Structure and function of a fungal AB toxin-like chimerolectin involved in anti-nematode defense

Stefanie S. Schmieder[1*‖], Gabriele Cordara[2*#], Flore Kersten[2,3], Kevin Steiner[1&], Clara H. Samin[2], David F. Plaza [1\\], Ahmad A. Ahmad[, 3], Andreas Boeggild[4], Jesper L. Karlsen[4], Blanka O. Sokolowska[1], Thomas Boesen[4], Ute Krengel[2#] and Markus Künzler[1#]

[1]Institute of Microbiology, Department of Biology, ETH Zürich, Zürich, Switzerland

[2]Department of Chemistry, University of Oslo, Norway

[3]Centre for Molecular Medicine Norway, University of Oslo, Norway

[4]Department of Molecular Biology and Genetics, Aarhus University, Denmark

*Contributed equally to the work

[‖]present address: Division of Gastroenterology, Boston Children's Hospital, Harvard Medical School, Boston, USA

[\\]present address: Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden

[&]present address: Institute of Medical Virology, University of Zürich, Zürich, Switzerland

[#]Corresponding authors: gabriele.cordara@kjemi.uio.no and ute.krengel@kjemi.uio.no (structure), mkuenzle@ethz.ch (biochemistry and function)

ABSTRACT

Fungal defense against predators largely relies on protein toxins, many of which are lectins. We previously showed that the production of the nematotoxin CCTX2 is upregulated in the Agaricomycete *Coprinopsis cinerea* upon predation by nematodes. Here, we classify CCTX2 as the founding member of a family of fungal chimerolectins. Cryo-EM analysis to 3.2 Å resolution reveals five domains. The four N-terminal β-trefoil fold (BTF) domains cradle a C-terminal domain, which exhibits a novel α+β protein fold. Mutational analysis shows that both N-terminal and C-terminal domains are required for nematotoxicity. While the biochemical function of the C-terminal domain remains unclear, the first two BTF domains enable CCTX2 to bind to glycosphingolipids with LacNAc or LacdiNAc glycoepitopes on nematode intestinal epithelial cells. Experiments in the model nematode *Caenorhabditis elegans* demonstrate that the chimerolectin CCTX2 exploits the endocytic and retrograde trafficking machinery of the target cell to exert its toxicity and obtain access to the yet-to-be-identified intracellular target of the non-lectin domain. The structure and mode of action of CCTX2 is reminiscent of bacterial and plant AB toxins.

35 **INTRODUCTION**

36 Proteinaceous toxins are involved in many inter-organismal conflicts and are especially prevalent in
37 plants and bacteria (Aravind et al, 2012; Zhang et al, 2012). AB toxins are one form of such defense
38 toxins and are formed by two protein components: the A subunit (**A**ctive subunit), responsible for
39 exerting toxicity inside the cell, and B subunits (**B**inding subunits), responsible for cell binding and
40 entry (Song, 2022). Ricin (plants), Shiga and cholera toxin (bacteria) are considered prime examples
41 of this class of toxins (Zhang et al, 2012; Melton-Celsa, 2014; Clemens et al, 2017; Polito et al, 2019).
42 AB toxins are found in nature with varying subunit stoichiometries, ranging from AB (*e.g.,* ricin), to
43 AB_2 (*e.g.,* cytolethal distending toxin, CDT), AB_5 (*e.g.,* Shiga and cholera toxins) or A_2B_5 (*e.g.,* typhoid
44 toxin) (Pickett & Whitehouse, 1999; Vanden Broeck et al, 2007; Song et al, 2013a; Melton-Celsa,
45 2014; Clemens et al, 2017; Polito et al, 2019).

46 The B subunit often recognizes a glycoepitope with high specificity, thus classifying it as a sugar-
47 binding protein, or lectin. This subunit mediates the binding of the AB-toxin to a glycoconjugate on
48 the target cell surface and triggers the entry of the toxin into the target cell by endocytosis
49 (Mancheño et al, 2005; Manna et al, 2017). Once inside, the A subunit exerts its toxicity by binding to
50 and/or modifying one or more intracellular target molecules (Teter, 2013). Catalytic activities of AB
51 toxins include ADP-ribosyltransferase (ART, cholera toxin) and rRNA *N*-glycosidase activity (ricin and
52 Shiga toxin) (Sánchez & Holmgren, 2011; Melton-Celsa, 2014; Polito et al, 2019).

53 The sessile lifestyle and the high concentration of macronutrients render fungi an attractive food
54 source for the soil-grazing fauna. To protect themselves from being foraged by *e.g.* nematodes, fungi
55 employ a combination of secondary metabolites and a protein-based defense system; the latter
56 including several lectins (Singh et al, 2010; Bleuler-Martínez et al, 2011; Kempken, 2011; Varrot et al,
57 2013). Fungivorous nematodes pierce fungal cells, and aspirate their cytoplasmic content;
58 consistently, defense lectins are cytoplasmically localized (Künzler, 2015). After ingestion, the lectin
59 travels through the digestive tract of nematodes to eventually bind to a cell surface receptor and
60 exert its toxic effect (Künzler, 2015).

61 Lectins include mero- (monomeric) and hololectins (multimeric), exclusively formed by sugar-binding
62 domains, and chimerolectins, where lectin and non-lectin domains coexist or assemble in the same
63 protein (Peumans & Van Damme, 1995; Peumans et al, 2001). Some chimerolectins are functionally
64 analogous to AB toxins. Studies of the fungal chimerolectin *Marasmius oreades* agglutinin (MOA)
65 suggest that the lectin domain acts as a B subunit and the non-lectin domain as an A subunit
66 (Cordara et al, 2011; Wohlschlager et al, 2011; Cordara et al, 2016). The toxicity mechanism of MOA
67 includes binding to glycosphingolipids exposed on the surface of the target cell as well as
68 endocytosis, retrograde trafficking, and (putative) modification of an intracellular target (Cordara et
69 al, 2011; Wohlschlager et al, 2011; Cordara et al, 2014; Künzler, 2015; Juillot et al, 2016). MOA exerts
70 strong toxicity against nematodes, including the bacterivorous model organism *Caenorhabditis*

71 *elegans*, and is thought to be part of the chemical defense of the host *M. oreades* against these
72 micropredators (Wohlschlager et al, 2011; Tayyrov et al, 2018).

73 CCTX2 was previously identified as a nematotoxic defense protein in the Agaricomycete *Coprinopsis*
74 *cinerea*, based on the induced expression of the *cctx2* gene in the vegetative mycelium of the fungus
75 in the presence of fungivorous nematodes (Plaza et al, 2016). In the present article, we performed a
76 structure/function analysis of the CCTX2 protein toxin. Within the *C. cinerea* genome, we additionally
77 identified the genes of two CCTX2 paralogs, which we termed *cctx1* and *cctx3*. The three genes differ
78 in their expression pattern regarding sexual development and bacterial or nematode challenge;
79 however, we demonstrate that the encoded proteins all exert strong nematotoxicity and show the
80 same domain partition: four N-terminal β-trefoil fold (BTF) domains followed by a C-terminal domain
81 of unknown function (DUF). The structure of CCTX2, determined to 3.2 Å resolution by single-particle
82 cryo–electron microscopy (cryo-EM), revealed that the BTF domains form a cradle carrying the C-
83 terminal domain. The C-terminal domain adopts a novel α+β protein fold, not represented in any
84 protein structures deposited in the Protein Data Bank (PDB). Deletion of the first two BTF domains of
85 CCTX2 leads to the loss of binding of the chimerolectin to LacNAc and LacdiNAc glycoepitopes. This,
86 together with the presence of experimental density at the carbohydrate-binding sites in the two N-
87 terminal BTF domains, suggests that the nematotoxicity of CCTX2 depends on the binding of the
88 protein to LacdiNAc-containing glycosphingolipids (GSLs) on the surface of nematode intestinal
89 epithelial cells. Upon binding, CCTX2 undergoes endocytosis and retrograde trafficking of the
90 protein-GSL complexes in these cells. The nematotoxicity of CCTX2 also depends on the presence and
91 integrity of the C-terminal, non-lectin domain. The intracellular target and the biochemical function
92 of this domain remain unknown. Based on orthologues encoded in the genome of other
93 Agaricomycetes and some ascomycetes, we conclude that CCTX2 is the founding member of a family
94 of fungal defense effectors against fungivorous nematodes.

95

RESULTS

CCTX2 belongs to a family of fungal nematotoxins

A query on the *Coprinopsis cinerea* genome (JGI Mycocosm *Coprinopsis cinerea* AmutBmut v2) using the *cctx2* gene sequence (gene model ID CopciAB_369589.T0) revealed the existence of two paralogues, aptly designated as *cctx1* and *cctx3* (gene model IDs CopciAB_369594.T0 and CopciAB_369576.T0, respectively), which are located on the same chromosome in close neighborhood of the *cctx2* gene (Fig. 1A). According to published data (Muraguchi et al, 2015; Tayyrov et al, 2018; Kombrink et al, 2019) (Fig. S1A-B), the three paralogues differ significantly in their expression patterns: *cctx2* is expressed at low levels in the vegetative mycelium and is induced upon predation by the fungivorous nematode *Aphelenchus avenae* but not upon bacterial challenge. Expression of *cctx1* is low also in vegetative mycelium, but the gene is strongly upregulated upon nematode challenge and in stage 1 primordia during sexual development, whereas *cctx3* is strongly expressed in the vegetative mycelium and downregulated during sexual development as well as upon nematode and bacterial challenge. A BLAST search using the full-length CCTX1, CCTX2 and CCTX3 sequences (JGI Mycocosm *Coprinopsis cinerea* AmutBmut v2 protein IDs 2257338, 2257340 and 2257352, respectively) revealed the existence of several CCTX homologues (Tayyrov et al, 2018; Kombrink et al, 2019) over the full length of the protein, specific to the fungal kingdom (Fig. 1B, Fig. S1D; Table S1). Most homologues were detected in Agaricomycetes, with a remarkable expansion in the genus *Pisolithus*. Additionally, a few homologues were encoded by Ascomycetes.

The specific induction of *cctx1* and *cctx2* by *A. avenae* suggests a function of these genes in the defense of *C. cinerea* against fungivorous nematodes. Indeed, we previously showed that CCTX2 is toxic towards the nematode model species, *Caenorhabditis elegans* (Plaza et al, 2016). To test if all three paralogous genes encode defense effector proteins, we produced the proteins in *E. coli* and tested their toxicity against *C. elegans.* All three paralogous proteins inhibited the development of *C. elegans* L1 larvae at low concentrations (<50 µg/ml, Fig. 1C). The inhibition of larval development was comparable to the effects of the *Marasmius oreades* chimerolectin MOA and much stronger than for the *C. cinerea* hololectin CCL2 (Fig. S2) (Wohlschlager et al, 2011). Feeding of TAMRA-labeled CCTX2 protein to L4 staged *C. elegans* GK70 larvae (harboring GFP-labeled apical plasma membrane marker protein PGP-1) revealed colocalization of the chimerolectin CCTX2 and PGP-1 with the intestinal epithelium (Fig. 1D). Accordingly, exposure to CCTX2 resulted in clear morphological changes of the intestine, such as gross distension of the intestinal lumen and an increasingly undulating phenotype of the apical plasma membrane after more than 2h of feeding (Fig. 1E). These morphological changes increased in severity over time, leading to death of the worms around 16 to 24 h after toxin administration (data not shown). Similar morphological changes to the intestinal phenotypes were previously described for other nematotoxic fungal lectins, such as the chimerolectin MOA and the two *C. cinerea* hololectins CGL2 and CCL2 (Butschi et al, 2010;

132 Wohlschlager et al, 2011; Schubert et al, 2012; Stutz et al, 2015). Besides *C. elegans*, we found that
133 CCTX2 was also toxic to the larvae of two other *Caenorhabditis* species (Fig. S1C). A similarly broad
134 specificity as CCTX2 was also observed for the chimerolectin MOA but not for the hololectin CGL2x.

135

136 ***CCTX2 fold comprises four β-trefoil domains and a C-terminal domain of unknown function (DUF)***

137 Sequence analysis revealed that CCTX1, CCTX2 and CCTX3 exhibit four ricin B chain-like domains at
138 the N-terminus, followed by a 220-230 amino acid domain with no homologue in the known
139 proteome (Fig. S1D). The first two domains of each orthologue contain the canonical $(QxW)_3$ motif,
140 characteristic of R-type lectins (Fig. S1D) (Yao et al, 2011), implying a possible lectin function (*i.e.*,
141 binding to glycans or glycoconjugates). An *in silico* structural model, generated in pre-AlphaFold era
142 (Jumper et al, 2021) using the Robetta server (Kim et al, 2004; Yang et al, 2020), showed that the
143 ricin B chain-like domains adopt the stereotypical β-trefoil fold (BTF); no fully reliable model could be
144 generated for the C-terminal domain (Fig. S3). To confirm these predictions and gain a better
145 understanding of the molecular function of the individual domains and their interplay, we
146 determined the CCTX2 structure by single-particle cryo-EM to a maximum resolution of 3.2 Å (Fig. 2,
147 Fig. S4; Table S2). Residues 1 to 787 (full sequence) were visible in the reconstructed volume; regions
148 with poorly resolved density are reported in Fig. S3D and Table S3. The structure shows a single
149 polypeptide chain with five identifiable domains connected by short linkers, matching extant
150 predictions (Fig. 2A); domain boundaries are reported in Table S4. Overall, the protein adopts a
151 compact, globular fold (Fig. 2A and 2B). *B*-factor analysis shows the presence of two distinct
152 structure-functional units: BTF domains 1-3, characterized by low flexibility (low *B*-factors), and
153 domain 4-5, showing much higher flexibility/disorder (Fig. 2C). The four ricin-B-chain-like domains
154 adopt the expected β-trefoil fold (BTF) and are arranged in a rhomboid shape, forming a cradle that
155 supports the fifth domain (Fig. 2B). Each BTF domain displays the characteristic arrangement of the
156 α, β and γ subdomains around a pseudo-threefold axis (Fig. S5A); α, β and γ subdomain boundaries
157 are reported in Table S4. A per-domain search for structural homologues using the DALI server
158 (Holm, 2022) returned matches among other fungal lectins and bacterial toxins (Table S6). Structural
159 alignment shows strong similarity among the four BTF domains, despite low sequence conservation
160 (Table S5); however, a visual inspection revealed that the fourth BTF domain contains an incomplete
161 β subdomain (Fig. S5A). The putative sugar-binding sites in all the BTF domains of CCTX2 are located
162 on the outer side of the 'cradle' and are thus fully exposed to the solvent (Fig. S5B and S5C). The
163 reconstructed Coulomb potential map showed additional density (Fig. S5D) at the sugar binding
164 pocket of subdomains β and γ of the first BTF domain and subdomain α of the second BTF domain
165 (sites marked in magenta in Fig. 2A and S5). The detected density can probably be traced to residual
166 D-galactose, used throughout the purification process. The canonical $(QxW)_3$ motifs of R-type lectins
167 are usually involved in the binding of glycans (Yao et al, 2011), which is in good agreement with the

168 presence of an intact $(QxW)_3$ motif only in the first two BTF domains. Finally, three Cys residues (514,
169 515, 550) form a putative Zn^{2+} binding site, well nestled into BTF domain 4 (Fig. 2B, inset).

170 The fifth domain (amino acids 557 to 787) adopts a unique α+β fold (Fig. 3), bearing no similarity to
171 any known protein structure. A search for structural homologues was performed using the DALI
172 server (Holm, 2022). Either the full domain or its core portion (residues 557-570 + 620-655 + 684-
173 745) did not return any significant hit, indicating the discovery of a new α+β-type protein fold. The
174 core of the C-terminal domain is formed by a seven-stranded antiparallel β-sheet, supported by two
175 α-helices; the β-sheet changes orientation between strands 6 and 7 (Fig. 3A and 3B). A large portion
176 of the β-sheet is solvent-exposed and forms a potential catalytic cleft. The core of the domain is
177 decorated with two large insertions. The first insert (residues 571-619) contains a helix and a long
178 stretch of random coil that tightly interact with the third BTF domain; the second insert (residues
179 656-683) forms a two-helix subdomain, packed at the sides of the central core (Fig. 3B and 3C). The
180 C-terminal residues (746-787) form a long random coil segment, circling the fourth BTF domain and
181 ending in a helix and another random coil, sandwiched between the first insert and the central core
182 (Fig. 3B and 3C). Interestingly, the C-terminal loop is in direct contact with the putative Zn^{2+} binding
183 site of CCTX2 (Fig. 2B, inset).

184

185 ***Nematotoxicity of CCTX2 is dependent on the two N-terminal lectin domains and the C-terminal***
186 ***DUF domain***

187 To characterize the molecular functions of the different CCTX2 domains, we generated both N- and
188 C-terminally truncated variants of the protein (Fig. 4A). Since both sequence and structural analysis
189 predicted a lectin function for the first two BTF domains, we produced a variant lacking these two
190 domains (residues 304 to 787), termed CCTX2ΔN. To assess carbohydrate-binding, we probed
191 AlexaFluor488-labeled versions of both full-length CCTX2 and CCTX2ΔN for binding to the CFG
192 mammalian glycan array v5.1, using the service provided by the Consortium of Functional Glycomics
193 (CFG, Emory University, USA). The data show that CCTX2 is indeed a lectin with specificity towards
194 LacNAc(Galβ1-4GlcNAc)- and LacdiNAc(GalNAcβ1-4GlcNAc)-containing glycans (Fig. S6 and
195 supplementary data files). No binding was detected for CCTX2ΔN (Supplementary data files), but the
196 protein was soluble suggesting structural and functional integrity. Overall, results indicate that the
197 first two BTF domains of CCTX2 are required for its carbohydrate-binding activity, in agreement with
198 the sequence predictions and the structural results. To test whether the carbohydrate-binding
199 activity of CCTX2 is necessary for its nematotoxicity, we analyzed CCTX2ΔN with our nematotoxicity
200 assay (Fig. 4B). We found that this protein variant was not nematotoxic, strongly suggesting that the
201 lectin function is essential for the nematotoxicity of CCTX2.

202 Although the C-terminal domain does not show any structural similarity to known proteins, it is well
203 conserved among the three CCTX paralogues, with a sequence identity of approximately 40%
204 (residues 555 to 787). To assess its role for nematotoxicity, we created a variant of CCTX2, termed
205 CCTX2ΔC, which lacks this domain (residues 570 to 787) (Fig. 4A). This variant exhibited severely
206 reduced nematotoxicity, indicating that this domain is also essential to mediate fungal defense (Fig.
207 4B). Currently, we can only speculate about the molecular function of this domain. An amino acid
208 sequence alignment between the members of the monophyletic clade of CCTX homologues revealed
209 several conserved amino acid residues within the domain (Fig. 1B and S1D). We individually replaced
210 sixteen of these residues with alanine and evaluated the toxicity of the resulting CCTX2 variant
211 proteins in our nematotoxicity assay against *C. elegans* (Fig. S1D and Table S7). We also mapped the
212 residues onto the structure (Fig. 3D). Nine of these CCTX2 variants showed a reduction in toxicity of
213 CCTX2 towards *C. elegans* (Table S7, Fig. 4C and S7A and S7B). Of these, substitutions D687A and
214 Q689A had the strongest effects with a reduction of the toxicity of CCTX2 towards *C. elegans* by two
215 orders of magnitude (Fig. 4C). A double replacement of both residues ('DxQ' variant), as well as a
216 triple variant ('RxDxQ' variant) that includes neighboring residue Arg685, whose individual
217 substitution caused only a weak reduction in toxicity (Fig. 4C), rendered the protein completely non-
218 toxic. We renamed these three amino acids the 'RxDxQ motif' (referring to Arg685, Asp687 and
219 Gln689 in CCTX2) (Fig. 4C). In the CCTX2 structure, this motif is exposed on the surface of the central
220 β-sheet (Fig. 3D). The single substitutions D585A and T739A in the C-terminal domain also resulted in
221 a significant reduction in toxicity (non-toxic up to a concentration of 500 mg/l), while the variants
222 R724A as well as Q635A and E656A were less strongly reduced, being non-toxic at a concentration
223 below 200 mg/l and 100 mg/l, respectively (Fig. S7A, Fig. S8 and Table S7).

224 Apart from the 'RxDxQ motif', we found that the C-terminal domain contains a conserved putative
225 cleavage site (Lys596 and Arg597 in CCTX2) for the Golgi-localized endoproteinase Kexin (Fig. 3B-E).
226 Substitution of these two residues (and the preceding Glu595) to alanine in CCTX2 resulted in a
227 reduction of toxicity of this protein variant, CCTX2-KR, against *C. elegans* (Fig. 4D). This site is
228 embedded in the second structural insert, on a solvent-exposed α-helix (Fig. 3C). To check whether
229 this motif was functional, purified CCTX2 protein was incubated with Kex2p protease from
230 *Saccharomyces cerevisiae* and the reaction was subjected to a subsequent pulldown experiment
231 using Sepharose beads. Analysis of the various fractions by SDS-PAGE and immunoblotting, using a
232 custom-made anti-CCTX2 antiserum, showed that CCTX2 was cleaved into two fragments of ≈23 and
233 ≈67 kDa, corresponding to the molecular masses expected from a proteolytic processing at the
234 presumptive cleavage site (Fig. S7C). Intriguingly, however, Kex2p cleavage alone was insufficient to
235 separate the two fragments, as we could not detect the C-terminal fragment, neither in the pulldown
236 flow-through nor in the wash. In conclusion, the C-terminal fragment of CCTX2 is tightly associated
237 with the N-terminal fragment and can only be separated under denaturing conditions, such as SDS-
238 PAGE (Fig. S7C).

239 Finally, we noticed a conserved motif, Asp572 to Leu 576, which we refer to as DDDxL, located in
240 between BTF4 and the 'KR motif' (Fig. 3D-E). Substitution of all three Asp residues and the terminal
241 Leu residue of this motif to alanine resulted in a protein (CCTX2-DDDxL) that was reduced in toxicity
242 (non-toxic at a concentration below 100 mg/l) (Fig. S7A, Fig. S8 and Table S7) to similar extent as the
243 single substitutions of residues Gln635 and Glu656 (see above).

244 Taken together, this mutational analysis confirmed that the C-terminus of CCTX2 is essential for its
245 nematotoxicity but did not pinpoint the biochemical function of this domain.

246

247 ***CCTX2 binds to C. elegans intestinal epithelial glycosphingolipids in vivo and in vitro***

248 The dependence of the nematotoxic activity of CCTX2 on two lectin domains (Fig. 4B) and the
249 localization of TAMRA-labeled CCTX2 to the intestinal epithelial membrane in *C. elegans* (Fig. 1C and
250 1D) suggest that this chimerolectin targets a glycoconjugate present in the apical intestinal plasma
251 membrane of nematode enterocytes. The specificity of CCTX2 towards LacNAc, and LacdiNAc
252 containing motifs (Fig. S6), the binding of nematotoxic chimerolectin MOA to *C. elegans*
253 glycosphingolipids (GSLs) (Wohlschlager et al, 2011) and the presence of LacdiNAc in the core of the
254 glycan part of *C. elegans* GSLs (Barrows et al, 2006) point to these glycoconjugates as putative target
255 molecules. The biosynthesis of the *C. elegans* GSL glycans occurs by the sequential action of the
256 enzymes BRE-3 (β1,4-mannosyltransferase), BRE-5 (β1,3-*N*-acetylglucosaminyl transferase), BRE-4
257 (β1, 4 *N*-acetylgalactosaminyl transferase) and BRE-2 (β1,3-galactosyltransferase) (Fig. 5A) (Barrows
258 et al, 2006). To test the role of GSLs in toxin binding, loss-of-function *C. elegans* mutants of the four
259 GSL glycosyltransferase-encoding genes were tested for resistance towards CCTX2. All the tested *C.*
260 *elegans* GSL mutant strains displayed increased resistance to CCTX2, except for the *bre-2(ye31)*-
261 mutant strain (Fig. 5B). This result strongly suggests that the LacdiNAc motif in the core of the *C.*
262 *elegans* GSLs is the glycoepitope targeted by CCTX2. To demonstrate direct binding of CCTX2 to *C.*
263 *elegans* GSLs, we performed CCTX2 protein overlays for GSLs isolated from wildtype and mutant *C.*
264 *elegans* strains. *C. elegans* GSLs were separated by TLC and overlayed with biotinylated CCTX2, using
265 biotinylated MOA as positive control. The results showed binding of CCTX2 to GSLs from *C. elegans*
266 N2 wild-type and the *bre-2(ye31)* mutant strain, similar to the MOA positive control (Fig. 5C). No
267 binding of CCTX2 or the MOA positive control to the GSLs isolated from *bre-3, -4 or -5* strains was
268 observed, consistent with the resistance of those strains to both proteins (Wohlschlager et al, 2011).
269 In addition, we could not observe any binding to GSLs for the CCTX2ΔN variant, which lacks the
270 carbohydrate-binding domains (Fig. 5C).

271

272 ***CCTX2 undergoes endocytosis and retrograde trafficking in C. elegans enterocytes***

273 Many bacterial GSL-binding AB toxins (*e.g.,* Shiga or cholera toxins (Sandvig et al, 2014)), are
274 endocytosed upon binding to a glycolipid receptor (Fig. 6A). We hypothesize therefore that the
275 fungal chimerolectins CCTX2 and MOA follow a similar pathway. In contrast, toxic hololectins, which
276 bind to and cluster *N*-glycoprotein receptors, *i.e.* protein-bound glycans (*e.g.,* CGL2 (Butschi et al,
277 2010) and CCL2 (Schubert et al, 2012)), might not require internalization and exert toxicity by a
278 different mechanism (Stutz et al, 2015; Bleuler-Martinez et al, 2017). To assess if endocytosis of
279 CCTX2 was required for its nematotoxicity, we fed TAMRA-labeled CCTX2 to L4 larvae of *C. elegans*
280 GK70, a strain carrying the apical plasma membrane marker PGP-1 fused to GFP and monitored the
281 localization of CCTX2 using confocal fluorescence microscopy. The fungal chimerolectin MOA and the
282 hololectin CCL2 were included for comparison. After 3h of feeding, intracellular TAMRA-positive
283 vesicles were observed in *C. elegans* enterocytes (Fig. 6B). Several endocytic vesicles co-stained with
284 PGP-1::GFP were detected, indicating active endocytosis of the chimerolectin CCTX2. Indeed, we also
285 observed such endocytic vesicles for TAMRA-labeled and GSL-binding chimerolectin MOA. In
286 contrast, we did not detect TAMRA-CCL2 in endocytic vesicles, which is in agreement with previous
287 studies (Stutz et al, 2015). Additionally, we could not detect endocytic vesicles with the CCTX2ΔN
288 variant (Fig. 6B), suggesting again that GSL receptor binding is necessary for endocytosis of the
289 protein toxin.

290 To confirm the endocytosis of CCTX2 in enterocytes, we used *C. elegans* strains with fluorescently
291 marked endocytic compartments (Fig. 6A and 6C). For this purpose, we fed TAMRA-labeled CCTX2 to
292 *C. elegans* RAB-5::GFP (Rab5 is a small GTPase present on early endosomes (Bucci et al, 1992)), RAB-
293 11::GFP (Rab11 is a small GTPase on recycling endosomes (Hoekstra et al, 2004; Jing & Prekeris,
294 2009; Sato et al, 2014)); and MANS::GFP (α-Mannosidase II, an enzyme localized in the Golgi (Shah et
295 al, 2008)) strains. After 6h of feeding double positive intracellular vesicles containing both, CCTX2
296 and the endocytic marker protein were observed in enterocytes of the RAB-5::GFP strain (Fig. 6C),
297 confirming that fluorescently labeled CCTX2 was endocytosed upon receptor binding. Colocalization
298 of CCTX2 and RAB-11::GFP was also observed (Fig. 6B). Localization of CCTX2 to downstream
299 organelles such as the Golgi apparatus (*C. elegans* MANS::GFP) was rarely detected.

300 After being endocytosed, many GSL-binding AB toxins hijack retrograde vesicle trafficking to reach
301 the endoplasmic reticulum (ER) (Teter, 2013). From this compartment, many AB toxins are released
302 into the cytoplasm by retro-translocation, where they exert their toxic effect (Sandvig & van Deurs,
303 2005; Sandvig et al, 2010; Sandvig et al, 2013). The retrograde trafficking of CCTX2-containing
304 vesicles and its requirement for nematotoxicity was evaluated using *C. elegans* strains deficient in
305 different components of the retrograde trafficking machinery. Deletions in the GARP (Golgi-
306 associated retrograde protein) complex subunits VPS-52 and VPS-54 rendered the *C. elegans* partially
307 resistant to intoxication with CCTX2 (Fig. 6D). In addition, a heterozygous knock-out mutant for an
308 ESCRT-II subunit, and the VPS-34-recruiting protein BEC-1 mutant led to increased resistance to
309 CCTX2 compared to the *C. elegans* N2 wild-type strain (Fig. 6D). Neither of these strains showed

310 increased resistance towards MOA or CCL2, indicating that their resistance towards CCTX2 was toxin-
311 specific and not a pleiotropic effect due to overall changes in membrane or trafficking homeostasis
312 (Fig. S2). Given that the intracellular target for the chimerolectin MOA is unknown, we speculate that
313 this chimerolectin might not require retrograde trafficking to reach its target or uses a different
314 trafficking route.

315 The three *C. cinerea* paralogues contain a conserved C-terminal four-residue sequence (HSEL) at their
316 C-termini, resembling the canonical ER-retention signal (KDEL) present e.g. in cholera toxin for
317 retention in the ER (Fig. S1D, Fig. 3B and 3C). Removal of the sequon from CCTX2 (CCTX2ΔHSEL)
318 abolished toxicity of the protein towards *C. elegans* N2 (Fig. 6E). This indicates that the presence of
319 an intact HSEL sequon is a prerequisite for toxicity, and its removal likely affects either translocation
320 into the ER or retention of the protein in this compartment.

321 Taken together, the results suggest that CCTX2 is internalized into early endosomes (Rab5-
322 dependent). From there, the toxin can be recycled back to the plasma membrane via Rab11-positive
323 vesicles or undergoes retrograde trafficking to the TGN, Golgi and the ER facilitated by the GARP
324 complex (VPS-52 and VPS-54) and the ESCRT machinery (BEC-1).

325

326 ## DISCUSSION

327 Fungal lectins have been implicated in defense against foraging nematodes or insects (Bleuler-
328 Martínez et al, 2011; Wohlschlager et al, 2011; Tayyrov et al, 2018). In the present work, we describe
329 CCTX1, CCTX2 and CCTX3, three nematotoxic chimerolectins from the basidiomycete *C. cinerea*. We
330 performed an in-depth investigation of CCTX2, the founding member of this new family of fungal
331 chimerolectins, and provide a structure-function analysis by characterizing its cryo-EM structure, its
332 target glycoconjugate and its internalization route in the model nematode *C. elegans*.

333 ***CCTX family of chimerolectins and CCTX2—its holotype***

334 CCTX1, CCTX2 and CCTX3 are the first known representatives of a new family of fungal defense
335 proteins with a mechanism of action reminiscent of bacterial AB toxins. AB toxins are relatively rare
336 in the kingdom of fungi, with MOA and its homologue PSL1a being the most prominent members
337 described so far (Wohlschlager et al, 2011; Juillot et al, 2016). Orthologues of the CCTX family are
338 restricted to the fungal kingdom (Fig. 1B), predominantly in the phyla Basidiomycota and
339 Ascomycota. Intriguingly, the closest sequence homologue of the CCTX toxins can be found in
340 *Trichoderma harzianum*, a pathogen of plants and other fungi without known sexual stage (Gams &
341 Meyer, 1998) (THTX2, UniProt ID: A0A0F9ZAH1).

342 The *cctx1*, *cctx2* and *cctx3* genes are located next to each other in the genome of *C. cinerea*,
343 suggesting that they are the result of gene duplication events. Similar tandem arrangements of
344 defense genes can be found in *C. cinerea* for nematotoxic hololectins CGL1-CGL2 and CCL1-CCL2
345 (Boulianne et al, 2000; Schubert et al, 2012) as well as bacteriocidal cysteine-stabilized β-defensin
346 CPP1 (Copsin)/CPP2 (Kombrink et al, 2019). Gene duplications allow organisms to diversify a gene
347 without losing its original function (Moore & Purugganan, 2003). Similar to the other known
348 examples of duplicated *C. cinerea* defense genes, we observed diversification of the *cctx* genes at the
349 level of transcriptional regulation. Whereas *cctx2* is expressed at low level only in the vegetative
350 mycelium and induced upon challenge with fungivorous nematodes (Plaza et al, 2016), *cctx1* and
351 *cctx3* are constitutively expressed at significant levels in young fruiting bodies and vegetative
352 mycelium, respectively (Fig. S1A and S1B).

353 *In silico* predictions of CCTX2 (Fig. S3A-B) suggested that CCTX family members are composed of five
354 domains: four N-terminal BTF domains and a C-terminal domain of unknown function. This
355 arrangement resembles that of AB_4 toxins, represented by MTX from *Bacillus sphaericus* and pierisin-
356 1 from *Pieris rapae*, the cabbage butterfly (Treiber et al, 2008; Oda et al, 2017). However, these
357 toxins display a swapped domain organization, with the four BTF domains placed *C-terminally* and
358 the effector domain at the *N-terminus.* The structure of CCTX2, determined by cryo-EM to 3.2 Å
359 resolution, confirms the β-trefoil fold for the N-terminal domains, and reveals their homology with
360 BTF domains from other fungal lectins and bacterial toxins (Table S6), including MOA (Grahn et al,
361 2007; Grahn et al, 2009), PSL1a (Kadirvelraj et al, 2011) and the *Clostridium botulinum* hemagglutinin

362 component (HA1, (Inoue et al, 2003)). The BTF domains of CCTX2 cluster in a rhomboid fashion,
363 forming a cradle for the fifth domain (Fig. 2A and 2B). This differs from MTX or pierisin-1, where the
364 four BTF domains are arranged in a linear fashion around one side of the effector domain (Treiber et
365 al, 2008) (Fig. S9). Furthermore, the C-terminal domain of CCTX2 does not show any sequence
366 conservation or structural similarity to the effector domain of MTX and pierisin-1, which in both
367 proteins has ADP-ribosyltransferase activity (Treiber et al, 2008; Oda et al, 2017).

368 BTF domains often have lectin activity, and, accordingly, CCTX2 was confirmed to be a lectin by
369 glycan array analysis, specifically recognizing the LacNAc and LacdiNAc glycoepitopes (Fig. S6). The
370 first two N-terminal BTF domains appear to be required for sugar binding, as the N-terminal deletion
371 construct CCTXΔN failed to bind to the array (Supplementary data files). Although we cannot exclude
372 that the N-terminal deletion causes structural changes that may compromise glycan binding activity,
373 our conclusion is supported by the fact that the protein behaves well upon production and
374 purification and that the putative sugar-binding sites are clustered on the same, solvent-exposed
375 molecular surface (Fig. S5). Moreover, QxW motifs β and γ on the first domain and site α on the
376 second domain show additional density, which likely corresponds to bound D-galactose from the
377 purification process. This indicates that the two N-terminal BTF domains 1 and 2 are important for
378 glycan binding on GSL receptors. The lack of residual density at the putative sugar-binding sites of the
379 third and fourth BTF domains and the less-conserved β subdomain in the fourth BTF domain suggests
380 that BTF domains 3 and 4 have a function different from receptor recognition—potentially for
381 structural support of the C-terminal domain.

382 ***Endocytosis and intracellular trafficking of CCTX2***

383 CCTX2 enters *C. elegans* enterocytes by binding to arthroseries glycosphingolipids carrying the
384 LacdiNAc moiety (Fig. 5). *C. elegans* GSLs are also targeted by the pore-forming nematicidal toxin
385 Cry5B from *Bacillus thuringensis* as well as by the fungal chimerolectin/AB toxin MOA (Griffitts et al,
386 2005; Wohlschlager et al, 2011). Glycoproteins might provide a secondary entry route, as indicated
387 by the residual toxicity detectable in GSL-deficient mutants *bre-3*, *bre-5* and *bre-4* (Fig. 5B). Receptor
388 binding is likely mediated by the two N-terminal BTF lectin domains, since no GSL binding was
389 detected for CCTXΔN (Fig. 4C). GSL binding triggers internalization of CCTX2, as the toxin colocalized
390 with the apical plasma membrane marker PGP-1 in vesicles within *C. elegans* enterocytes (Fig. 6B).
391 Colocalization of CCTX2 with fluorescent endocytic markers, such as Rab-5 or Rab-11 (*C. elegans* Rab-
392 5::GFP and Rab-11::GFP mutants), was conclusive evidence that CCTX2 undergoes endocytosis (Fig.
393 6C).

394 Several AB toxins (*e.g.*, Shiga toxin, cholera toxin, *etc.*) take the same internalization route, followed
395 by retrograde trafficking to the ER and retro-translocation to the cytoplasm (Sandvig & van Deurs,
396 2005; Cho et al, 2012; Teter, 2013). CCTX2 retrograde trafficking was demonstrated using several *C.*
397 *elegans* mutants deficient for intracellular trafficking components (Fig. 6/Table S8). These include a

398 mutant in ESCRT-II, mutants for subunits VPS-52 and VPS-54 of the GARP (Golgi associated
399 retrograde protein) complex and a *bec-1* heterozygous mutant (Fig. 6D). In *C. elegans,* BEC-1
400 mediates autophagy and is required for retrograde transport from endosomes to the TGN (Takacs-
401 Vellai et al, 2005; Ruck et al, 2011). The toxicity resistance of the *bec-1* mutant links the toxic effect
402 of CCTX2 to its trafficking from endosome to the Trans-Golgi-Network (TGN), a necessary step in
403 retrograde trafficking. Moving backward, the GARP complex is required for tethering endosomal-
404 derived vesicles to the Golgi apparatus and is localized within the TGN (Luo et al, 2011). The toxin
405 resistance of the *vps-52* and *vps-54* mutants indicates a strong requirement for TGN transit in CCTX2
406 trafficking. None of the retrograde trafficking mutants showed resistance to the chimerolectin/AB
407 toxin MOA. This suggests that the observed effects were specific to sorting and trafficking of CCTX2,
408 and not pleiotropic effects (Fig. 6D). Apparently, the retrograde trafficking pathways of CCTX2 and
409 MOA substantially diverge from the endosome onwards. Retrograde trafficking of bacterial AB toxins
410 is often required for toxicity, with ER processing being a fundamental step in toxin activation
411 (Nowakowska-Gołacka et al, 2019). Several bacterial toxins, such as cholera toxin or *Pseudomonas*
412 exotoxin A rely on a KDEL sequon for efficient ER targeting and accumulation by Erd2, the ER protein
413 retention receptor (Lencer & Saslowsky, 2005). CCTX paralogues carry a conserved HSEL sequon at
414 the C-terminus, similar to the ER-retention signal HDEL/KDEL in yeast/mammalian cells. Removal of
415 the sequon decreased the toxicity of the protein substantially (Fig. 6E), suggesting that CCTX2
416 exploits the ER protein retention receptor Erd2 for vesicular trafficking. However, it is worth noting
417 that not all CCTX2 homologues share this C-terminal KDEL-like sequon.

418 Structural data provide further insights into the possible translocation of CCTX2 to the ER. The HSEL
419 sequence is hidden inside the bulk of the C-terminal domain (Fig. 4C), shielded from direct contact
420 with the solvent by an α-helix (residues 589-600). However, the helix carries a putative kexin
421 cleavage site (K596-R597), which is cleaved by the *S. cerevisiae* homologue Kex2p in *in vitro* assays
422 (Fig. S7C). Kexin is a TGN protease and a lower eukaryote orthologue of furin, a peptidase involved in
423 the activation of several toxins (*e.g.*, Shiga toxin) (Garred et al, 1995). The fragments resulting from
424 Kex2p cleavage do not dissociate after proteolysis (Fig. S7C), which can be explained by the highly
425 compact fold of CCTX2. Thus, we hypothesize that proteolytic processing causes a conformational
426 change in the protein, exposing the HSEL sequon to the solvent, allowing recognition by the KDEL
427 receptor and transportation back to the ER.

428 **CCTX2 effector domain**

429 The nematotoxicity of CCTX2 is dependent on its C-terminal domain of unknown function (Fig. 4B
430 and 4C), thus identifying it as the defense effector domain. Neither BLAST nor structural searches
431 revealed any similarities to proteins with known functions, leaving its biological role unclear.
432 Sequence comparisons among fungal homologues and functional analysis revealed a conserved
433 RxDxQ motif that is essential for the nematotoxicity of CCTX2 (Figs. 1B, S1D and 4C). This finding

434 does, however, not help to draw conclusions regarding the function of the domain as such a motif
435 has, to our knowledge, not been described for any other protein.

436 The cytotoxic domain adopts an α+β-fold: a large, mixed seven-stranded β-sheet supported by two
437 α-helices constitutes its core folding motif (Fig. 3). Searching the PDB for structural matches did not
438 return any hit, revealing a new protein fold. The β-sheet exhibits a large solvent-exposed surface,
439 which could act as a catalytic cleft, and is decorated by two insertions, mostly composed of α-helices
440 and random coil segments. Site-directed mutagenesis targeting 23 amino acid residues conserved
441 within the CCTX family, identified D687 and Q689 as essential for *C. elegans* toxicity (Fig. 4C and S7).
442 Their placement at the putative catalytic cleft (Fig 3D and 3E) is consistent with a potential role in
443 substrate binding. However, the fold observed in the cryo-EM structure might not be the active form
444 of the effector domain of CCTX2. As discussed in the context of the putative ER retention signal, the
445 proteolytic processing by kexin could trigger a structural rearrangement that considerably alters the
446 domain topology and eventually also the overall fold of CCTX2. Interestingly, an Ala variant targeting
447 residue Asp585, positioned in the high-*B*-factor loop immediately preceding the kexin-cleaved helix,
448 shows a decrease in activity (Fig. S7A and Fig. S8A), suggesting its active role in dynamic changes
449 affecting that region, such as a conformational change.

450 Finally, several features hint to a role of metals in the DUF domain activity. A putative Zn^{2+}-binding
451 site formed by Cys residue 514, 515 and 550 has been detected at the interface between BTF domain
452 4 and the DUF domain (Fig. 2B). The Cys arrangement suggests a structural, rather than an enzymatic
453 role for the zinc ion (Pace & Weerapana, 2014). The three cysteines are not conserved in CCTX1 and
454 CCTX3; however, they are present in other putative members of the CCTX family (Fig. S1D). More
455 intriguingly, the mutation of Thr739 to an Ala leads to a sharp decrease in toxicity. Thr739 lies at the
456 bottom of a solvent-exposed pocket, placed at the base of the DUF domain (Fig. S8C) and mostly
457 lined by negatively-charged amino acids (Glu786 and aspartates from the DDDxL motif) - an ideal
458 binding spot for divalent cations. Metal-binding has also been found to be essential for the activity of
459 other chimerolectins/toxins, *e.g*. MOA and PSL1a (Cordara et al, 2011; Wohlschlager et al, 2011;
460 Cordara et al, 2016; Cordara et al, 2017; Manna et al, 2020).

461 **Concluding remarks**

462 We provide an in-depth investigation of CCTX2, the holotype of a new family of fungal
463 chimerolectins/AB toxins. While CCTX2 undergoes the same GSL-mediated internalization and
464 retrograde trafficking as many plant and bacterial AB toxins, it reveals a unique domain topology and
465 effector domain. This domain adopts a protein fold not found in any structure deposited in the PDB.
466 Further investigation will be needed to identify the biochemical function of this domain and the
467 detailed molecular mechanism of this protein toxin.

468

469 **MATERIALS AND METHODS**

470 **Strains and cultivation conditions**

471 *Escherichia coli* strains DH5α and Turbo (New England Biolabs Inc.) were used for cloning and plasmid
472 amplification, while strains BL21(DE3) and C41(DE3) (Sigma) were used for protein production. *E. coli*
473 strains were cultivated in Luria broth (LB) medium, using well-known protocols described *e.g.* in
474 Sambrook *et al*. (Sambrook & Russell, 2001). *E. coli* strain OP50 was used as food source for *C.*
475 *elegans* (Stiernagle, 2006). The Bristol isolate N2 was used as the wild-type *C. elegans* strain; *C.*
476 *elegans* strains were propagated on nematode growth medium (NGM) as previously described
477 (Stiernagle, 2006). All the *C. elegans* strains included in this study were obtained from *Caenorhabditis*
478 Genetics Center (CGC; University of Minnesota) and are listed in Table S8.

479 **Cloning and site-directed mutagenesis**

480 The full-length coding sequences for CCTX2, CCTX1 and CCTX3 were amplified from *Coprinopsis*
481 *cinerea* AmutBmut fruiting body (CCTX2, CCTX1) and vegetative mycelium cDNA (CCTX3), using
482 primer pairs CCTX1for_His_NdeI/CCTX1rev_NotI (CCTX1), CCTX2for_NdeI/CCTX2rev_NotI (CCTX2),
483 and CCTX3for_His_NdeI/CCTX3rev_NotI (CCTX3), respectively (Table S9). cDNA synthesis was
484 performed using the Transcriptor Universal cDNA Master kit (Roche) and Phusion DNA polymerase
485 (NEB) as described in Plaza *et al.* (Plaza et al, 2014). The resulting PCR amplicons were cloned into the
486 expression vector pET24b (Novagen) using *Nde*I (Thermo Scientific) and *Not*I (Thermo Scientific)
487 restriction sites. The list of constructs used in this work and their use is reported in Table S7.

488 *Constructs for functional studies*

489 The pET24_CCTX2 plasmid served as the template for the construction of N- and C-terminal
490 truncations of the CCTX2 coding region, including the deletion of the C-terminal HSEL sequon, and for
491 the site-directed mutagenesis of specific residues to alanine. Constructs pET24_CCTX2ΔN (lacking
492 residues 2-303) and pET24_CCTX2ΔC (lacking residues 570-787), both with a N-terminal His_8 tag,
493 were generated using primer pairs CCTX2ΔNfor_His_NdeI/CCTX2rev_NotI and
494 CCTX2for_His_NdeI/CCTX2ΔC_rev_NotI, respectively. The pET24-CCTX2ΔHSEL construct was cloned
495 using primer pair CCTX2for_NdeI/CCTX2_ΔHSEL_rev_NotI (Table S9); the corresponding PCR
496 products were cloned into the pET24b expression vector using *Nde*I and *Not*I restriction sites. Site-
497 directed mutagenesis of pET24-CCTX2 was performed by overlap extension PCR using Phusion DNA
498 polymerase and a combination of CCTX2for_NdeI, CCTX2rev_NotI and mutagenesis primers with the
499 respective single, double and triple mutations. All plasmids were amplified in *E. coli* strain DH5α,
500 purified using mini- and midi-preparation kits (Macherey-Nagel) and verified by Sanger sequencing
501 (Microsynth, Switzerland). Plasmids carrying the correct mutations were then transformed into *E. coli*
502 strain BL21(DE3) for protein production and purification.

503 *Constructs for structural studies*

504 A different expression vector was generated for structural studies, carrying CCTX2 fused to a tobacco
505 etch virus (TEV) protease-cleavable His_8 tag. Starting from the pET24_CCTX2 construct, the *cctx2*
506 gene was subcloned into the pET22b(+) expression vector (Novagen) to generate the pET22b(+)-8H-
507 TEV-CCTX2 construct (Table S7). The cloning was carried out using the NEBuilder HiFi DNA Assembly
508 kit (New England Biolabs), following guidelines from the manufacturer. Cloning primer pairs for the
509 gene (CCTX2_fwd/CCTX2_rev; Table S9) and the vector (pET22_fwd/pET22_rev; Table S9) were
510 designed using the NEBuilder Assembly tool (https://nebuilder.neb.com). PCR was performed using
511 Q5 High-Fidelity DNA Polymerase (New England Biolabs Inc.) following the protocol provided by the
512 manufacturer. Template DNA was degraded using endonuclease *Dpn*I (Thermo Scientific) and
513 assembly of the pET22b(+)-8H-TEV-CCTX2 vector was carried out using the HiFi DNA Assembly
514 Master Mix (New England Biolabs Inc.). Turbo *E. coli* cells (New England Biolabs Inc.) were
515 transformed using the assembly digest and grown o/n at 37°C on LB-agar plates, complemented with
516 100 µg/mL ampicillin. Single colonies were used to inoculate 5 mL LB media cultures supplemented
517 with 100 µg/mL ampicillin; DNA extraction and purification was performed using the Nucleospin
518 Plasmid kit (Macherey-Nagel). Constructs were verified by Sanger sequencing (Eurofins Genomics
519 GmbH, Germany). Clones carrying plasmids with the correct sequence were amplified by o/n growth
520 at 37°C in 200 mL LB cultures, supplemented with 100 µg/mL ampicillin. DNA extraction and
521 purification for long-term storage was performed using the NucleoBond Xtra Midi kit (Macherey-
522 Nagel).

523 **Protein production and purification**

524 *Protein for functional studies*

525 Constructs were transformed into *E. coli* BL21(DE3) for protein production and purification. *E. coli*
526 BL21(DE3) transformants were grown in LB medium to an OD_{600} ≈0.7 at 37°C. Cultures were put on
527 ice for 10 to 15 min before protein production was induced with 0.5 mM IPTG (isopropyl β-D-1-
528 thiogalactopyranoside) for 16 h at 20°C under constant shaking at 120 rpm (Multitron incubator-
529 shaker, Infors HT). All proteins were tested for solubility according to the method described by
530 Künzler *et al*. (Künzler et al, 2010). Bacteria were harvested by centrifugation and the pellets were
531 resuspended in 20 mL of 20 mM HEPES, 500 mM NaCl, pH 7.5 and lysed using a French press
532 (Avantec). Thereafter, the lysates were centrifuged at 16'000 rpm (Sorvall SS34 rotor) for 30 min to
533 remove cell debris. The resulting supernatant, containing the soluble whole protein fraction, was
534 incubated for 1 h at 4°C with Sepharose™ CL 6B (GE Healthcare Life Sciences) beads or Ni-NTA beads
535 (Macherey-Nagel), for untagged and His-tagged proteins, respectively. Proteins were batch eluted
536 from Sepharose™ CL 6B or Ni-NTA beads by gravity flow using lysis buffer supplemented with either
537 200 mM lactose or 250 mM imidazole, respectively. Proteins were desalted on a PD-10 column (GE
538 Healthcare Life Sciences) and the buffer exchanged for 20 mM HEPES and 150 mM NaCl, pH 7.5.

539 MOA, *Coprinopsis cinerea* galectin 2 (CGL2) and *Coprinopsis cinerea* lectin 2 (CCL2) were expressed
540 and purified as described previously (Butschi et al, 2010; Wohlschlager et al, 2011; Schubert et al,
541 2012).

542 *Protein for structural studies*

543 *E. coli* C41 (DE3) cells were transformed with pET22b(+) following a standard protocol, like the one
544 described *e.g.* in Sambrook *et al*. (Sambrook & Russell, 2001). C41 (DE3) was selected among a panel
545 of several *E. coli* strains, as it minimized CCTX2 cleavage by endogenous proteases. Glycerol stocks
546 were prepared from successful transformants, cultured o/d at 37°C in LB supplemented with 100
547 µg/mL ampicillin and stored at -80°C. 2 mL precultures were inoculated with multiple colonies, taken
548 from LB-agar plates streaked with glycerol stocks, and grown o/d at 37°C, shaking at 120 rpm
549 (Multitron incubator-shaker, Infors HT). 50 mL LB medium cultures containing 100 µg/mL ampicillin
550 were inoculated with the 2 mL precultures and grown at 37˚C, 120 rpm for 16 hours. The 50 mL
551 cultures were then used to inoculate 1 L of LB medium complemented with 100 µg/mL ampicillin to
552 an optical density at 600 nm (OD_{600}) of 0.1, and incubated at 37˚C, 120 rpm. When OD_{600} reached ≈
553 0.6, the culture was cooled down in iced water for 10 to 15 minutes. Expression of the gene was
554 induced with 0.5 mM IPTG for 18-20 hours at 20˚C, shaking at 120 rpm (Multitron incubator-shaker,
555 Infors HT).

556 Cells were harvested at 4,000 x *g*, 4˚C for 30 minutes and resuspended in 5 mL of lysis buffer per
557 gram of cell paste. The lysis buffer was prepared to a final concentration of 50 mM HEPES, 200 mM
558 NaCl, 2 mM EDTA, 5 mM DTT; the pH was adjusted to pH 7.5. The buffer was complemented with
559 c*O*mplete EDTA-free protease Inhibitor cocktail (Roche Life Science) to the working concentration
560 recommended by the manufacturer. Cells were lysed using an ultrasound homogenizer (Qsonica
561 Q500) at 20% amplitude for 1 minute (10 second cycles: 3 seconds 'on', 7 seconds 'off'). Cell lysates
562 were clarified at 40,000 x *g*, 4°C for 40 minutes. The supernatants were recovered and directly used
563 for purification.

564 Galactose-affinity chromatography was carried out at RT on an ÄKTA Start (GE Life Sciences), using
565 Immobilized D-galactose gel (Thermo Scientific) packed into a 5 mL Tricorn 10/50 column (GE Life
566 Sciences). The affinity column was equilibrated with loading buffer (20 mM HEPES, 500 mM NaCl, pH
567 7.5), loaded with cell lysate and washed with six column volumes (CV) of loading buffer. Bound
568 protein was eluted with ten CV of elution buffer (20 mM HEPES, 500 mM NaCl, 1 M D-galactose, pH
569 7.5). Fraction content was assessed by SDS-PAGE; fractions containing CCTX2 were pooled and
570 concentrated by ultrafiltration at 3,500 x *g*, 4˚C using Vivaspin® 20 Centrifugal Filter Units 10K
571 Molecular Weight Cut Off (MWCO) (VWR). Size exclusion chromatography (SEC) was performed, to
572 ensure sample purity and monodispersity. A Superdex 200 Increase 10/300 GL (Cytiva) gel-filtration
573 column was equilibrated with SEC buffer (50 mM HEPES, 150 mM NaCl, 100 mM Galactose; pH
574 adjusted to 7.5) and loaded with the concentrated sample from ultrafiltration. The content of eluted

575 fractions was assessed by SDS-PAGE; fractions containing pure CCTX2 were pooled and loaded onto a
576 5 mL HiTrap desalting column (Cytiva), previously equilibrated with dialysis buffer (20 mM HEPES, pH
577 7.5). The dialysed protein was kept at 4˚C until its use for structural studies.

578 **Biotin, Alexa Fluor®488 and TAMRA protein-labeling**

579 For thin-layer chromatography (TLC), glycan array analysis and *C. elegans in vivo* imaging, proteins
580 were purified as described above ('Protein purification for functional studies') and biotinylated with
581 the EZ-Link®-sulfo-NHS-biotin kit (Thermo Scientific, USA). The proteins were labeled following the
582 manufacturer instructions, using a 20-fold molar excess of labeling agent. Protein labeling with
583 Alexa®488 (Thermo Scientific, USA) and carboxytetramethylrhodamine (TAMRA; Thermo Scientific,
584 USA) was performed as described previously (Bleuler-Martínez et al, 2011; Wohlschlager et al, 2011;
585 Stutz et al, 2015).

586 **Preparation and analysis of glycosphingolipids from *C. elegans***

587 Total lipids were extracted, and upper phase lipids containing glycosphingolipids (GSLs) were
588 separated by thin-layer chromatography (TLC) according to Barrows *et al.*(Barrows et al, 2006;
589 Wohlschlager et al, 2011). Binding of separated lipids to CCTX2 was assessed by overlaying the TLC
590 plates with biotinylated CCTX2 and HRP-coupled avidin. The biotinylated GSL-binding chimerolectin
591 MOA (Wohlschlager et al, 2011) and the N-terminally truncated version of CCTX2 (CCTX2ΔN) were
592 used as controls. GSLs were extracted from *C. elegans* wild-type N2 strain as well as from the mutant
593 strains *bre-2, bre-3, bre-4* and *bre-5* and separated using TLC, as described by Barrows *et al.* (Barrows
594 et al, 2006). In brief, L4-staged larvae from each strain were collected and washed three times in 10
595 ml H_20. The resulting pellet was flash-frozen in liquid nitrogen and worm pellets were sonicated to
596 break the worm cuticle. 1 volume of homogenized worm pellet was mixed first with methanol and
597 then chloroform to a final ratio of 4:8:3 (chloroform:methanol:water). Samples were incubated for 2
598 h at 37°C and afterwards centrifuged 5 min at 1400 x *g*. Supernatant was transferred into a new glass
599 vial. 0.147 volumes of water were added to the supernatant. Samples were centrifuged again as
600 above to achieve phase separation. The upper phase (containing the glycolipids) was transferred
601 onto a Sep-Pak cartridge (Millipore) for reversed-phased chromatography. For TLC overlay analysis,
602 150 nM biotinylated MOA, CCTX2, or CCTX2ΔN were used following the protocol described by
603 Wohlschlager *et al.* (Wohlschlager et al, 2011).

604 ***C. elegans* toxicity assay in liquid medium with purified protein**

605 *C. elegans* toxicity assays were carried out as previously described (Künzler et al, 2010; Stutz et al,
606 2015) with the following modifications: Assays were performed in 96-well plates, with approximately
607 30 L1 staged larvae. Larvae were incubated in 100 µl 1x phosphate-buffered saline (PBS; pH 7.4)
608 containing *E. coli* OP50 OD_{600} = 2 and the indicated protein concentrations for 72 h at 20°C. Larval
609 development into L4 stage or adulthood was scored as full development.

610 **Microscopy of *C. elegans***

611 *C. elegans* strains were grown into L4 stage and fed with 150 µg/ml of 5-TAMRA-labeled MOA or
612 CCTX2 or with 500 µg/ml of 5-TAMRA-labeled CCL2 for the indicated time, before washing 3x with 1x
613 PBS to remove excess protein. Worms were immobilized using 10 mM levamisole and mounted on a
614 3% w/v agarose pad. Fluorescence images were taken using a Zeiss LSM 780 upright confocal
615 microscope. Image acquisition was performed using the ZEN system imaging software (Zeiss,
616 Germany). Image analysis was performed using the software Fiji (Schindelin et al, 2012), a
617 distribution of the open-source imaging software ImageJ (Schneider et al, 2012).

618 **Kex2p proteolysis assay**

619 100 µg of CCTX2 were incubated at 37°C overnight (o/n) with 10 µg recombinant Kex2p from
620 *Saccharomyces cerevisiae* (PeproTech). Samples from the reactions were collected, diluted 1:10 and
621 separated by SDS-PAGE. Gel bands were excised and identified by tryptic digest and peptidomics
622 (Functional Genomics Center Zürich). The remainder of the reactions was incubated with
623 Sepharose™ CL 6B or Ni-NTA beads for 1 h at RT on a rotating wheel. Thereafter, the slurry was
624 loaded onto Mobicol spin columns (MoBiTec, Germany) and the flow-through was collected. The
625 beads were washed with 50 column volumes of 20 mM HEPES, 150 mM NaCl, pH 7.5 and boiled in
626 Laemmli buffer. The eluates were subjected to an immunoblot analysis using a polyclonal rabbit
627 antiserum raised against recombinant CCTX2 (Seramun Diagnostica; Fig. S7C).

628 **Sequence analysis, phylogeny and structure prediction**

629 Domain predictions were performed using the SMART software (Simple Modular Architecture
630 Research Tool, biobyte solutions GmbH, Germany; http://smart.embl-heidelberg.de). CLUSTALW
631 amino acid sequence alignments were performed using the Pole Bio-Informatique Lyonnais (PBIL)-
632 based software NPS@CLUSTALW (Combet et al, 2000) (Network Protein Sequence @nalysis, Lyon,
633 France; https://npsa-prabi.ibcp.fr/NPSA/npsa_clustalw.html). The multiple sequence alignment of
634 the three *C. cinerea* paralogues CCTX1, CCTX2 and CCTX3, as well as their fungal full-length
635 homologues, was done using the MUltiple Sequence Comparison by Log-Expectation (MUSCLE)
636 algorithm (Edgar, 2004). The alignment was visualized in Unipro UGENE software (Okonechnikov et
637 al, 2012) and manually trimmed (Fig. S1D). The phylogenetic tree was constructed using the RAxML-
638 ng algorithm (Kozlov et al, 2019) which implements maximum-likelihood (ML) optimality criterion,
639 setting 500 bootstrap replicates under a Jones-Taylor-Thornton matrix. The phylogenetic tree with
640 the highest bootstrap support was visualized (Fig. 1B) using ITOL (Letunic & Bork, 2024). *In silico*
641 structure prediction for CCTX2 was initially carried out using the Robetta server (Kim et al, 2004)
642 (https://robetta.bakerlab.org), based on the Rosetta simulation framework (Baek et al, 2021),
643 applying the 'comparative modeling' protocol and automatic template selection (Song et al, 2013b).
644 Later on, a model was generated using AlphaFold 3 (Abramson et al, 2024). Homology models of

645 CCTX1 and CCTX3, used for domain boundary prediction, were generated using the Robetta server
646 (Kim et al, 2004) (https://robetta.bakerlab.org), applying the 'comparative modeling' protocol and
647 uploading the cryo-EM structure of CCTX2 as a template.

648 **Cryo-EM grid preparation and data collection**

649 Cryo-EM grids were prepared at the SciLife lab cryo-EM Swedish National facility, Umeå node
650 (Sweden) according to standard procedures described in *e.g.* Passmore and Russo (Passmore &
651 Russo, 2016). Quantifoil R2/1 300 mesh carbon grids (Quantifoil Micro Tools, Germany) were glow-
652 discharged at 15 µA for 30 s with a Pelco easiGlow glow-discharge cleaning system (Ted Pella, Inc.). 4
653 µL of 1 mg/mL CCTX2 were applied to the grids using a Vitrobot Mark IV (Thermo Scientific)
654 equilibrated at 4°C, 100% humidity. 7,500 movies were collected on a FEI Titan Krios (Thermo
655 Scientific) operating at 300 kV and equipped with a Gatan K2 camera, at a nominal magnification of
656 215,000 x and with a physical pixel size of 0.63 Å. The dose rate was set to 4.7 e^-/px/s, and the total
657 exposure time was 5 s, resulting in a total dose of 59.5 $e^-/Å^2$. Nominal defocus range was -1.5 µm to -
658 3.0 µm in 0.3 µm steps. Data collection statistics are reported in Table S2.

659 **Image processing and model building**

660 The data set was processed using *cryoSPARC* 3 (Punjani et al, 2017) as described in Fig. S4. In brief,
661 patch motion correction and contrast transfer function (CTF) estimation was performed. After
662 manual particle picking, homogenous refinement was carried out to generate a volume for template-
663 based particle picking, generating a stack of 1.8 million particles. After three rounds of 2D
664 classification, 641,000 particles remained with a final extraction at a box size of 480 pixels and
665 cropping to 256 pixels. These particles were used for *ab initio* reconstruction, where the best class
666 was chosen for heterogeneous refinement. 204,000 particles were selected from two rounds of
667 heterogeneous refinement. The volume was subjected to local refinement and sharpening to
668 increase the quality of the map, which extended to 3.2 Å resolution.

669 Predictive models were generated using the Robetta server (Kim et al, 2004) for all the CCTX2
670 domains, using either template-based (BTF domains 1 to 4) or with *ab initio* (C-terminal domain)
671 protocols. Directly fitting the models into the reconstructed volume proved impossible, due to the
672 strong structural homology among the BTF domains (plus the presence of a pseudo-threefold axis)
673 and the poor quality of the C-terminal domain prediction; automated protein chain building was not
674 successful either. However, a polyalanine trace, manually built using *Coot* (Casañal et al, 2020), a
675 component of the *CCP4* software suite (Agirre et al, 2023), revealed the correct domain topology.
676 The first three BTF domains were easily docked onto their respective polyAla trace, whereas the
677 fourth domain had to be manually aligned in PyMOL (Schrödinger, Inc.), only relying on the backbone
678 segments built into the discontinuous Coulomb potential density (Fig. S3A). Poorly predicted
679 segments were removed from the model of the C-terminal domain, keeping only a core region

680 (residues 619-740; Figs. S3B and S3C). The CCTX2 model was completed using *Coot*, by accurately
681 fitting the models into the cryo-EM volume and building missing residues; an *in silico* model
682 generated using AlphaFold 3 (Abramson et al, 2024) was used to improve areas with poor density fit.
683 Refinement was carried out by alternating cycles of real-space building using *Coot* (Casañal et al,
684 2020) and real-space refinement with *phenix.refine*, a component of the *Phenix* software suite for
685 structural biology (Liebschner et al, 2019). The final model includes the entire CCTX2 chain (residues
686 1-787); the C-terminal amino acid residues from the His_8-TEV tag were poorly defined and were not
687 modeled. Similarly, galactose molecules at the putative sugar-binding sites of the first (subdomains β
688 and γ) and second (subdomain α) BTF domains were not modeled due to poor fit density. Refinement
689 statistics are reported in Table S2: although the model displays a remarkably high number of
690 Ramachandran outliers (8, or about the 1%, evaluated with RAMPAGE (Lovell et al, 2003)), their
691 geometry is generally close to the allowed regions of the plot. All structural illustrations were
692 prepared using PyMOL (Schrödinger, Inc.).

693

694 **ACKNOWLEDGEMENTS**

695 We would like to thank Ken Teter for comments on the manuscript. Cryo-EM experiments were
696 performed at SciLife lab (Umeå, Sweden), many thanks to Michael Hall for collecting the cryo-EM
697 data. The structural biology work was carried out at the UiO Structural Biology core facilities, which
698 are part of the Norwegian Macromolecular Crystallography Consortium (NORCRYST) and received
699 funding from the Norwegian INFRASTRUKTUR-program (project no. 245828) as well as from UiO
700 (core facility funds). The genetics, biochemistry and cell biology work was carried out at ETH Zürich
701 and funded by the Swiss National Science Foundation (Grants no. 31003A-130671 and 310030-
702 212435). The PhD position of F.K. was financed by UiO.

703

704 **AUTHOR CONTRIBUTIONS**

705 M.K. conceived the study, and G.C., S.S.S. and U.K. additionally contributed to designing the
706 experiments. S.S.S., K.B. and D.P. produced and purified the recombinant CCTX1-3 variants and
707 performed the toxicity and cell biology assays, B.S. prepared the sequence alignments and
708 phylogenetic tree, under the supervision of M.K.. C.H.K. and F.K. produced and purified recombinant
709 CCTX2 for structural biology studies, under the supervision of G.C. and U.K.. F.K. and G.C. analyzed
710 the cryo-EM data, in collaboration with and supervised by A.A.A., A.B., J.L.K. and T.B.. The manuscript
711 was initially written by S.S.S. and M.K., substantially revised by G.C., F.K. and U.K., and finalized by all
712 authors.

713

714 **ABBREVIATIONS LIST**

715 BSA, bovine serum albumin; BTF, β-trefoil fold; CCL2, *Coprinopsis cinerea* lectin 2; CCTX1, *Coprinopsis*
716 *cinerea* toxin 1; CCTX2, *Coprinopsis cinerea* toxin 2; CCTX3, *Coprinopsis cinerea* toxin 3; CGL1,
717 *Coprinopsis cinerea* galectin 1; CGL2, *Coprinopsis cinerea* galectin 2; cryo-EM, cryogenic electron
718 microscopy; CV, column volumes; DUF, Domain of Unknown Function; EDTA,
719 ethylenediaminetetraacetic acid; ER, endoplasmic reticulum; GSL, glycosphingolipid; HEPES, 4-(2-
720 hydroxyethyl)piperazine-1-ethane-sulfonic acid; IPTG, isopropyl β-D-1-thiogalactopyranoside; Kex-2,
721 kexin 2 from *Saccharomyces cerevisiae*; LB, lysogeny broth; MANS, α-Mannosidase II; MOA,
722 *Marasmius oreades* agglutinin; MS, mass spectrometry; MWCO, molecular weight cut-off: NAD,
723 nicotinamide adenine dinucleotide; NGM, nematode growth medium; o/d, over day; o/n, overnight;
724 OD, optical density; PARP, poly(ADP-ribose) polymerase; PBS, phosphate-buffered saline; PDB,
725 Protein Data Bank; rpm, rotation/round per minute; TAMRA, tetramethylrhodamine; TLC, thin-layer
726 chromatography; TGN, trans-Golgi network

727

728 **REFERENCES**

729

730 Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ,
731 Bambrick J et al (2024) Accurate structure prediction of biomolecular interactions with
732 AlphaFold 3. Nature 630: 493-500

733 Agirre J, Atanasova M, Bagdonas H, Ballard CB, Baslé A, Beilsten-Edmands J, Borges RJ, Brown DG,
734 Burgos-Mármol JJ, Berrisford JM et al (2023) The *CCP*4 suite: integrative software for
735 macromolecular crystallography. Acta Crystallogr D Struct Biol 79: 449-461

736 Aravind L, Anantharaman V, Zhang D, de Souza RF, Iyer LM (2012) Gene flow and biological conflict
737 systems in the origin and evolution of eukaryotes. Front Cell Infect Microbiol 2: 89

738 Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN,
739 Schaeffer RD et al (2021) Accurate prediction of protein structures and interactions using a
740 three-track neural network. Science 373: 871-876

741 Barrows BD, Griffitts JS, Aroian RV (2006) *Caenorhabditis elegans* carbohydrates in bacterial toxin
742 resistance. In: Functional Glycomics, Fukuda M. (ed.) pp. 340-358. Academic Press: San Diego,
743 California, United States of America

744 Bleuler-Martínez S, Butschi A, Garbani M, Wälti MA, Wohlschlager T, Potthoff E, Sabotiĉ J, Pohleven J,
745 Lüthy P, Hengartner MO et al (2011) A lectin-mediated resistance of higher fungi against
746 predators and parasites. Mol Ecol 20: 3056-3070

747 Bleuler-Martinez S, Stutz K, Sieber R, Collot M, Mallet J-M, Hengartner M, Schubert M, Varrot A,
748 Künzler M (2017) Dimerization of the fungal defense lectin CCL2 is essential for its toxicity
749 against nematodes. Glycobiology 27: 486-500

750 Boulianne RP, Liu Y, Aebi M, Lu BC, Kües U (2000) Fruiting body development in Coprinus cinereus:
751 regulated expression of two galectins secreted by a non-classical pathway. Microbiology 146:
752 1841-1853

753 Bucci C, Parton RG, Mather IH, Stunnenberg H, Simons K, Hoflack B, Zerial M (1992) The small GTPase
754 rab5 functions as a regulatory factor in the early endocytic pathway. Cell 70: 715-728

755 Butschi A, Titz A, Wälti MA, Olieric V, Paschinger K, Nöbauer K, Guo X, Seeberger PH, Wilson IBH, Aebi
756 M et al (2010) *Caenorhabditis elegans N*-glycan core ß-galactoside confers sensitivity towards
757 nematotoxic fungal galectin CGL2. PLoS Pathog 6: e1000717

758 Casañal A, Lohkamp B, Emsley P (2020) Current developments in *Coot* for macromolecular model
759 building of electron cryo-microscopy and crystallographic data. Protein Sci 29: 1069-1078

760 Cho JA, Chinnapen DJ-F, Aamar E, te Welscher YM, Lencer WI, Massol R (2012) Insights on the
761 trafficking and retro-translocation of glycosphingolipid-binding bacterial toxins. Front Cell
762 Infect Microbiol 2: 51

763 Clemens JD, Nair GB, Ahmed T, Qadri F, Holmgren J (2017) Cholera. Cholera 390: 1539-1549

764 Combet C, Blanchet C, Geourjon C, Deléage G (2000) *NPS@*: network protein sequence analysis.
765 Trends Biochem Sci 25: 147-150

766 Cordara G, Egge-Jacobsen W, Johansen HT, Winter HC, Goldstein IJ, Sandvig K, Krengel U (2011)
767 *Marasmius oreades* agglutinin (MOA) is a chimerolectin with proteolytic activity. Biochem
768 Biophys Res Commun 408: 405-410

769 Cordara G, Manna D, Krengel U (2017) Family of papain-like fungal chimerolectins with distinct Ca^{2+}-
770 dependent activation mechanism. Biochemistry 56: 4689-4700

771 Cordara G, van Eerde A, Grahn EM, Winter HC, Goldstein IJ, Krengel U (2016) An unusual member of
772 the papain superfamily: mapping the catalytic cleft of the *Marasmius oreades* agglutinin (MOA)
773 with a caspase inhibitor. PloS one 11: e0149407

774 Cordara G, Winter HC, Goldstein IJ, Krengel U, Sandvig K (2014) The fungal chimerolectin MOA
775 inhibits protein and DNA synthesis in NIH/3T3 cells and may induce BAX-mediated apoptosis.
776 Biochem Biophys Res Commun 447: 586-589

777 Edgar RC (2004) *MUSCLE*: a multiple sequence alignment method with reduced time and space
778 complexity. BMC Bioinformatics 5: 113

779 Gams W, Meyer W (1998) What exactly is *Trichoderma harzianum*? Mycologia 90: 904-915

780 Garred Ø, van Deurs B, Sandvig K (1995) Furin-induced cleavage and activation of Shiga toxin. J Biol
781 Chem 270: 10817-10821

782 Grahn E, Askarieh G, Holmner A, Tateno H, Winter HC, Goldstein IJ, Krengel U (2007) Crystal structure
783 of the *Marasmius oreades* mushroom lectin in complex with a xenotransplantation epitope. J
784 Mol Biol 369: 710-721

785 Grahn EM, Winter HC, Tateno H, Goldstein IJ, Krengel U (2009) Structural characterization of a lectin
786 from the mushroom *Marasmius oreades* in complex with the blood group B trisaccharide and
787 calcium. J Mol Biol 390: 457-466

788 Griffitts JS, Haslam SM, Yang T, Garczynski SF, Mulloy B, Morris H, Cremer PS, Dell A, Adang MJ,
789 Aroian RV (2005) Glycolipids as receptors for *Bacillus thuringiensis* crystal toxin. Science 307:
790 922-925

791 Hoekstra D, Tyteca D, van IJzendoorn SCD (2004) The subapical compartment: a traffic center in
792 membrane polarity development. J Cell Sci 117: 2183-2192

793 Holm L (2022) *Dali* server: structural unification of protein families. Nucleic Acids Res 50: W210-
794 W215

795 Inoue K, Sobhany M, Transue TR, Oguma K, Pedersen LC, Negishi M (2003) Structural analysis by X-
796 ray crystallography and calorimetry of a haemagglutinin component (HA1) of the progenitor
797 toxin from *Clostridium botulinum*. Microbiology 149: 3361-3370

798 Jing J, Prekeris R (2009) Polarized endocytic transport: the roles of Rab11 and Rab11-FIPs in
799 regulating cell polarity. Histol Histopathol 24: 1171-1180

800 Juillot S, Cott C, Madl J, Claudinon J, van der Velden NSJ, Künzler M, Thuenauer R, Römer W (2016)
801 Uptake of *Marasmius oreades* agglutinin disrupts integrin-dependent cell adhesion. Biochim
802 Biophys Acta 1860: 392-401

803 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek
804 A, Potapenko A et al (2021) Highly accurate protein structure prediction with AlphaFold.
805 Nature 596: 583-589

806 Kadirvelraj R, Grant OC, Goldstein IJ, Winter HC, Tateno H, Fadda E, Woods RJ (2011) Structure and
807 binding analysis of *Polyporus squamosus* lectin in complex with the Neu5Acα2-6Galβ1-4GlcNAc
808 human-type influenza receptor. Glycobiology 21: 973-984

809 Kempken F (2011) Fungal defences against animal antagonists - lectins & more. Mol Ecol 20: 2876-
810 2877

811 Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the *Robetta* server.
812 Nucleic Acids Res 32: W526-W531

813 Kombrink A, Tayyrov A, Essig A, Stöckli M, Micheller S, Hintze J, van Heuvel Y, Dürig N, Lin C-W, Kallio
814 PT et al (2019) Induction of antibacterial proteins and peptides in the coprophilous mushroom
815 *Coprinopsis cinerea* in response to bacteria. ISME J 13: 588-602

816 Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) *RAxML-NG*: a fast, scalable and user-
817 friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35: 4453-4455

818 Künzler M (2015) Hitting the sweet spot: glycans as targets of fungal defense effector proteins.
819 Molecules 20: 8144-8167

820 Künzler M, Bleuler-Martinez S, Butschi A, Garbani M, Lüthy P, Hengartner MO, Aebi M (2010)
821 Biotoxicity assays for gruiting body lectins and other cytoplasmic proteins. In: Glycobiology,
822 Fukuda M. (ed.) pp. 141-150. Academic Press: San Diego, California, United States of America

823 Lencer WI, Saslowsky D (2005) Raft trafficking of AB_5 subunit bacterial toxins. Biochim Biophys Acta
824 1746: 314-321

825 Letunic I, Bork P (2024) Interactive *Tree of Life* (*iTOL*) v6: recent updates to the phylogenetic tree
826 display and annotation tool. Nucleic Acids Res 52: W78-W82

827 Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung L-W, Jain S, McCoy
828 AJ et al (2019) Macromolecular structure determination using X-rays, neutrons and electrons:
829 recent developments in *Phenix*. Acta Crystallogr D Struct Biol 75: 861-877

830 Lovell SC, Davis IW, Arendall WBI, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson
831 DC (2003) Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. Proteins 50: 437-450

832 Luo L, Hannemann M, Koenig S, Hegermann J, Ailion M, Cho M-K, Sasidharan N, Zweckstetter M,
833 Rensing SA, Eimer S (2011) The *Caenorhabditis elegans* GARP complex contains the conserved
834 Vps51 subunit and is required to maintain lysosomal morphology. Mol Biol Cell 22: 2564-2578

835 Mancheño JM, Tateno H, Goldstein IJ, Martínez-Ripoll M, Hermoso JA (2005) Structural analysis of
836 the *Laetiporus sulphureus* hemolytic pore-forming lectin in complex with sugars. J Biol Chem
837 280: 17251-17259

838 Manna D, Cordara G, Krengel U (2020) Crystal structure of MOA in complex with a peptide fragment:
839 a protease caught *in flagranti*. Curr Res Struct Biol 2: 56-67

840 Manna D, Pust S, Torgersen ML, Cordara G, Künzler M, Krengel U, Sandvig K (2017) *Polyporus*
841 *squamosus* lectin 1a (PSL1a) exhibits cytotoxicity in mammalian cells by disruption of focal
842 adhesions, inhibition of protein synthesis and induction of apoptosis. PloS one 12: e0170716

843 Melton-Celsa AR (2014) Shiga toxin (Stx) classification, structure, and function. Microbiol Spectr 2:
844 EHEC-0024-2013

845 Moore RC, Purugganan MD (2003) The early stages of duplicate gene evolution. Proc Natl Acad Sci
846 USA 100: 15682-15687

847 Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, Sakai K, Shimizu M, Nakahori K,
848 Sakamoto Y et al (2015) Strand-specific RNA-Seq analyses of fruiting body development in
849 *Coprinopsis cinerea*. PloS one 10: e0141586

850 Nowakowska-Gołacka J, Sominka H, Sowa-Rogozińska N, Słomińska-Wojewódzka M (2019) Toxins
851 utilize the endoplasmic reticulum-associated protein degradation pathway in their intoxication
852 process. Int J Mol Sci 20: 1307

853 Oda T, Hirabayashi H, Shikauchi G, Takamura R, Hiraga K, Minami H, Hashimoto H, Yamamoto M,
854 Wakabayashi K, Shimizu T et al (2017) Structural basis of autoinhibition and activation of the
855 DNA-targeting ADP-ribosyltransferase pierisin-1. J Biol Chem 292: 15445-15455

856 Okonechnikov K, Golosova O, Fursov M, team U (2012) *Unipro UGENE*: a unified bioinformatics
857 toolkit. Bioinformatics 28: 1166-1167

858 Pace NJ, Weerapana E (2014) Zinc-binding cysteines: diverse functions and structural motifs.
859 Biomolecules 4: 419-434

860 Passmore LA, Russo CJ (2016) Specimen preparation for high-resolution cryo-EM. In: The resolution
861 revolution: recent advances in cryo-EM, Crowther R.A. (ed.) pp. 51-86. Academic Press:
862 Cambridge, Massachusetts, United States of America

863 Peumans WJ, Van Damme EJM (1995) Lectins as plant defense proteins. Plant Physiol 109: 347-352

864 Peumans WJ, Van Damme EJM, Barre A, Rouge P (2001) Classification of plant lectins in families of
865 structurally and evolutionary related proteins. Adv Exp Med Biol 491: 27-54

866 Pickett CL, Whitehouse CA (1999) The cytolethal distending toxin family. Trends Microbiol 7: 292-297

867 Plaza DF, Lin C-W, van der Velden NS, Aebi M, Künzler M (2014) Comparative transcriptomics of the
868 model mushroom *Coprinopsis cinerea* reveals tissue-specific armories and a conserved circuitry
869 for sexual development. BMC genomics 15: 492

870 Plaza DF, Schmieder SS, Lipzen A, Lindquist E, Künzler M (2016) Identification of a novel nematotoxic
871 protein by challenging the model mushroom *Coprinopsis cinerea* with a fungivorous nematode.
872 G3 6: 87-98

873 Polito L, Bortolotti M, Battelli MG, Calafato G, Bolognesi A (2019) Ricin: an ancient story for a
874 timeless plant toxin. Toxins 11: 324

875 Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA (2017) *cryoSPARC*: algorithms for rapid unsupervised
876 cryo-EM structure determination. Nat Methods 14: 290-296

877 Ruck A, Attonito J, Garces KT, Nuñez L, Palmisano NJ, Rubel Z, Bai Z, Nguyen KCQ, Sun L, Grant BD et
878 al (2011) The Atg6/Vps30/Beclin 1 ortholog BEC-1 mediates endocytic retrograde transport in
879 addition to autophagy in *C. elegans*. Autophagy 7: 386-400

880 Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor
881 Laboratory Press, Cold Spring Harbor, N.Y.

882 Sánchez J, Holmgren J (2011) Cholera toxin - A foe & a friend. Indian J Med Res 133: 153-163

883 Sandvig K, Bergan J, Kavaliauskiene S, Skotland T (2014) Lipid requirements for entry of protein toxins
884 into cells. Prog Lipid Res 54: 1-13

885 Sandvig K, Skotland T, van Deurs B, Klokk TI (2013) Retrograde transport of protein toxins through
886 the Golgi apparatus. Histochem Cell Biol 140: 317-326

887 Sandvig K, Torgersen ML, Engedal N, Skotland T, Iversen T-G (2010) Protein toxins from plants and
888 bacteria: probes for intracellular transport and tools in medicine. FEBS Lett 584: 2626-2634

889 Sandvig K, van Deurs B (2005) Delivery into cells: lessons learned from plant and bacterial toxins.
890 Gene Ther 12: 865-872

891 Sato K, Norris A, Sato M, Grant BD (2014) *C. elegans* as a model for membrane traffic. In: WormBook,
892 Labouesse M., Moerman D.G. (eds.) pp. 1-47. The *C. elegans* Research Community: Pasadena
893 (CA)

894 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C,
895 Saalfeld S, Schmid B et al (2012) *Fiji*: an open-source platform for biological-image analysis. Nat
896 Methods 9: 676-682

897 Schneider CA, Rasband WS, Eliceiri KW (2012) *NIH Image* to *ImageJ*: 25 years of image analysis. Nat
898 Methods 9: 671-675

899 Schubert M, Bleuler-Martinez S, Butschi A, Wälti MA, Egloff P, Stutz K, Yan S, Collot M, Mallet JM,
900 Wilson IB et al (2012) Plasticity of the ß-trefoil protein fold in the recognition and control of
901 invertebrate predators and parasites by a fungal defence system. PLoS Pathog 8: e1002706

902 Shah N, Kuntz DA, Rose DR (2008) Golgi α-mannosidase II cleaves two sugars sequentially in the same
903 catalytic site. Proc Natl Acad Sci USA 105: 9570-9575

904 Singh RS, Bhari R, Kaur HP (2010) Mushroom lectins: current status and future perspectives. Crit Rev
905 Biotechnol 30: 99-126

906 Song J (2022) Bacterial AB toxins and host-microbe interactions. In: Poole R.K., Kelly D.J. (eds.) pp. 67-
907 109. Academic Press:

908 Song J, Gao X, Galán JE (2013a) Structure and function of the *Salmonella* Typhi chimaeric A_2B_5
909 typhoid toxin. Nature 499: 350-354

910 Song Y, DiMaio F, Wang RY-R, Kim D, Miles C, Brunette TJ, Thompson J, Baker D (2013b) High-
911 resolution comparative modeling with *RosettaCM*. Structure 21: 1735-1742

912 Stiernagle T (2006) Maintenance of *C. elegans*. In: WormBook, Fay D., Ambros V. (eds.)The *C. elegans*
913 Research Community:

914 Stutz K, Kaech A, Aebi M, Künzler M, Hengartner MO (2015) Disruption of the *C. elegans* Intestinal
915 brush border by the fungal lectin CCL2 phenocopies dietary lectin toxicity in mammals. PloS
916 one 10: e0129381

917 Takacs-Vellai K, Vellai T, Puoti A, Passannante M, Wicky C, Streit A, Kovacs AL, Müller F (2005)
918 Inactivation of the autophagy gene bec-1 triggers apoptotic cell death in *C. elegans*. Curr Biol
919 15: 1513-1517

920 Tayyrov A, Schmieder SS, Bleuler-Martinez S, Plaza DF, Künzler M (2018) Toxicity of potential fungal
921 defense proteins towards the fungivorous nematodes *Aphelenchus avenae* and
922 *Bursaphelenchus okinawaensis*. Appl Environ Microbiol 84: e02051-02018

923 Teter K (2013) Toxin instability and its role in toxin translocation from the endoplasmic reticulum to
924 the cytosol. Biomolecules 3: 997-1029

925 Treiber N, Reinert DJ, Carpusca I, Aktories K, Schulz GE (2008) Structure and mode of action of a
926 mosquitocidal holotoxin. J Mol Biol 381: 150-159

927 Vanden Broeck D, Horvath C, De Wolf MJS (2007) *Vibrio cholerae*: cholera toxin. Int J Biochem Cell
928 Biol 39: 1771-1775

929 Varrot A, Basheer SM, Imberty A (2013) Fungal lectins: structure, function and potential applications.
930 Curr Opin Struct Biol 23: 678-685

931 Wohlschlager T, Butschi A, Zurfluh K, Vonesch SC, Auf dem Keller U, Gehrig P, Bleuler-Martinez S,
932 Hengartner MO, Aebi M, Künzler M (2011) Nematotoxicity of *Marasmius oreades* agglutinin
933 (MOA) depends on glycolipid binding and cysteine protease activity. J Biol Chem 286: 30337-
934 30343

935 Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D (2020) Improved protein structure
936 prediction using predicted interresidue orientations. Proc Natl Acad Sci USA 117: 1496-1503

937 Yao J, Nellas RB, Glover MM, Shen T (2011) Stability and sugar recognition ability of ricin-like
938 carbohydrate binding domains. Biochemistry 50: 4097-4104

939 Zhang D, de Souza RF, Anantharaman V, Iyer LM, Aravind L (2012) Polymorphic toxin systems:
940 comprehensive characterization of trafficking modes, processing, mechanisms of action,
941 immunity and ecology using comparative genomics. Biol Direct 7: 18

942

Main Figures

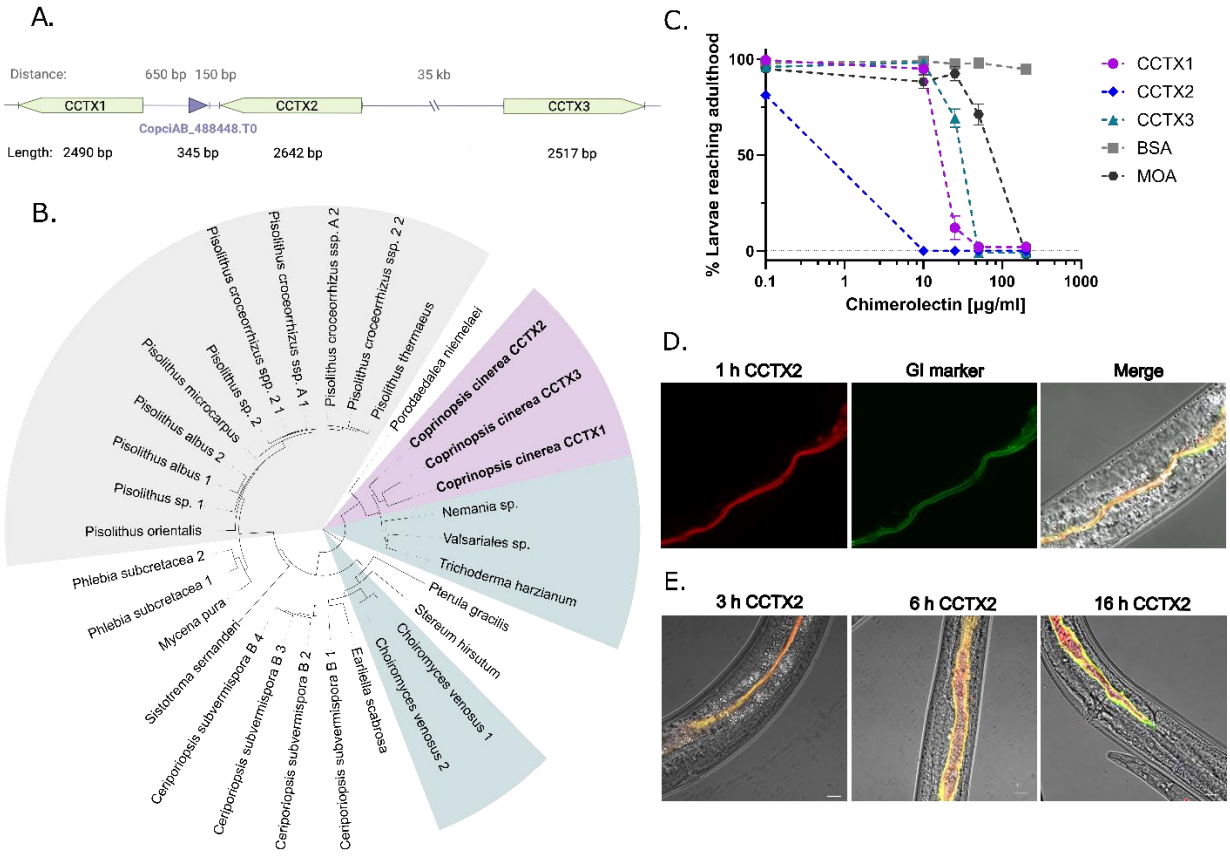

Figure 1. The nematotoxin CCTX2 is a member of a large family of fungal chimerolectins. A. Location of CCTX paralogs in the *C. cinerea* genome (*Coprinopsis cinerea* AmutBmut v.2 annotation from JGI). **B.** Phylogenetic analysis of full-length homologues of CCTX2. Paralogues in *C. cinerea* are highlighted in pink, homologues in *Pisolithus* spp. in gray and homologues in Ascomycetes in turquoise. See Table S1 for the JGI protein IDs of all displayed homologues and Fig. S1A for a sequence alignment of homologues from the monophyletic clade including CCTX1-3. **C.** Nematotoxicity assays of the paralogous CCTX1, CCTX2 and CCTX3. Development of *C. elegans* N2 larvae was arrested in L1 larvae in the presence of either of the three proteins at concentrations >10 µg/ml. The MOA chimerolectin was included as positive control, BSA served as negative control. Data points with error bars indicate means of N = 4 biological replicates with standard error of the mean (SEM). Overlapping data points were nudged by +/- 2 data units in Y for better readability. **D** and **E.** Localization of CCTX2-TAMRA upon feeding to L4 larvae of *C. elegans* strain GK70. This strain expresses PGP-1-GFP as intestinal apical membrane marker. Images in panel D were taken 1 h after feeding with CCTX2-TAMRA. **E.** Time course of intoxification by CCTX2. After 3h, feeding of CCTX2-TAMRA leads to aberrant dilatation of the intestinal lumen. Pictures were taken at the rear end of the animal. Scale bar = 10 µm.

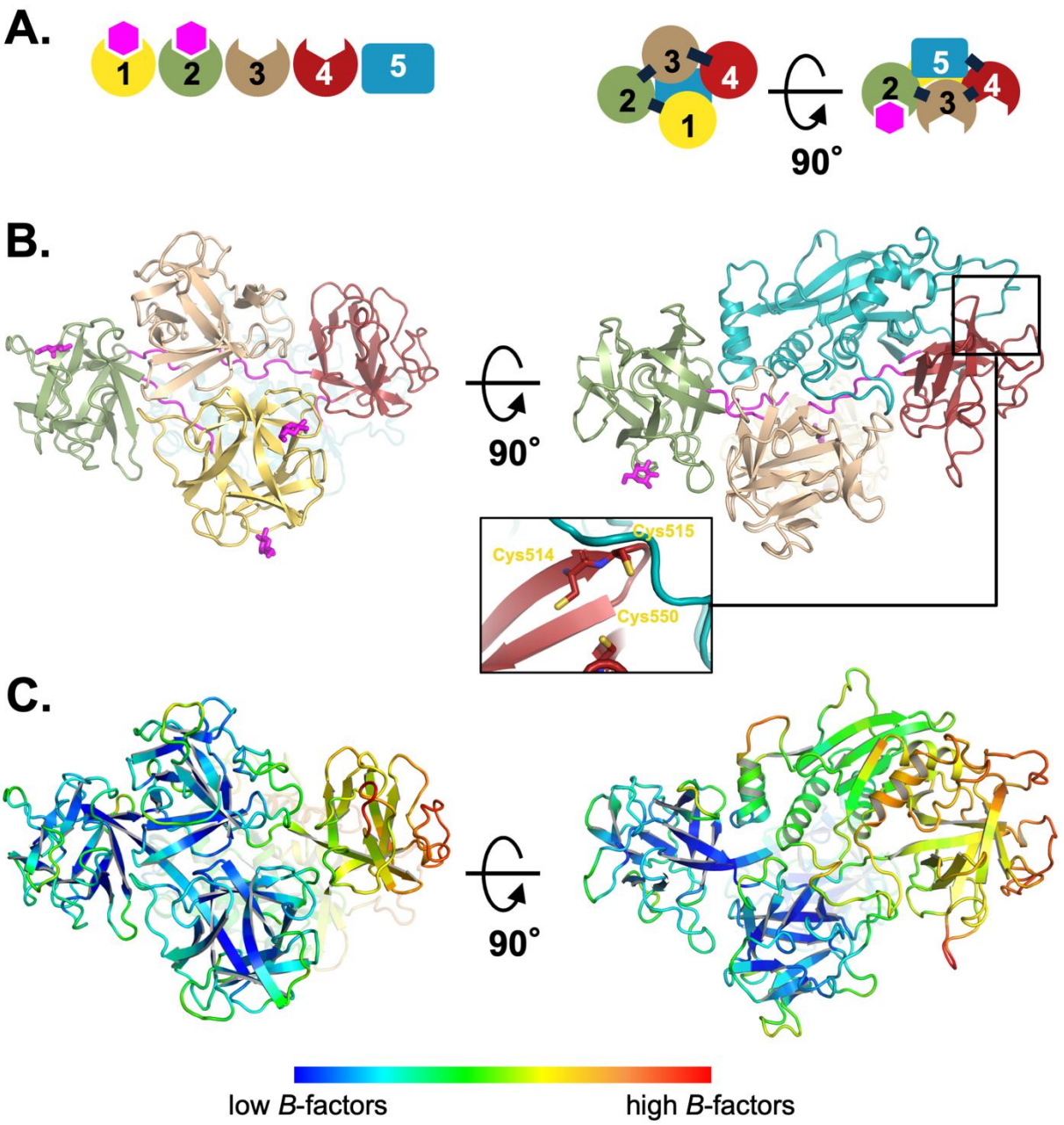

Figure 2. Cryo-EM structure of CCTX2. A. Schematic diagram of the domains of CCTX2 and their topology in the protein tertiary structure. Each domain is represented by a different color to be easily located in the folded protein: yellow (β-trefoil domain/BTF domain 1), green (BTF domain 2), wheat (BTF domain 3), red (BTF domain 4) and teal (C-terminal domain of unknown function; DUF). The hexagons in magenta represent sugars observed bound to the BTF domains with proposed lectin activity (see also panel B and Fig. S5). **B**. Cartoon representation of the CCTX2 structure. The domains follow the color code assigned in A, with linkers between domains colored in magenta. D-galactose molecules (see Fig. S5)—likely remnants from the purification procedure—are depicted as magenta-colored sticks. They represent the Coulomb potential density at the β and γ sugar binding sites of domain 1 and the α sugar-binding site of domain 2. The inset shows the position of a putative zinc-binding site **C.** Schematic diagram of CCTX2, colored by *B*-factors of Cα atoms, according to the color chart on the bottom.

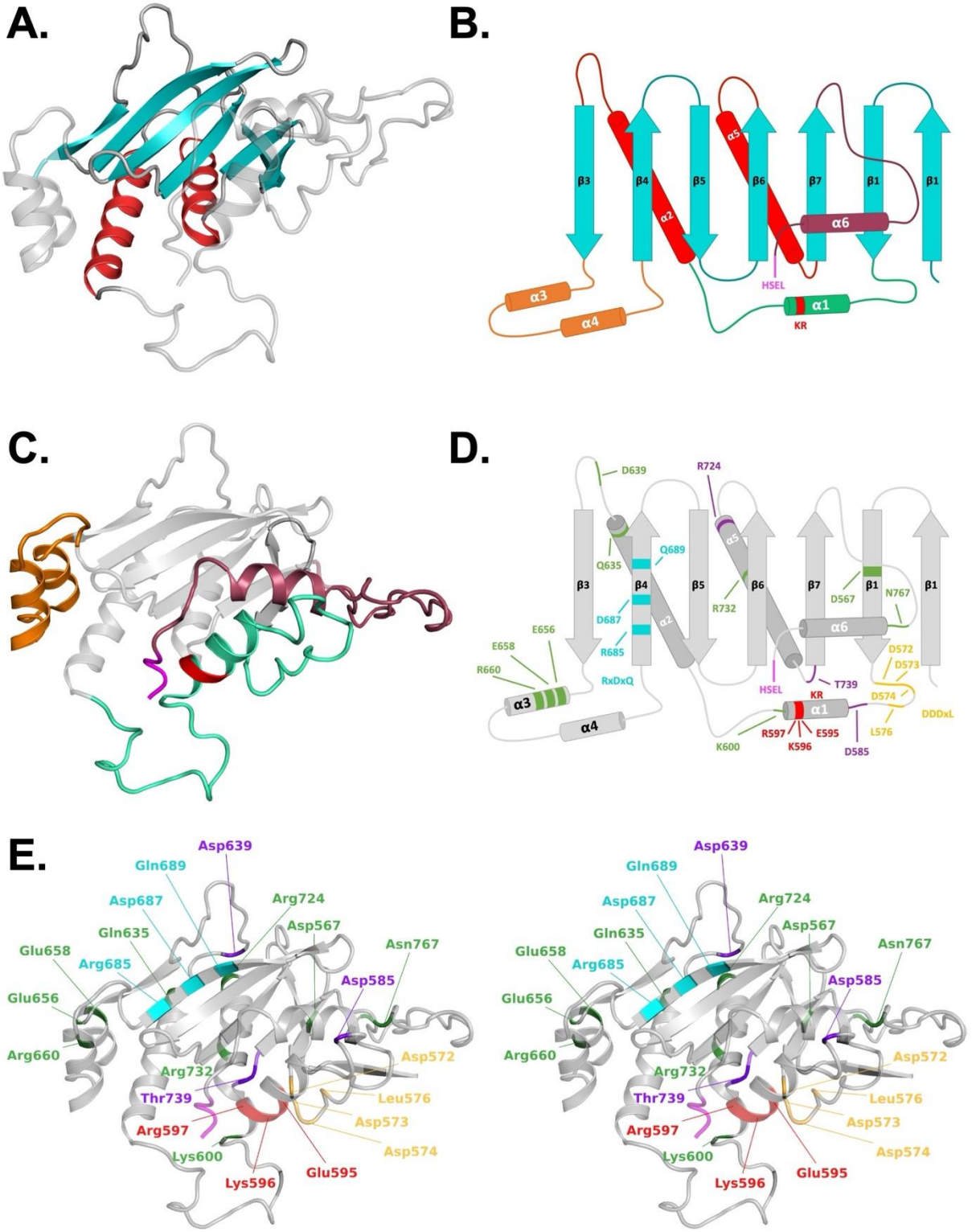

Figure 3. CryoEM structure of CCTX2 C-terminal Domain of Unknown Function (DUF). A. Cartoon representation of the C-cryo-EM structure of the C-terminal DUF. The core elements of the fold, the central seven-stranded β-sheet and the two supporting α-helices are indicated in teal and red, respectively. **B.** Topology diagram of the C-terminal DUF indicating insertions and conserved sequons. The KR sequon for kexin cleavage (residues 596-597) in the first insertion (residues 571-619, light green) is highlighted in red. The two-helix insertion (residues 656-683) is shown in orange. The C-terminus (residues 746-783) is colored dark red, with the HSEL sequon (residues 784-787) highlighted in magenta. **C**. Cartoon representation of the DUF fold indicating the same insertions and conserved sequons as in panel B. **D.** Topology diagram of

the C-terminal DUF reporting the position of amino acids targeted by site mutagenesis. **E.** Stereo cartoon representation of the C-terminal DUF, with residues targeted by Ala-scanning substitutions (green) mapped onto the cryo-EM structure. Positions corresponding to the key functional motifs KR (Glu595, Lys596, Arg597; red), RxDxQ (Arg685, Asp687, Gln689; cyan), DDDxL (Asp572, Asp573, Asp574, Leu576; yellow), HSEL (C-terminal residues 784-787; magenta) and three variants significantly affecting toxicity (Asp585, Arg724, Thr739; purple) have been marked in different colors.

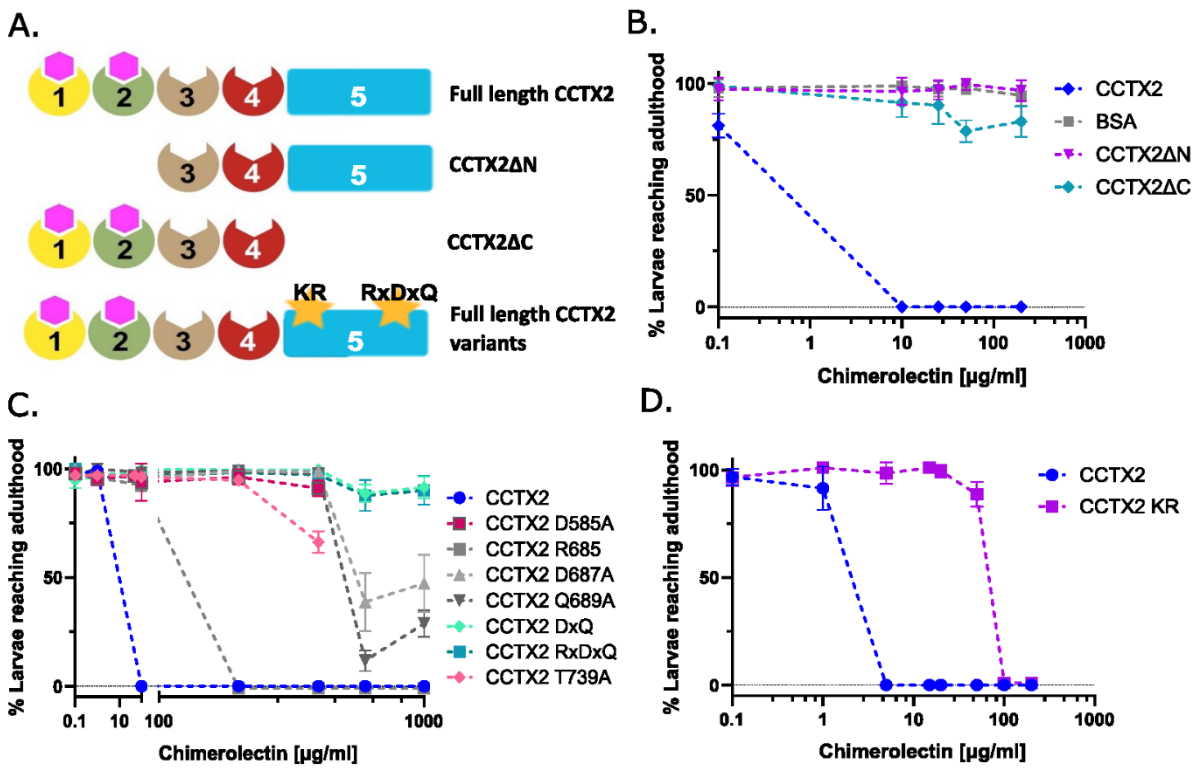

Figure 4. The two N-terminal BTF lectin domains and the C-terminal DUF domain are required for CCTX2 nematotoxicity. **A.** Schematic representation of the CCTX2 architecture with KR and RxDxQ motifs indicated by yellow stars (see main text for details). **B.** *C. elegans* nematotoxicity assay with full-length CCTX2 (blue), a CCTX2 variant lacking the two N-terminal BTF domains (CCTX2ΔN, magenta) and a variant lacking the C-terminal domain of unknown function (CCTX2ΔC, teal), with BSA used as negative control. **C**. *C. elegans* nematotoxicity assay with CCTX2 variants in the RxDxQ motif either as single substitutions (grey, R685A, D687A and Q689A) or double (DxQ, light teal) or triple alanine replacements (RxDxQ, teal). Point mutations D585A and T739A also conferred resistance to the toxin. **D.** Resistance of *C. elegans* to CCTX2-KR (E595A, K596A, R597A) in nematotoxicity assay. Data points with error bars indicate means of N = 4 biological replicates with standard error of the mean (SEM). Overlapping data points were nudged by +/- 2 data units in Y for better readability.

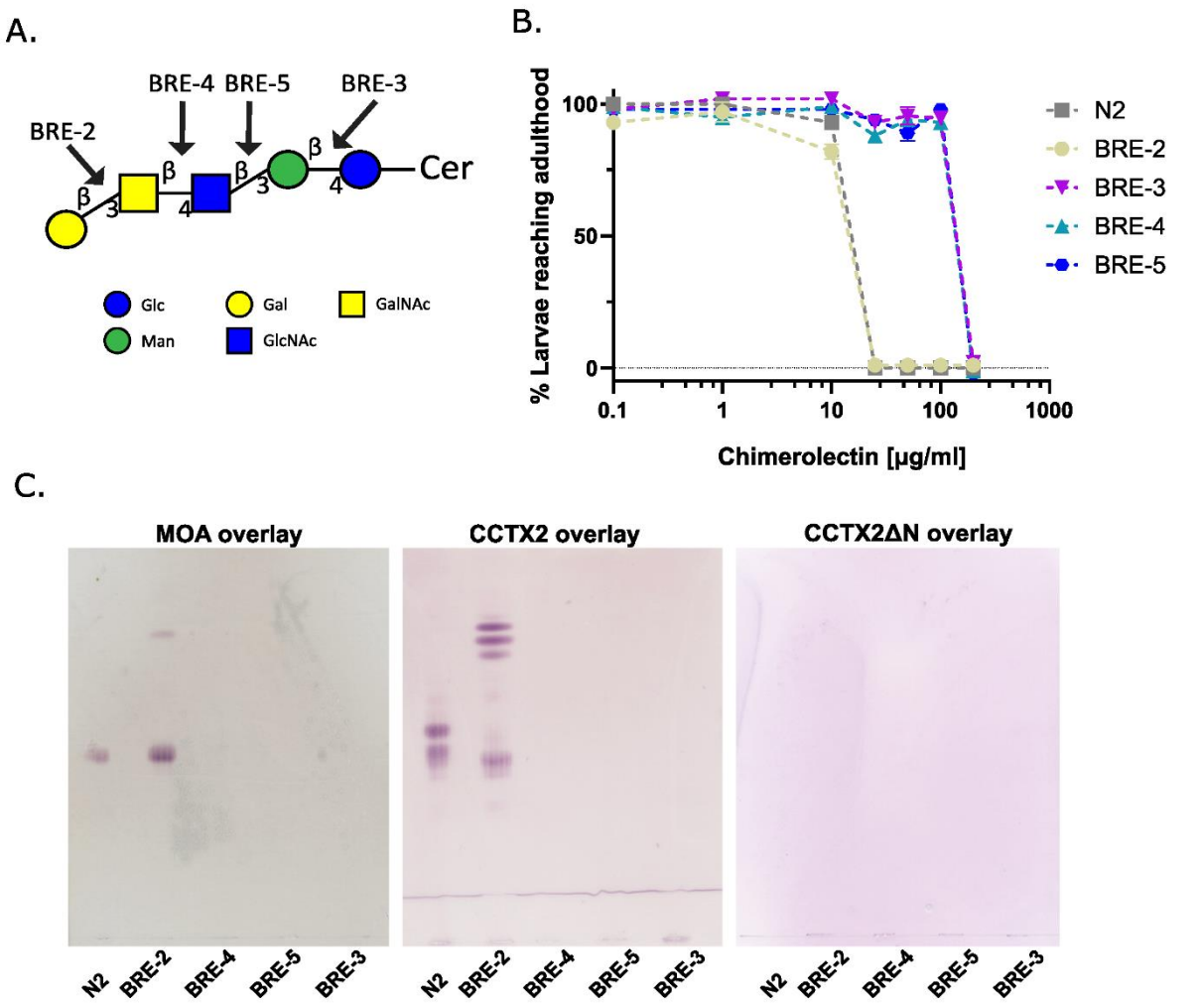

Figure 5. CCTX2 binds to *C. elegans* intestinal epithelial glycosphingolipids *in vivo* and *in vitro.* A. Schematic representation of the arthroseries GSL core including the *C. elegans* enzymes responsible for the addition of individual monosaccharides. The ceramide-linked core is synthesized by sequential action of the enzymes BRE-3, BRE-5 and BRE-4; GalNAc is transferred by BRE-2. Monosaccharides are represented as introduced by the Consortium for Functional Glycomics (CFG) (Varki et al, 2015). **B.** Toxicity assay on *C. elegans* GSL biosynthesis mutants *bre-3(ye26), bre-5(ye-17), bre-4(ye27)* and *bre-2(ye31).* L1 larvae of mutant and wild-type worms were fed with different concentrations of CCTX2 (up to 200 µg/ml) and scored for the percentage of larvae reaching the L4 stage/adulthood. *C. elegans* N2 and the mutant for the GalNAc transferase *bre-2(ye31)* were sensitive to CCTX2 at low concentrations. Data points with error bars indicate means of N = 4 biological replicates with standard error of the mean (SEM). Overlapping data points were nudged by +/- 2 data units in Y for better readability. **C.** TLC overlays of upper phase GSLs isolated from the respective *C. elegans* GSL biosynthesis mutants and N2 wild-type strains, blotted with 150 nM of the indicated biotinylated protein.

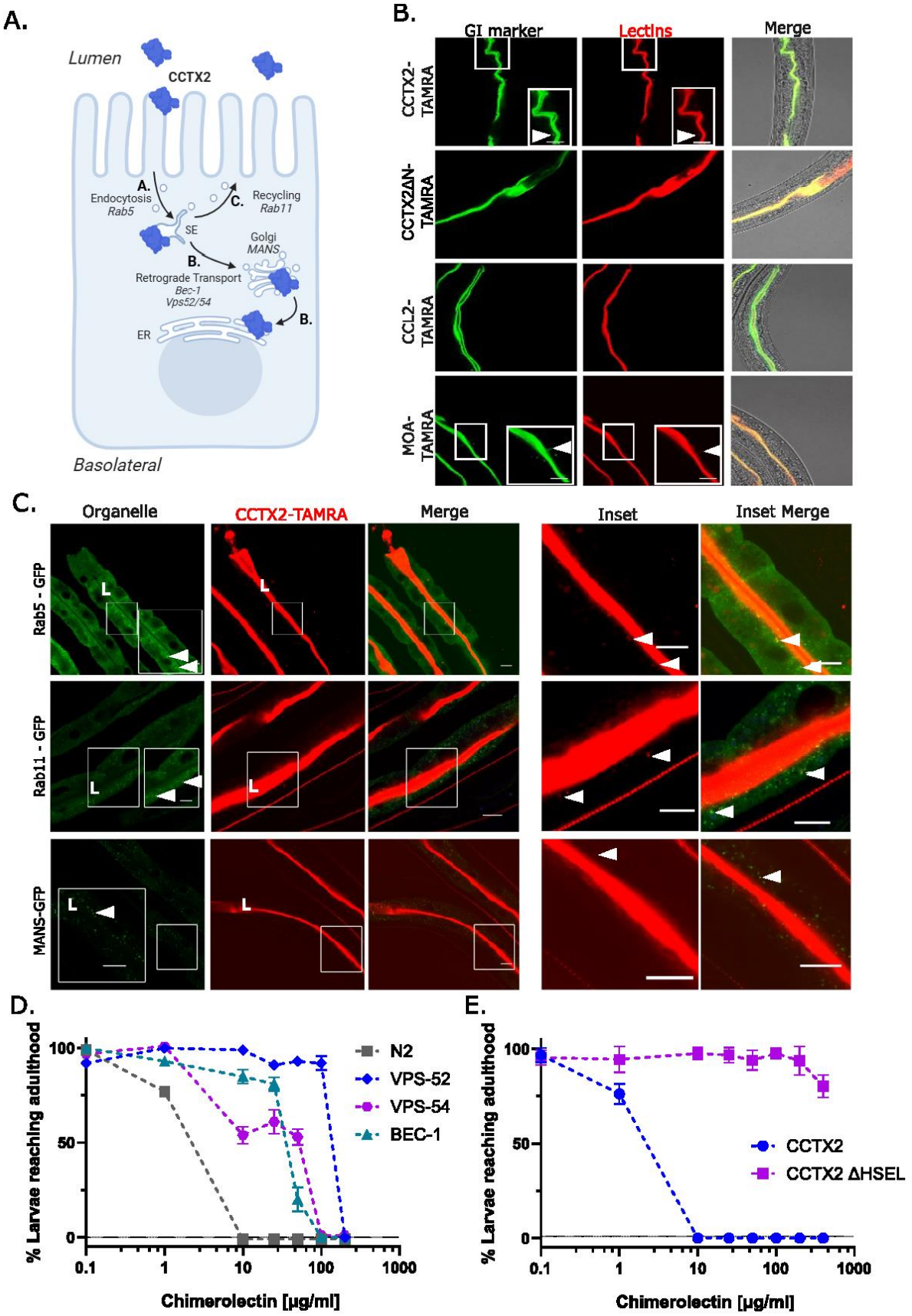

Figure 6. Endocytosis and retrograde trafficking of CCTX2 in *C. elegans* enterocytes. A. Schematic representation of the endocytic network in *C. elegans* enterocytes. After endocytosis and fusion of the endocytic vesicles with the sorting endosome (SE) (A.), the cargo is sorted either for recycling back to the plasma membrane (C.) or for retrograde trafficking to

the trans Golgi network (TGN), Golgi and endoplasmic reticulum (ER) (B.). Cargo destined for degradation is sorted into the late endosome (LE) and the lysosomal pathway. **B.** Internalization of CCTX2 into *C. elegans* enterocytes. Fluorescently-labeled CCTX2 (red), truncated CCTX2ΔN (red), the GSL-binding chimerolectin MOA (red) or the N-glycan binding lectin CCL2 (red) were fed to *C. elegans* GK70 (green) for 3 h at 20°C. Arrowheads indicate fluorescent puncta (vesicles) as a sign of endocytosis. These vesicles appear red (due to the TAMRA-labelled cargo), green (due to the intestinal apical membrane marker PGP-1::GFP) or yellow (merge). **C.** Co-localization of CCTX2 with endocytic markers. *C. elegans* RAB-5::GFP, RAB-11::GFP as well as the Golgi marker α-Mannosidase II (MANS::GFP) were used as endocytic and retrograde trafficking markers to test for co-localization with CCTX2. TAMRA-labeled CCTX2 was fed to the respective *C. elegans* L4 larvae for 6 h at 20°C and monitored for co-localization. Images show the intestinal cross-sections. Arrowheads indicate overlapping puncta (vesicles). 'L' indicates the position of the intestinal lumen. Scale bars = 10 µm. **D.** Nematotoxicity assay with CCTX2 with different *C. elegans* retrograde trafficking mutants: VPS-52 (VC625) and VPS-54 (VC985) bear mutations in the GARP complex subunits VPS-52 and VPS-54, respectively. BEC-1 is required for retromer localization to endosomes. VPS-36 (VC947) carries mutation in the ESCRT II subunit. **E.** Nematotoxicity assay of CCTX2 variant lacking the terminal HSEL sequon. Data points with error bars indicate means of N = 4 biological replicates with standard error of the mean (SEM). Overlapping data points were nudged by +/- 2 data units in Y for better readability.