rounds). Structure solution was subsequently performed using the CCP4i2 package. The number of molecules in the asymmetric unit were estimated using Matthews' analysis. Two distinct populations of Cry8Ba2 crystals were identified during data processing with CrystFEL, both of which exhibited tetragonal symmetry and could be indexed in space group $P4_12_12$ (Table S1). The first crystal form (PDB: 9H9B) was found to exhibit a shorter more compact c-axis (a=b=93.32 Å, c=275.51 Å), and is hence referred to as Cry8Ba2-Com. The final refined model had R_{work}/R_{free} of 0.188/0.252 at 2.27 Å. The second crystal form (PDB: 9H9A) was found to exhibit a longer more extended c-axis (a=b=92.23 Å, c=308.23 Å), and is hence referred to as Cry8Ba2-Ext. The final refined model had R_{work}/R_{free} of 0.188/0.235 at 2.27 Å. The following steps were performed with the CCP4 integrative software for macromolecular crystallography.[70] Structure solution of Cry8Ba2-Com was performed in PHASER[86] using *B. thuringiensis* Cry8Ea1 (PDB: 3EB7) and *B. thuringiensis* Cry1Ac-Δ14C (PDB: 4W8J), broken into the component domains, as starting models. The initial model was incomplete and, therefore, subjected to several rounds of automated model building in Buccaneer.[87] This was followed by iterative refinement and manual model rebuilding cycles in Refmac5[88] and Coot[89] respectively until the model reached convergence. Phasing of the Cry8Ba2-Ext model was also performed in PHASER using the Cry8Ba2-Com structure, solved as part of this study. The model was subjected to iterative cycles of refinement and manual model building cycles in Refmac5 and Coot respectively. Cry1Ca18 appeared in only one crystalline form, also in $P4_12_12$, a=b=88.82 Å, c=267.85 Å, and its structure was solved using the same strategy as above (Table S1).

X-ray fluorescence spectra of Cry1Ca18

Microcrystalline sample was loaded into thin nylon loops normally used for single crystal X-ray diffraction. Best results were achieved with repeated aliquoting of the sample slurry onto the loop, then drying it in a warm air stream to reduce water content. The loop was placed in the vacuum chamber on beamline I23 at Diamond Light Source, where scanning fluorescence spectra were recorded upon illumination of the sample with a fixed energy X-ray beam. Two energies were chosen, 11keV and 7keV, on either side of the Zn absorption edge. Ca absorbs at both energies, but the signal is weaker at the higher energy, the same as the inherent S in the protein. Other elements in the sample mount, e.g. Ti and Si, also produce peaks and can be discounted.

Sequence and model analysis of Cry protein cysteine content

Sequences for all Cry proteins, grouped by subfamily (Cry1, Cry2, Cry3 etc.) were obtained from the BPPRC database. Identical sequences were filtered out and, in large subfamilies (eg Cry1), a subsample of sequences was selected for analysis (Data S1). Overall, approximately half the Cry sequences available in the current BPPRC database were analysed. For these proteins, cysteine content was noted along with the 'form' of Cry protein: those with a 'crystallisation domain' denoted as *Long* (>~1100 residues), those without as *Short* (~600-800 residues), those intermediate as *Medium* (800 < x < 1100 residues), and those with separate CDS for the toxin core and 'crystallisation domains' as *Split*. Alignments of the whole Cry family, each individual Cry group (Cry1, Cry2 etc.), *Long* forms, and *Split* forms were generated using the MUSCLE algorithm in Unipro UGENE version 51 on an Apple Macbook Air 2020 M1, to determine conserved cysteine residues. To determine localisation of cysteine residues across the structural family, models for the holotype of each class (totalling 58 models) were generated using AlphaFold3 via the AlphaFold server. The server was used in default mode, inputting sequences, and running predictions. Models were visually inspected and compared to experimental

structures to assess the quality of predictions e.g. correct number of domains and similar domain architecture compared to published structures. *Split* form proteins were predicted by inputting the two separate CDS as individual amino acid sequences. Model superposition was performed in PyMOL 3.1 using the inbuilt align function, outlier removal function enabled. Structural alignments and comparisons were made of the toxin cores and crystallisation domains independently to avoid problematic positioning of domains relative to each other.

Cry protein pH digest analysis

For both Cry1 and Cry8, 1.5 mL of crystal suspension were pelleted, resuspended in 1.5 mL dH_2O. Pelleting and resuspension were repeated, before crystals were resuspended in a final 1.5 mL volume of dH_2O. Aliquots (100 μL) were then pelleted and resuspended in 100 μL of a range of buffers, varying in pH. Buffers used were citrate pH 3.0, acetate pH 5.0, MilliQ water, HEPES pH 7.0, Tris-Base pH 8.0, Tris-Base pH 8.5, CAPS pH 9.0, CAPS pH 9.5, CHES pH 10.0, and CHES pH 11.0. Samples were incubated in the presence of 2 μl of protease inhibitor (1 cOmplete Protease Inhibitor Cocktail (EDTA-Free) tablet (Roche) dissolved in 300 μL MilliQ water). All buffers were prepared at 50 mM concentration, with ionic strength maintained at 50 mM by addition of sodium chloride. Samples were incubated at 27°C, with agitation, overnight before being quenched by the addition of an equivalent volume of 250 mM HEPES pH 7.0. Samples were centrifuged at 15,000 x *g* for 10 minutes, and the supernatant analysed on 4-20% gradient SDS-PAGE under reducing conditions (5% (v/v) β-mercaptoethanol). A solubilisation control was prepared by mixing crystals with SDS buffer. To detect protein bands, gels were stained with Coomassie Blue R-250 Brilliant Blue.

QUANTIFICATION AND STATISTICAL ANALYSIS

Crystallography data collection and refinement statistics are summarized in Table 1.