

1 **Novel anionic cecropins from the spruce budworm feature a poly-L-aspartic acid**
2 **C-terminus**

3

4 Halim Maaroufi^{a*}, Michel Cusson^b, Roger C. Levesque^c

5

6 ^a Institut de biologie intégrative et des systèmes (IBIS), Université Laval, Quebec City,
7 Canada

8 ^b Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre,
9 Quebec City, Canada

10 ^c Institut de biologie intégrative et des systèmes (IBIS) and Faculté de médecine,
11 Université Laval, Quebec City, Canada

12

13

14 *Corresponding author

15 Dr Halim Maaroufi

16 Institut de biologie intégrative et des systèmes (IBIS), Université Laval, Pavillon Charles-
17 Eugène Marchand, 1030, avenue de la médecine, Québec, Québec G1V 0A6, Canada.
18 E-mail : halim.maaroufi@ibis.ulaval.ca

19 **Abstract**

20 Cecropins form a family of amphipathic α -helical cationic peptides with broad-spectrum
21 antibacterial properties and potent anticancer activity. The emergence of bacteria and
22 cancer cells showing resistance to cationic antimicrobial peptides (CAMPs) has fostered
23 a search for new, more selective and more effective alternatives to CAMPs. With this
24 goal in mind, we looked for cecropin homologs in the genome and transcriptome of the
25 spruce budworm, *Choristoneura fumiferana*. Not only did we find paralogs of the
26 conventional cationic cecropins ($Cfcec^+$), our screening also led to the identification of
27 previously uncharacterized anionic cecropins ($Cfcec^-$), featuring a poly-L-aspartic acid
28 C-terminus. Comparative peptide analysis indicated that the C-terminal helix of $Cfcec^-$ is
29 amphipathic, unlike that of $Cfcec^+$, which is hydrophobic. Interestingly, molecular
30 dynamics simulations pointed to the lower conformational flexibility of $Cfcec^-$ peptides,
31 relative to that of $Cfcec^+$. Phylogenetic analysis suggests that the evolution of distinct
32 $Cfcec^+$ and $Cfcec^-$ peptides may have resulted from an ancient duplication event within
33 the Lepidoptera. Our analyses also indicated that $Cfcec^-$ shares characteristics with
34 entericidins, which are involved in bacterial programmed cell death, lunasin, a peptide of
35 plant origins with antimitotic effects, and APC15, a subunit of the anaphase-promoting
36 complex. Finally, we found that both anionic and cationic cecropins contain a BH3-like
37 motif (G-[KQR]-[HKQNR]-[IV]-[KQR]) that could interact with Bcl-2, a protein
38 involved in apoptosis; this observation is congruent with previous reports indicating that
39 cecropins induce apoptosis. Altogether, our observations suggest that cecropins may
40 provide templates for the development of new anticancer drugs.

41

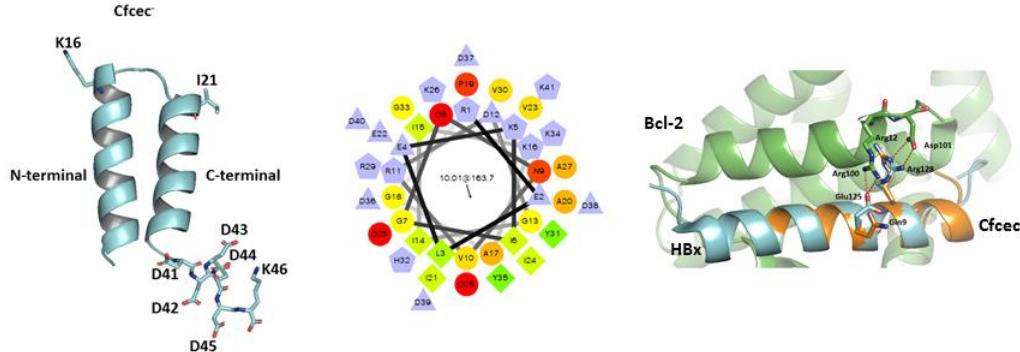
42 **Keywords:** *Choristoneura fumiferana*; anionic cecropins; C-terminal poly-L-aspartic
43 acid; ancient duplication; apoptotic motif; anticancer peptide.

44

45

46 **Graphical abstract**

47



48

49

50 **Highlights**

- 51 1. Genes encoding novel anionic cecropins (Cfcec⁻), featuring a C-terminal poly-L-
- 52 aspartic acid, were found in the genome of the spruce budworm, *Choristoneura*
- 53 *fumiferana*.
- 54 2. Divergence between Cfcec⁺ and Cfcec⁻ could be the result of an ancient duplication
- 55 event within the Lepidoptera.
- 56 3. There is an apparent relationship between motifs observed in cecropin peptides and
- 57 apoptosis.
- 58 4. Anionic cecropins from the spruce budworm display characteristics suggesting they
- 59 could have anticancer activity

60

61 **Abbreviations:** Cfcec⁺, *C. fumiferana* cationic cecropins; Cfcec⁻, *C. fumiferana* anionic
62 cecropins; PAA, C-terminal poly-L-aspartic acid; AMPs, antimicrobial peptides;
63 CAMPs, cationic antimicrobial peptides; ACPs, anticancer-peptides; APD, Antimicrobial
64 Peptide Database; LPS, lipopolysaccharides; Pxcec⁺, cationic *Papilio Xuthus* cecropin;
65 HccecA, *H. cecropina* cecropin-A; Pxcec⁻, anionic *P. Xuthus* cecropin; CED-9, cell-death
66 abnormal 9; Bcl-2, B-cell lymphoma 2; APC15, subunit anaphase-promoting complex
67 15; MCC, the mitotic checkpoint complex; SAC, spindle assembly checkpoint; hepG-2,
68 human liver hepatocellular carcinoma cell line.

69

70 **1. Introduction**

71 Antimicrobial peptides (AMPs) constitute an important component of the innate immune
72 system of insects and, as such, may have played a role in the evolutionary success of this
73 highly speciose taxon, whose members occupy almost all habitats in nature. To date, over
74 150 insect AMPs have been identified (Yi et al., 2014), with cecropin being the first one
75 to have been purified, in 1980, from pupae of the cecropia moth, *Hyalophora cecropia*
76 (Hultmark et al., 1980; Steiner et al., 1981). Cecropins are now known to form a family
77 of amphipathic α -helical peptides containing 34–39 amino acid residues. They display a
78 broad spectrum of antibacterial properties and act as modulator of the innate immune
79 system. They also show potent anticancer activity (Suttmann et al., 2008; Hoskin and
80 Ramamoorthy, 2008; Huang et al., 2015).

81 Cecropins characterized to date fall in the category of cationic antimicrobial
82 peptides (CAMPs), whose distinguishing feature is an excess in basic amino acids
83 (positive charge, cationic; Otvos, 2000). Their amphipathic properties allow interactions
84 with membranes (Lee et al., 2013) and their toxicity toward bacteria is considered to be
85 primarily due to an initial electrostatic interaction between the peptide and the anionic
86 phospholipid head groups in the outer layer of the bacterial cytoplasmic membrane,
87 ultimately causing membrane disruption (Zasloff, 2002).

88 For the present work, we took advantage of genomic and transcriptomic resources
89 recently developed by our group for the spruce budworm, *Choristoneura fumiferana*, to
90 identify and characterize the repertoire of cecropins in this important lepidopteran conifer
91 pest. Our analyses led to the identification of novel, anionic cecropins (*Cfcec⁻*) featuring a
92 poly-L-aspartic acid (PAA) C-terminus. We used *in silico* approaches to compare the
93 physico-chemical properties of the anionic peptides with those of the more conventional
94 cationic cecropins, some of which were also found in the *C. fumiferana* genome (*Cfcec⁺*).
95 In addition, we compared *C. fumiferana* cecropins with other bioactive peptides whose
96 biochemical properties have been well characterized; in the process, we identified some
97 peptides that share features with cecropins, including motifs that could explain why the
98 latter can induce apoptosis.

99

100 **2. Materials and Methods**

101 *2.1. C. fumiferana genomic and transcriptomic resources*

102 To identify spruce budworm cecropins, we queried a draft assembly of the *C. fumiferana*
103 genome, generated from Roche 454 (GS-FLX+) whole genome shotgun reads, assembled
104 using the Newbler software (Roche); sequencing was performed using DNA extracted
105 from a single male pupa (Cusson et al. unpublished). Similar searches were conducted by
106 querying a *C. fumiferana* transcriptome. The latter was generated using both Illumina and
107 Roche 454 (GS-FLX+) reads following sequencing of a normalized cDNA library,
108 generated from a pool of mRNAs collected from all *C. fumiferana* life stages. Contig
109 assembly was carried out using the MIRA (Chevreux et al. 2004) assembler (Brandão et
110 al. unpublished).

111

112 *2.2. Blast*

113 Similarity searches in our genomic and transcriptomic databases were performed locally
114 using the tblastn algorithm (<http://www.ncbi.nlm.nih.gov/blast>) and the sequence of the
115 *H. cecropia* cecropin-A (HcceCA; Uniprot: P01507) as query. We also conducted similar
116 tblastn searches against a public *C. fumiferana* EST database (NCBI). To determine if
117 other cecropin homolog sequences with PAA exist in other species, we searched in all
118 sequenced genomes in GenBank using BLASTp, tBLASTn and HMM profiles.

119

120 *2.3. Chemical properties of cecropins*

121 The chemical structures and properties of cecropin peptides were investigated *in silico*
122 using PepDraw (<http://www.tulane.edu/~biochem/WW/PepDraw/index.html>).
123 Hydrophobicity, as determined by PepDraw, is the free energy associated with
124 transitioning a peptide from an aqueous environment to a hydrophobic environment such
125 as octanol. The scale used is the Wimley-White scale, an experimentally determined
126 scale, where the hydrophobicity of the peptide is the sum of Wimley-White
127 hydrophobicities and measured in Kcal/mol (White and Wimley, 1998). Neutral pH is
128 assumed.

129

130

131 2.4. Structure predictions and molecular dynamics simulations

132 The secondary structures of cecropins and lunasin were predicted using Psipred software
133 (Jones, 1999), while helical wheel projections for the same peptides were performed
134 using the tool available at <http://rzlab.ucr.edu/scripts/wheel/wheel.cgi>.

135 3D homology models of *C. fumiferana* cecropins and lunasin were built using the
136 crystal structures of papiliocin of *Papilio xuthus* (PDB id: 2LA2) and allergen Ara h6 of
137 *Arachis hypogaea* (PDB id: 1W2Q) as templates, respectively, using the modeling
138 software Modeller (Webb and Sali, 2014). Model quality was assessed by Ramachandran
139 plot analysis through PROCHECK (Laskowski et al., 1993). Structure images were
140 generated using PyMOL (<http://www.pymol.org>).

141 In order to evaluate conformational changes of cecropins, molecular dynamics
142 (MD) simulations were performed in GROMACS (v5.1.4) using the OPLS-AA/L all-
143 atom force field (Kaminski et al. 2001). Cecropins were solvated in a cubic box as the
144 unit cell, using SPC/E water model with the box edge distance from the molecule set to
145 1.0 nm. The system was neutralized by replacing solvent molecules with Cl⁻ and Na⁺
146 ions. Energy minimization was conducted using the steepest descent method to ensure
147 that the system has no steric clashes or inappropriate geometry. Equilibration of the
148 solvent and ions around the peptide was conducted under NVT (300 K) and NPT (1.0
149 bar) ensembles for 100 ps. Cecropin MD simulations were conducted for 1 ns.

150

151 2.5. Phylogenetic analysis

152 We used the amino acid sequences of cationic and anionic cecropins to search by BlastP
153 for close homologs in insects. To assess phylogenetic relationships among cecropins of
154 *C. fumiferana* and those of Diptera, Coleoptera and other Lepidoptera, sequences were
155 aligned using Muscle (Edgar, 2004) and a phylogenetic tree was constructed using the
156 Neighbor-Joining method (Saitou and Nei, 1987). Evolutionary distances were computed
157 using the Equal Input method (Tajima and Nei, 1984) and are shown as the number of
158 amino acid substitutions per site. All positions displaying less than 95% site coverage
159 were eliminated. Phylogenetic analyses were conducted in MEGA6 (Tamura et al.,
160 2013).

161

162 **3. Results and discussion**

163

164 ***3.1. Search for C. fumiferana cecropins in genomic and transcriptomic resources***

165 To identify cecropin orthologs in the sequenced genome and transcriptome of *C.*
166 *fumiferana* (unpublished data), we conducted tblastn searches using cecropin-A of *H.*
167 *cecropina* (Uniprot: P01507) as query. We found two types of cecropin genes that were
168 designated *Cfcec⁺* (cationic cecropins) and *Cfcec⁻* (anionic cecropins). *Cfcec⁺* and *Cfcec⁻*
169 peptides were again used as queries to carry out tblastn iteratively until no new hit
170 occurred. Two *Cfcec⁺* and two *Cfcec⁻* genes were found in the genome and transcriptome
171 of *C. fumiferana*. These genes were designated *Cfcec⁺1*, *Cfcec⁺2* and *Cfcec⁻1*, *Cfcec⁻2*,
172 for cationic and anionic cecropins, respectively (Table 1; Fig. 1). Cecropin genes are
173 composed of two exons and one intron (Fig. 1A, B). In the draft assembly of the *C.*
174 *fumiferana* genome, the *Cfcec⁺1* and *Cfcec⁻1* genes were localized on the same scaffold
175 with an interval of 3588 bp, whereas the *Cfcec⁺2* and *Cfcec⁻2* genes were found on
176 distinct scaffolds. Interestingly, another scaffold contains the exon 1 of *Cfcec⁺2* and the
177 C-terminal poly-L-aspartic acid (PAA) of *Cfcec⁻1* (Fig. 1B). It is tempting to speculate
178 that the ancestor of *Cfcec⁻* acquired the PAA C-terminus through exon shuffling.

179 To confirm that *Cfcec⁺* and *Cfcec⁻* are actually transcribed, we searched *C.*
180 *fumiferana* transcriptomic and EST databases, where we found the corresponding
181 transcripts. In addition, NCBI's EST database revealed that *Cfcec⁺* and some anionic
182 cecropins of Lepidoptera (Fig. S1) were expressed in frontline defense tissues such as the
183 fat body, epidermis and midgut, as shown for *Plutella xylostella* (Jin et al., 2012).

184

185 ***3.2. Peptide analysis***

186 ***3.2.1. Sequence analysis***

187 Alignment of *Cfcec⁺* and *Cfcec⁻* showed that the major differences between these two
188 peptides are found at the N- and C-termini of the mature peptides (Fig. 1D). *Cfcec⁺2*
189 displays characteristics similar to those of other cecropins of moths where the first
190 residue, preceding the conserved Trp2, is a Lys or an Arg (Otvos, 2000). This is not the
191 case for *Cfcec⁻* where these amino acid residues are absent (Fig. 1D). The Trp2 residue of

192 HccecA was shown to be important for activity against all tested bacteria (Andreu et al.,
193 Trp2 and Phe5 of *Papilio xuthus* cecropin (Pxcec⁺) were also shown to interact
194 with LPS (Kim et al., 2011). These two amino acid residues are conserved in Cfcec⁺ but
195 not in Cfcec⁻ (Fig.1D). In addition, the hinge region of cecropins (Gly-Pro) is conserved
196 in Cfcec⁺2 and Cfcec⁻2 as described by Efimova et al. (2014) and provides
197 conformational flexibility (Oh et al., 2000) (Fig.1D). Analysis of amino acid residues of
198 both types of *C. fumiferana* cecropins suggests that Cfcec⁺ shares many characteristics
199 with HccecA (Table 1).

200 Blast searches were conducted to determine whether anionic cecropins such as
201 Cfcec⁻ were present in other insects. Not surprisingly, we found a few cecropins bearing
202 three or four acidic amino acid residues at their C-termini, conferring a negative net
203 charge (-2 to -4) to the peptides (Fig. S1). The characteristics of anionic *P. xuthus*
204 cecropin (Pxcec⁻) are presented in Table 1; this peptide shares several characteristics with
205 Cfcec⁻, including a similar length, a negative net charge, and comparable pI and
206 hydrophobicity values. To our knowledge, the activity and structure of these anionic
207 cecropins has never been examined experimentally. Finally, it is worth noting that the
208 PAA of Cfcec⁻ ends with a lysine residue (Fig. 2). Interestingly, a lysine residue was
209 similarly found at the end of the poly-L-aspartic acid peptides (VDDDDK, APDDDDK
210 and TDDDK) studied by Brogden et al. (1997). It is conceivable that the role of this
211 positively charged residue is to interact with negative charges on membrane. Indeed, the
212 long nonpolar region of the side chain of lysine was shown to extend or snorkel into the
213 hydrophobic core of the target membrane (Li et al., 2013).

214

215 3.2.2. *Structure and molecular dynamics simulation*

216 2D and 3D structures of Cfcec⁺ and Cfcec⁻ were estimated using Psipred and homology
217 modeling (Fig. 1C), respectively. For both types of peptides, the structure comprises two
218 helices, as shown by NMR for other cecropins (PDB id: 2LA2 and 2MMM). To illustrate
219 the properties of α -helices in these peptides, we constructed helical wheel diagrams.
220 Thus, the N-terminal helices of Cfcec⁺ and Cfcec⁻ both display amphipathic properties
221 (amino acids 1-21 of Cfcec⁺ and 1-17 of Cfcec⁻), whereas their C-terminal helices (amino

222 acids 25–37 of Cfcec⁺ and 21–34 of Cfcec⁻) are hydrophobic and amphipathic,
223 respectively (Fig. S2). In comparison, lunasin, a peptide of plant origins that shares some
224 characteristics with Cfcec⁻ (Table 1), has N- and C-terminal helices that are hydrophilic
225 and amphipathic, respectively (Fig. S2). The high antibacterial activity of cecropin-like
226 model peptides has been shown to require a basic, amphipathic N-terminal helix and a
227 hydrophobic C-terminal helix, connected by a flexible hinge region (Fink et al., 1989).
228 Cfcec⁺ possesses these characteristics and can therefore be classified as a CAMP.
229 However, Cfcec⁻ does not fit this pattern, with its amphipathic C-terminal helix, a feature
230 shared with lunasin (Dia and de Mejia, 2011), which displays anti-cancer activity.

231 Molecular dynamics simulations showed that Cfcec⁻ peptides have less
232 conformational flexibility at their N- and C-termini than Cfcec⁺ (Fig. S3), with only the
233 PAA of Cfcec⁻ displaying significant flexibility. Flexibility appears to be provided
234 primarily by glycine residues (Fig. S3). Since CAMP activity is dependent upon the
235 presence of conformational flexibility (Amos et al. 2016), differences in this variable
236 between the Cfcec⁺ and Cfcec⁻ peptides, in addition to the different physico-chemical
237 properties of their C-termini, suggests that Cfcec⁺ and Cfcec⁻ could have different
238 biological activities.

239

240 3.3. Phylogeny

241 To infer phylogenetic relationships among cecropins of *C. fumiferana* and those of other
242 Lepidoptera, Diptera, and Coleoptera, sequences reported here and others gleaned from
243 public databases were aligned using Muscle (Edgar, 2004), and a phylogenetic tree was
244 constructed using the Neighbor-Joining method (Saitou and Nei, 1987). The branching
245 pattern obtained suggests that Cfcec⁺ peptides are more closely related to their *H.*
246 *cecropia* and *D. plexippus* CecA and CecB orthologs than to their *C. fumiferana* paralogs
247 (Cfcec⁻; Fig. 2B). Again, this observation suggests that anionic cecropins could have
248 arisen following an ancient duplication event within the Lepidoptera. Not surprisingly,
249 lepidopteran cecropins with negative and positive net charges formed two separate clades
250 (Figure 2B). It has been observed that the net charge of α -helical ACPs has an effect on
251 the anticancer activity of the peptide, with an increase in the net charge enhancing
252 anticancer activity (Huang et al., 2015).

253

254 *3.4. Peptides that display similarity to Cfcec⁺ and Cfcec⁻*

255 In recent years, AMPs have become promising molecules to fight cancer (Hoskin and
256 Ramamoorthy, 2008; Riedl et al., 2011). The Antimicrobial Peptide Database (APD,
257 <http://aps.unmc.edu/AP/main.php>) currently contains 193 peptides that are identified as
258 anticancer-peptides (ACPs). These peptides, however, are from different sources and
259 display limited similarity to one another. Nonetheless, they share some characteristics,
260 including a positive charge and an amphipathic helix, and they exhibit a large spectrum
261 of anticancer activity (Gaspar et al., 2013; Lu et al., 2016).

262 BLAST searches against the AMP database using Cfcec⁻ as query revealed that
263 the PAA of mature lunasin (amino acids 56-64) is similar to that of Cfcec⁻. Moreover,
264 lunasin shares other features with Cfcec⁻, including net charge, pI, hydrophobicity and
265 secondary structure (Table 1). Lunasin has antimitotic properties attributed to the binding
266 of its PAA C-terminus to regions of hypoacetylated chromatin (Galvez and de Lumen,
267 1999; Galvez et al., 2001). In the AMP database, we also found other short peptides
268 containing mostly aspartic acid residues, including the peptide DEDDD, which shows
269 inhibitory activity against breast cancer cells (Li et al., 2016), as well as three other
270 AMPs: DDDDDDD, GDDDDDD and GADDDDD (Brogden et al., 1996). To refine our
271 search for peptides displaying amino acid patterns similar to those of Cfcec⁻, we turned
272 our attention to profile-profile alignment methods (Xu et al., 2014), which are more
273 accurate and more sensitive than sequence-sequence and sequence-profile alignment
274 methods. In this way, we identified HBx, a hepatitis B virus protein whose C-terminus is
275 similar to the N-terminus of Cfcec⁻ (Fig. 3A). HBx contains a BH3-like motif that folds
276 to form an amphipathic α -helix that binds to the conserved BH3-binding groove of Bcl-2,
277 an anti-apoptotic protein (Ma et al., 2008; Jiang et al., 2016). The BH3-like motif may
278 inhibit the anti-apoptotic action of Bcl-2 through interactions with its conserved BH3-
279 binding groove. Glu125 and Arg128 of HBx (Fig. 3C) each makes a pair of charge-
280 stabilized hydrogen bonds to residues Arg100 and Asp101 in Bcl-2 (Jiang et al., 2016).
281 The amino acids corresponding to Glu125 and Arg128 in the N-terminus of Cfcec⁻ are
282 Gln9 and Arg12 (Fig. 3A and C). To determine if Cfcec⁻ could interact with Bcl-2, we
283 superimposed the 3D model of Cfcec⁻1 onto the crystal structure of the HBx-Bcl-2

284 complex (PDBid: 5FCG). Indeed, the N-terminus of Cfcec⁻¹ superimposed well onto
285 HBx (Fig. 3C) and the amino acids Gln9 and Arg12 could form hydrogen bonds with
286 Arg100 and Asp101 of Bcl-2. Interestingly, a G124L/I127A double mutation (conserved
287 in cecropins; Fig. 3A) in HBx abolished interactions between HBx and CED-9, the
288 functional ortholog of Bcl-2 in *C. elegans* (Geng et al., 2012), pointing to an important
289 functional role of these two residues in HBx, and possibly in cecropins too.

290 Alignment of cecropins from different species showed that BH3-like motifs are
291 also present at their N-termini. In addition, we found this motif to be present in other
292 AMPs that have been shown to induce apoptosis (Table 2). Interestingly, Choi and Lee
293 (2013) showed that when the first six amino acid residues (GWGSFF) of pleurocidin
294 (Table 2) were truncated, the peptide lost its effect on ROS production (apoptosis),
295 apparently because the N-terminus of pleurocidin contains a BH3-like motif (Table 2,
296 underlined amino acids). Similarly, cecropin-P17 (Table 2) was shown to suppress the
297 proliferation of HepG-2 cells by inducing apoptosis, which was dependent (among other
298 things) on inhibiting Bcl-2 (Wu et al., 2015).

299 An intensive search for proteins displaying similarity to Cfcec's led to the finding
300 of entericidins A (EcnA) and B (EcnB) (Fig. 3B), two interesting peptides from bacteria.
301 Entericidins are small lipoproteins encoded by tandem genes whose products function as
302 toxin/antidote in programmed cell death in bacteria. EcnA inhibits the apoptotic action of
303 EcnB by a mechanism that is not yet understood (Bishop et al., 1998). Amino acid
304 sequences of EcnA and EcnB display similarities to AMPs. They have a signal peptide,
305 adopt amphipathic α -helical structures and reciprocally modulate membrane stability
306 (Fig. 3B). Moreover, EcnA and EcnB, like many AMPs, contain a BH3-like motif (Table
307 2). This supports the hypothesis that programmed cell death genes may have originated in
308 bacteria from a pool of antibiotic genes (Ameisen, 1996).

309 Lastly, we found that Cfcec⁻ displays sequence similarity to the subunit anaphase-
310 promoting complex 15 (APC15), which plays a role in the release of the mitotic
311 checkpoint complex (MCC) from the APC/C (Fig. 3D). The function of APC15 in human
312 cells seems to be primarily linked to the spindle assembly checkpoint (SAC), and its
313 depletion prevents mitotic slippage (Mansfeld et al., 2011).

314

315 *3.5. Prospective function of PAA*

316 The number of aspartic acid residues differs between the PAA of Cfcec⁻¹ and Cfcec⁻²,
317 with five in the latter and eight in the former. This variation changes the net charge of
318 these peptides. It has been reported that variation in the net charge of α -helical ACPs has
319 an effect on the anticancer activity of the peptide (Huang et al., 2015). Moreover,
320 Brogden et al. (1996) showed that the antimicrobial activity of Asp homopolymers
321 increases with the number of Asp residues in the peptide. In addition, lunasin contains a
322 unique Arg-Gly-Asp (RGD) cell adhesion motif just upstream its PAA (Dia and de
323 Mejia, 2011). Peptides with an RGD motif bind integrins with high specificity, leading to
324 antiangiogenic and anti-inflammatory effects (Kuphal et al., 2005). Cfcec⁻¹ has an NGD
325 motif at the corresponding position (Fig. 1D); as we are now set to assess the putative
326 anticancer activity of *C. fumiferana* anionic cecropins, it will be interesting to examine
327 the impact of mutating Asn60 to Arg in order to generate the RGD motif found in
328 lunasin.

329

330 **Acknowledgements**

331 We thank the bioinformatics platform and the next generation sequencing platform at
332 IBIS (Université Laval) for their assistance.

333

334

335

336 **Supplementary data**

337 Figures S1-S3.

338 References

- 339 Aerts, A.M., Carmona-Gutierrez, D., Lefevre, S., Govaert, G., François, I.E., Madeo,
340 F., Santos, R., Cammue, B.P., Thevissen, K., 2009. The antifungal plant defensin
341 RsAFP2 from radish induces apoptosis in a metacaspase independent way in
342 *Candida albicans*. FEBS Lett 583, 2513–2516. doi:10.1016/j.febslet.2009.07.004
- 343 Ameisen J.C., 1996. The origin of programmed cell death. Science 272, 1278-1279.
- 344 Amos S.T, Vermeer L.S, Ferguson P.M, Kozlowska J, Davy M, Bui T.T, et al. 2016.
345 Antimicrobial Peptide Potency is Facilitated by Greater Conformational Flexibility
346 when Binding to Gram-negative Bacterial Inner Membranes. Sci Rep. 6, 37639.
- 347 Andreu, D., Merrifield, R.B., Steiner, H., Boman, H.G., 1985. N-terminal analogues of
348 cecropin A: synthesis, antibacterial activity, and conformational properties.
349 Biochemistry 24, 1683–1688.
- 350 Bishop R.E., Leskiw B.K., Hodges R.S., Kay C.M., Weiner J.H., 1998. The entericidin
351 locus of *Escherichia coli* and its implications for programmed bacterial cell death. J
352 Mol Biol 280, 583-596.
- 353 Brogden, K.A., De Lucca, A.J., Bland, J., Elliott, S., 1996. Isolation of an ovine
354 pulmonary surfactant-associated anionic peptide bactericidal for *Pasteurella*
355 *haemolytica*. Proc. Natl. Acad. Sci. U. S. A. 93, 412–6. doi:10.1073/pnas.93.1.412
- 356 Brogden, K.A., Ackermann, M., Huttner, K.M., 1997. Small, anionic, and charge-
357 neutralizing propeptide fragments of zymogens are antimicrobial. Antimicrob.
358 Agents Chemother. 41, 1615–1617.
- 359 Chen, J.Y., Lin, W.J., Wu, J.L., Her, G.M., Hui, C.F., 2009. Epinecidin-1 peptide induces
360 apoptosis which enhances antitumor effects in human leukemia U937 cells. Peptides
361 30, 2365–2373. doi:10.1016/j.peptides.2009.08.019
- 362 Chevreux, B., Pfisterer, T., Drescher, B., Diesel, A.J., Müller, W.E., Wetter, T., et al.,
363 2004. Using the miraEST assembler for reliable and automated mRNA transcript
364 assembly and SNP detection in sequenced ESTs. Genome Res 14, 1147-1159.
- 365 Cho, J., Lee, D.G., 2011. The antimicrobial peptide arenicin-1 promotes generation of
366 reactive oxygen species and induction of apoptosis. Biochim Biophys Acta 1810,
367 1246–1251. doi:10.1016/j.bbagen.2011.08.011
- 368 Choi, H., Lee, D.G., 2013. The influence of the N-terminal region of antimicrobial peptide
369 pleurocidin on fungal apoptosis. J Microbiol Biotechnol 23, 1386–1394.

- 370 Dia, V.P., de Mejia, E., 2011. Lunasin induces apoptosis and modifies the expression of
371 genes associated with extracellular matrix and cell adhesion in human metastatic
372 colon cancer cells. *Mol Nutr Food Res* 55, 623–634. doi:10.1002/mnfr.201000419
- 373 Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time
374 and space complexity. *BMC Bioinformatics* 5, 113. doi:10.1186/1471-2105-5-113
- 375 Efimova, S.S., Schagina, L. V, Ostroumova, O.S., 2014. Channel-forming activity of
376 cecropins in lipid bilayers: effect of agents modifying the membrane dipole potential.
377 *Langmuir* 30, 7884–7892. doi:10.1021/la501549v
- 378 Fink, J., Boman, A., Boman, H.G., Merrifield, R.B., 1989. Design, synthesis and
379 antibacterial activity of cecropin-like model peptides. *Int J Pept Protein Res* 33, 412–
380 421.
- 381 Galvez, A.F., Chen, N., Macasieb, J., de Lumen, B.O., 2001. Chemopreventive property
382 of a soybean peptide (lunasin) that binds to deacetylated histones and inhibits
383 acetylation. *Cancer Res* 61, 7473–7478.
- 384 Galvez, A.F., de Lumen, B.O., 1999. A soybean cDNA encoding a chromatin-binding
385 peptide inhibits mitosis of mammalian cells. *Nat Biotechnol* 17, 495–500.
386 doi:10.1038/8676
- 387 Gaspar, D., Veiga, A.S., Castanho, M.A., 2013. From antimicrobial to anticancer peptides.
388 A review. *Front Microbiol* 4, 294. doi:10.3389/fmicb.2013.00294
- 389 Geng, X., Harry, B.L., Zhou, Q., Skeen-Gaar, R.R., Ge, X., Lee, E.S., Mitani, S., Xue, D.,
390 2012. Hepatitis B virus X protein targets the Bcl-2 protein CED-9 to induce
391 intracellular Ca²⁺ increase and cell death in *Caenorhabditis elegans*. *Proc Natl Acad
392 Sci U S A* 109, 18465–18470. doi:10.1073/pnas.1204652109
- 393 Hoskin, D.W., Ramamoorthy, A., 2008. Studies on anticancer activities of antimicrobial
394 peptides. *Biochim Biophys Acta* 1778, 357–375. doi:10.1016/j.bbamem.2007.11.008
- 395 Hsu, J.C., Lin, L.C., Tzen, J.T., Chen, J.Y., 2011. Characteristics of the antitumor activities
396 in tumor cells and modulation of the inflammatory response in RAW264.7 cells of a
397 novel antimicrobial peptide, chrysophsin-1, from the red sea bream (*Chrysophrys
398 major*). *Peptides* 32, 900–910. doi:10.1016/j.peptides.2011.02.013
- 399 Huang, Y., Feng, Q., Yan, Q., Hao, X., Chen, Y., 2015. Alpha-helical cationic anticancer
400 peptides: a promising candidate for novel anticancer drugs. *Mini Rev Med Chem* 15,
401 73–81.
- 402 Hultmark, D., Steiner, H., Rasmuson, T., Boman, H.G., 1980. Insect immunity.
403 Purification and properties of three inducible bactericidal proteins from hemolymph
404 of immunized pupae of *Hyalophora cecropia*. *Eur J Biochem* 106, 7–16.

- 405 Hwang, B., Hwang, J.S., Lee, J., Kim, J.K., Kim, S.R., Kim, Y., Lee, D.G., 2011. Induction
406 of yeast apoptosis by an antimicrobial peptide, Papiliocin. Biochem Biophys Res
407 Commun 408, 89–93. doi:10.1016/j.bbrc.2011.03.125
- 408 Hwang, B., Hwang, J.S., Lee, J., Lee, D.G., 2011b. The antimicrobial peptide, psacotheasin
409 induces reactive oxygen species and triggers apoptosis in *Candida albicans*. Biochem
410 Biophys Res Commun 405, 267-271.
- 411 Jiang, T., Liu, M., Wu, J., Shi, Y., 2016. Structural and biochemical analysis of Bcl-2
412 interaction with the hepatitis B virus protein HBx. Proc Natl Acad Sci U S A 113,
413 2074–2079. doi:10.1073/pnas.1525616113
- 414 Jin, F., Sun, Q., Xu, X., Li, L., Gao, G., Xu, Y., Yu, X., Ren, S., 2012. CDNA cloning and
415 characterization of the antibacterial peptide cecropin 1 from the diamondback moth,
416 *Plutella xylostella* L. Protein Expr. Purif. 85, 230–238. doi:10.1016/j.pep.2012.08.006
- 417 Jin, X., Mei, H., Li, X., Ma, Y., Zeng, A.H., Wang, Y., Lu, X., Chu, F., Wu, Q., Zhu, J.,
418 2010. Apoptosis-inducing activity of the antimicrobial peptide cecropin of *Musca*
419 *domestica* in human hepatocellular carcinoma cell line BEL-7402 and the possible
420 mechanism. Acta Biochim Biophys Sin 42, 259–265.
- 421 Jones, D.T., 1999. Protein secondary structure prediction based on position-specific
422 scoring matrices. J Mol Biol 292, 195–202. doi:10.1006/jmbi.1999.3091
- 423 Kaminski G.A, Friesner R.A, Tirado-Rives J, Jorgensen W.L., 2001. Evaluation and
424 Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with
425 Accurate Quantum Chemical Calculations on Peptides. J Phys Chem B 105, 6474-
426 6487.
- 427 Kim, J.K., Lee, E., Shin, S., Jeong, K.W., Lee, J.Y., Bae, S.Y., Kim, S.H., Lee, J., Kim,
428 S.R., Lee, D.G., Hwang, J.S., Kim, Y., 2011. Structure and function of papiliocin with
429 antimicrobial and anti-inflammatory activities isolated from the swallowtail butterfly,
430 *Papilio xuthus*. J Biol Chem 286, 41296–41311. doi:10.1074/jbc.M111.269225
- 431 Kuphal, S., Bauer, R., Bosserhoff, A.K., 2005. Integrin signaling in malignant melanoma.
432 Cancer Metastasis Rev 24, 195–222. doi:10.1007/s10555-005-1572-1
- 433 Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a
434 program to check the stereochemical quality of protein structures. J. Appl. Crystallogr.
435 26, 283–291. doi:10.1107/S0021889892009944
- 436 Lee, E., Jeong, K.W., Lee, J., Shin, A., Kim, J.K., Lee, D.G., Kim, Y., 2013. Structure-
437 activity relationships of cecropin-like peptides and their interactions with
438 phospholipid membrane. BMB Rep 46, 282–287.

- 439 Lee, J., Hwang, J.S., Hwang, I.S., Cho, J., Lee, E., Kim, Y., Lee, D.G., 2012. Coprisin-
440 induced antifungal effects in *Candida albicans* correlate with apoptotic mechanisms.
441 *Free Radic Biol Med* 52, 2302–2311. doi:10.1016/j.freeradbiomed.2012.03.012
- 442 Lee, W., Lee, D.G., 2014. Magainin 2 induces bacterial cell death showing apoptotic
443 properties. *Curr Microbiol* 69, 794–801. doi:10.1007/s00284-014-0657-x
- 444 Li, L., Vorobyov, I., Allen, T.W., 2013. The different interactions of lysine and arginine
445 side chains with lipid membranes. *J Phys Chem B* 117, 11906–11920.
446 doi:10.1021/jp405418y
- 447 Li, S., Hao, L., Bao, W., Zhang, P., Su, D., Cheng, Y., Nie, L., Wang, G., Hou, F., Yang,
448 Y., 2016. A novel short anionic antibacterial peptide isolated from the skin of
449 *Xenopus laevis* with broad antibacterial activity and inhibitory activity against breast
450 cancer cell. *Arch Microbiol.* doi:10.1007/s00203-016-1206-8
- 451 Lin, H.J., Huang, T.C., Muthusamy, S., Lee, J.F., Duann, Y.F., Lin, C.H., 2012. Piscidin-
452 1, an antimicrobial peptide from fish (hybrid striped bass *morone saxatilis* x *M.*
453 *chrysops*), induces apoptotic and necrotic activity in HT1080 cells. *Zool. Sci* 29, 327–
454 332. doi:10.2108/zsj.29.327
- 455 Lu, Y., Zhang, T.F., Shi, Y., Zhou, H.W., Chen, Q., Wei, B.Y., Wang, X., Yang, T.X.,
456 Chinn, Y.E., Kang, J., Fu, C.Y., 2016. PFR peptide, one of the antimicrobial peptides
457 identified from the derivatives of lactoferrin, induces necrosis in leukemia cells. *Sci
458 Rep* 6, 20823. doi:10.1038/srep20823
- 459 Ma, N.F., Lau, S.H., Hu, L., Xie, D., Wu, J., Yang, J., Wang, Y., Wu, M.C., Fung, J., Bai,
460 X., Tzang, C.H., Fu, L., Yang, M., Su, Y.A., Guan, X.Y., 2008. COOH-terminal
461 truncated HBV X protein plays key role in hepatocarcinogenesis. *Clin Cancer Res* 14,
462 5061–5068. doi:10.1158/1078-0432.CCR-07-5082
- 463 Mansfeld, J., Collin, P., Collins, M.O., Choudhary, J.S., Pines, J., 2011. APC15 drives the
464 turnover of MCC-CDC20 to make the spindle assembly checkpoint responsive to
465 kinetochore attachment. *Nat Cell Biol* 13, 1234–1243. doi:10.1038/ncb2347
- 466 Meng, M.X., Ning, J.F., Yu, J.Y., Chen, D.D., Meng, X.L., Xu, J.P., Zhang, J., 2014.
467 Antitumor activity of recombinant antimicrobial peptide penaeidin-2 against kidney
468 cancer cells. *J Huazhong Univ Sci Technol. Med Sci* 34, 529–534.
469 doi:10.1007/s11596-014-1310-4
- 470 Oh, D., Shin, S.Y., Lee, S., Kang, J.H., Kim, S.D., Ryu, P.D., Hahm, K.S., Kim, Y., 2000.
471 Role of the hinge region and the tryptophan residue in the synthetic antimicrobial
472 peptides, cecropin A(1-8)-magainin 2(1-12) and its analogues, on their antibiotic
473 activities and structures. *Biochemistry* 39, 11855–11864.

- 474 474 Otvos, L., 2000. Antibacterial peptides isolated from insects. *J Pept Sci* 6, 497–511.
475 475 doi:10.1002/1099-1387(200010)6:10<497::AID-PSC277>3.0.CO;2-W
- 476 476 Park, C., Lee, D.G., 2010. Melittin induces apoptotic features in *Candida albicans*.
477 477 *Biochem Biophys Res Commun* 394, 170-172.
- 478 478 Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., Yao, X., 2009. DOG 1.0: illustrator of protein
479 479 domain structures. *Cell Res* 19, 271–273. doi:10.1038/cr.2009.6
- 480 480 Ren, S.X., Cheng, A.S., To, K.F., Tong, J.H., Li, M.S., Shen, J., Wong, C.C., Zhang, L.,
481 481 Chan, R.L., Wang, X.J., Ng, S.S., Chiu, L.C., Marquez, V.E., Gallo, R.L., Chan, F.K.,
482 482 Yu, J., Sung, J.J., Wu, W.K., Cho, C.H., 2012. Host immune defense peptide LL-37
483 483 activates caspase-independent apoptosis and suppresses colon cancer. *Cancer Res* 72,
484 484 6512–6523. doi:10.1158/0008-5472.CAN-12-2359
- 485 485 Riedl, S., Zweytk, D., Lohner, K., 2011. Membrane-active host defense peptides--
486 486 challenges and perspectives for the development of novel anticancer drugs. *Chem
487 487 Phys Lipids* 164, 766–781. doi:10.1016/j.chemphyslip.2011.09.004
- 488 488 Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing
489 489 phylogenetic trees. *Mol Biol Evol* 4, 406–425.
- 490 490 Steiner, H., Hultmark, D., Engström, A., Bennich, H., Boman, H.G., 1981. Sequence and
491 491 specificity of two antibacterial proteins involved in insect immunity. *Nature* 292, 246–
492 492 248.
- 493 493 Suttmann, H., Retz, M., Paulsen, F., Harder, J., Zwergel, U., Kamradt, J., Wullich, B.,
494 494 Unteregger, G., Stöckle, M., Lehmann, J., 2008. Antimicrobial peptides of the
495 495 Cecropin-family show potent antitumor activity against bladder cancer cells. *BMC
496 496 Urol* 8, 5. doi:10.1186/1471-2490-8-5
- 497 497 Tajima, F., Nei, M., 1984. Estimation of evolutionary distance between nucleotide
498 498 sequences. *Mol Biol Evol* 1, 269–285.
- 499 499 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular
500 500 Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30, 2725–2729.
501 501 doi:10.1093/molbev/mst197
- 502 502 Wang, Y., Qu, L., Gong, L., Sun, L., Gong, R., Si, J., 2013. Targeting and eradicating
503 503 hepatic cancer cells with a cancer-specific vector carrying the Buforin II gene. *Cancer
504 504 Biother Radiopharm* 28, 623–630. doi:10.1089/cbr.2012.1469
- 505 505 Webb, B., Sali, A., 2014. Comparative Protein Structure Modeling Using MODELLER.
506 506 *Curr Protoc Bioinforma.* 47, 5.6.1–32. doi:10.1002/0471250953.bi0506s47

- 507 White, S.H., Wimley, W.C., 1998. Hydrophobic interactions of peptides with membrane
508 interfaces. *Biochim Biophys Acta* 1376, 339–352.
- 509 Wu, C., Geng, X., Wan, S., Hou, H., Yu, F., Jia, B., Wang, L., 2015. Cecropin-P17, an
510 analog of Cecropin B, inhibits human hepatocellular carcinoma cell HepG-2
511 proliferation via regulation of ROS, Caspase, Bax, and Bcl-2. *J Pept Sci* 21, 661–668.
512 doi:10.1002/psc.2786
- 513 Xu, D., Jaroszewski, L., Li, Z., Godzik, A., 2014. FFAS-3D: improving fold recognition
514 by including optimized structural features and template re-ranking. *Bioinformatics* 30,
515 660–667. doi:10.1093/bioinformatics/btt578
- 516 Yi, H.Y., Chowdhury, M., Huang, Y.D., Yu, X.Q., 2014. Insect antimicrobial peptides and
517 their applications. *Appl Microbiol Biotechnol* 98, 5807–5822. doi:10.1007/s00253-
518 014-5792-6
- 519 Zasloff, M., 2002. Antimicrobial peptides of multicellular organisms. *Nature* 415, 389–
520 395. doi:10.1038/415389a
- 521 Zhang, H.T., Wu, J., Zhang, H.F., Zhu, Q.F., 2006. Efflux of potassium ion is an important
522 reason of HL-60 cells apoptosis induced by tachyplesin. *Acta Pharmacol Sin* 27,
523 1367–1374. doi:10.1111/j.1745-7254.2006.00377.x
- 524
- 525
- 526

527 **Figure legends**

528

529 **Figure 1.** (A) Comparison of *Cfcec⁺*1 and (B) *Cfcec⁻*1 genes. The genes are composed of
530 two exons and one intron. Amino acids sequence of prepeptides are indicated below each
531 gene. SP, signal peptide; PP, propeptide. The figure was generated with DOG 1.0 (Ren et
532 al. 2009). (C) 3D model of anionic cecropin *Cfcec⁻*2. (D) Alignment of *Cfcec⁺* and *Cfcec⁻*
533 peptides. Green and blue arrows identify the C-terminal residue of the signal peptide (SP)
534 and the propeptide (PP), respectively. Black stars represent functionally important amino
535 acids. Muscle (Edgar 2004) was used to create multiple alignments. The figure was
536 prepared with ESPript (<http://escript.ibcp.fr>).

537

538

539 **Figure 2.** (A) Multiple sequence alignment and (B) phylogenetic relationship among
540 cecropins (based on amino acid sequences). Green letters: Coleoptera (represented by
541 *Acalolepta luxuriosa* (ACALU); blue letters: Diptera (represented by *Drosophila*
542 *melanogaster* (DROME), *Hermetia illucens* (HERIL) and *Culex quinquefasciatus*
543 (CULQU). Black letters: Lepidoptera (represented by *Antheraea pernyi* (ANTPE),
544 *Bombyx mori* (BOMMO), *Danaus plexippus* (DANPL), *Hyalophora cecropia* (HYACE)
545 *Manduca sexta* (MANSE), and the *C. fumiferana* cecropins identified in this work. Net
546 charge of each mature peptide is in brackets. The distinct anionic and cationic cecropin
547 clades are highlighted. The evolutionary history was inferred using the Neighbor-Joining
548 method. The percentage of replicate trees in which the associated cecropins clustered
549 together in the bootstrap test (1000 replicates) are shown next to the branches. The tree
550 was rooted using a coleopteran cecropin as outgroup. Evolutionary analyses were
551 conducted in MEGA6 (Tamura et al. 2013).The multiple alignment figure was prepared
552 with ESPript (<http://escript.ibcp.fr>).

553

554 **Figure 3.** HBx (hepatitis B virus protein), entericidins and APC15 (subunit anaphase-
555 promoting complex 15) sequences display similarities to *Cfcec* peptides. (A) Alignment
556 of *Cfcec⁺* and *Cfcec⁻* with HBx. The HBx-like motif is indicated by black stars. (B)
557 Alignment of *Cfcec⁺* and *Cfcec⁻* with entericidin A and B of *Thalassospira mesophila*
558 (Uniprot: A0A1Y2L2K6 and A0A1Y2L5A4). Green arrow identifies the C-terminal
559 residue of the signal peptide (SP). (C) A close-up of the superposition of 3D models of
560 the N-terminus of *Cfcec⁻* and the complex HBx-Bcl-2 (PDBid: 5FCG). N-termini of
561 *Cfcec⁻*, HBx and Bcl-2 are colored in orange, cyan and green, respectively. (D)
562 Alignment of *Cfcec⁻* and APC15. Highly conserved amino acid residues are shown in red
563 and boxed in blue. The sequence alignments and the image of 3D model were prepared
564 with ESPript (<http://escript.ibcp.fr>) and PyMol (www.pymol.org), respectively.

565

566 **Tableau 1.** Physico-chemical characteristics of HccecA, Cfcec⁺, Cfcec⁻, Pxcec⁻ and lunasin.

	Mature peptide	Length	Net charge	pI	Hydrophobicity (Kcal/mol)	Structure
HccecA	KWKLFKKIEKVQGNIRDGIIKAGPAVAVVGQATQIAK	37	+6	10.39	34.74	α-helix
Cfcec ⁺ 1	RWNPFKKLERVGQNI RDGIIKAAPAPA VAVVGQAAIAKG	38	+5	11.59	32.55	α-helix
Cfcec ⁺ 2	RWKPFKKLERVGQHIRDGIIKAGPAVQVVGQAATIAKG	38	+7	11.12	36.65	α-helix
Cfcec ⁻ 1	GRELERIGQQI RDGII SARPAL D VIRDQ KI YNGDDDDDDDK	43	-6	4.0	69.32	α-helix
Cfcec ⁻ 2	GRELEKIGQNVRD GIIKAGPAIEVIQKAQRIYHGKYDDDDDK	42	-1	5.54	62.63	α-helix
Pxcec ⁻	WNPFKELERAGQNI RDAIISAGPAV D VVVARAQK IARGEDVDEDE	44	-4	4.17	56.65	α-helix
Lunasin	SKWQHQ QDSCRKQLQGVNLTPCEKHIMEKIQGRGDDDDDDDD	43	-6	4.23	73.27	α-helix

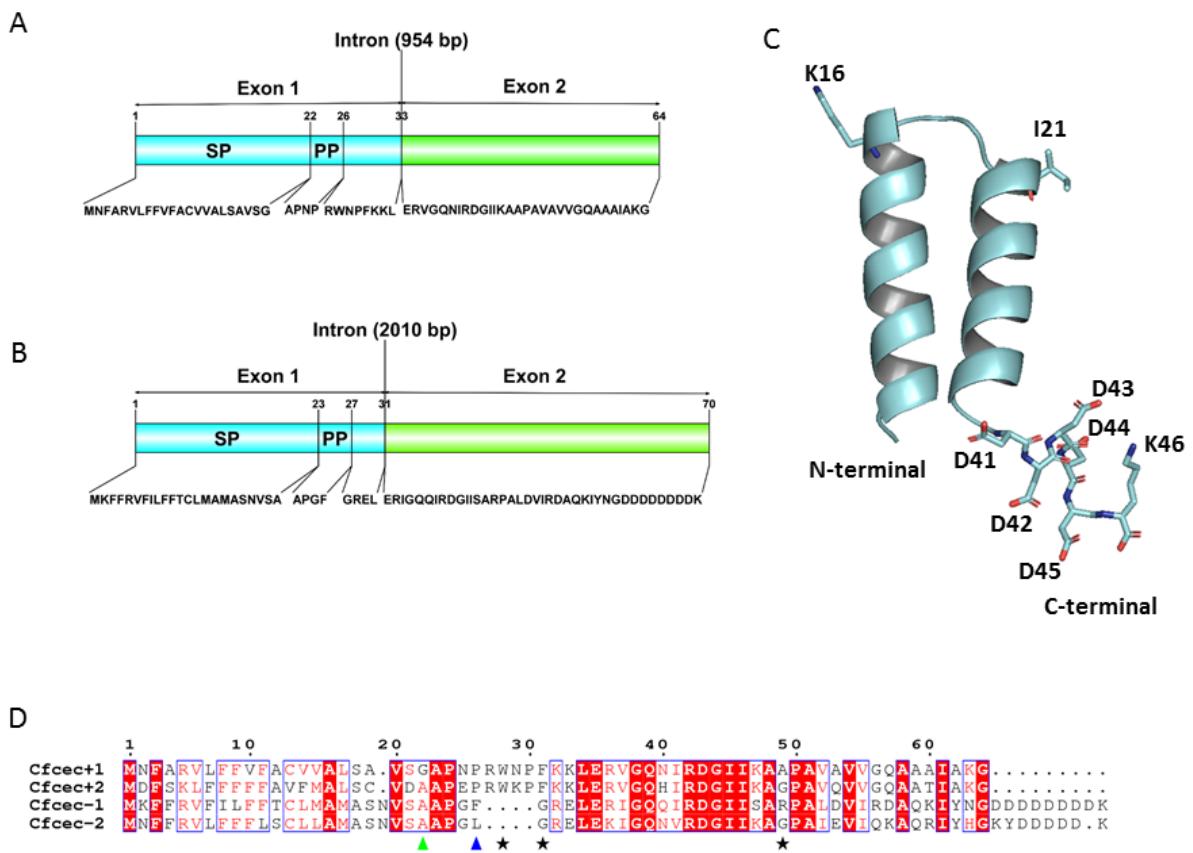
567

568 **Tableau 2.** AMPs with shared (G-[KQR]-[HKQNR]-[IV]-[KQR]) motif known to induce apoptosis.

Peptides	Source	sequences	References
Cecropin	<i>Musca domestica</i>	GWLKKIGKKIERV <u>GQHTRD</u> ATIQTIGVAQQAANVAATLKG	Jin et al. 2010
Cecropin-P17	<i>Hyalophora cecropia</i>	FKKKKKV <u>GRNIRNGIIK</u>	Wu et al. 2015
Papiliocin	<i>Papilio xuthus</i>	WKIFKKIEKV <u>GRNVRDGIIKAGPAVAVVGQAA</u> TVVKG	Hwang et al. 2011
Melittin	<i>Apis mellifera</i>	<u>GIGAVLKVLTTGLP</u> ALISWIK RKRQQ	Park et al. 2010
Arenicin-1	<i>Arinicola marina</i>	RWCVYAYVVR <u>VRGVLVRYRRCW</u>	Cho and Lee 2011
Coprisin	<i>Copris tripartitus</i>	VTCDVLSFEAKGIAVNHSACALHCIALRKGGSCQNG <u>VCVRN</u>	Lee et al. 2012
Pleurocidin	<i>Pleuronectes americanus</i>	<u>GWGSFFKKAAHV</u> GKHVGAALTHYL	Choi and Lee 2013
plant defensin RsAFP2	<i>Raphanus sativus</i>	QKLCQRPSGTWS <u>CGVCGNNNACKNQCIRLEKARHGSC</u>	Aerts et al. 2009
Psacothearin	<i>Psacothearia hilaris</i>	CIAKGNGCQPSGV <u>QGNCCSGHCHKEPGWVAGYCK</u>	Hwang et al. 2011b
Cathelicidin LL-37	<i>Homo sapiens</i>	<u>LLGDFFRKSKEKIGKEF</u> KRIVQRIKDFLRNLVPRTES	Ren et al. 2012
Magainin 2	<i>Xenopus laevis</i>	<u>GIGKFLHS</u> AKKFGKAFVGEIMNS	Lee and Lee 2014
Tachyplesin I	<i>Tachyplesius tridentatus</i>	KWCFRVCYRG <u>ICYRRCR</u>	Zhang et al. 2006
Buforin II	<i>Bufo gargarizans</i>	TRSSRAGLQFPVGRV <u>HRLLRK</u>	Wang et al. 2013
Chrysophsin-1	<i>Chrysophrys major</i>	<u>FFGWLIK</u> GAIHAGKAIHGLIHRHH	Hsu et al. 2011
Penaeidin-2a	<i>Penaeus vannamei</i>	YRGGYT <u>GPIPRPPIGRPPFRPV</u> CNACYRLSVDARNCCI <u>KFGSCCHLVK</u>	Meng et al. 2014
Piscidin 1	<i>Perca saxatilis</i>	FFHHIFRGIV <u>HVGKTIHRLVTG</u>	Lin et al. 2012
Epinecidin-1	<i>Epinephelus coioides</i>	<u>GFIFHI</u> IKGLFHAGKMIHGLV	Chen et al. 2009

569

570



571

572 Figure 1

573

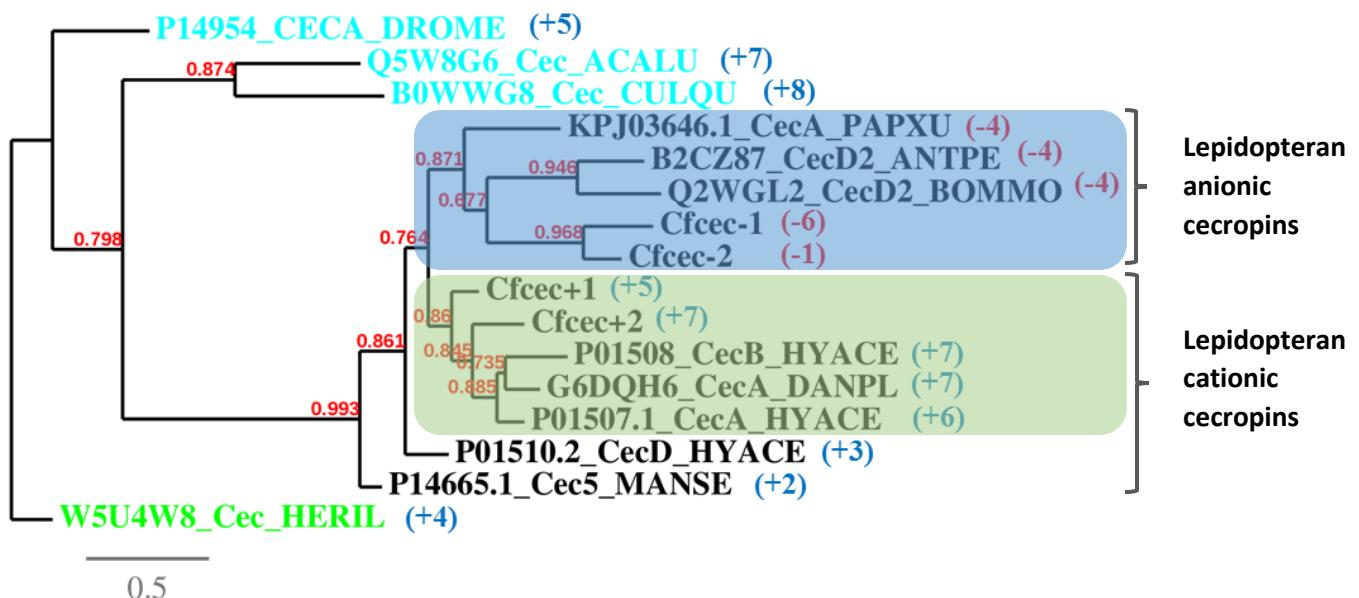
574 A

	1	10	20	30	40	50	60	70
KPJ03646.1 CecA_PAPXU	M	Y	V	I	L	F	V	F
B2CZ87 CecD2_ANTPE	M	N	S	V	R	I	L	I
Q2WGL2 CecD2_BOMMO	M	Y	F	T	K	I	V	F
Cfcec-1	M	K	F	R	V	I	L	F
Cfcec-2	M	N	F	R	V	I	L	F
P01510.2 CecD_HYACE	M	N	F	T	K	I	L	F
P14665.1 Cec5_MANSE	M	N	F	S	R	V	L	F
P01508 CecB_HYACE	M	N	F	S	R	I	F	F
Cfcec+2	M	N	F	S	R	I	F	F
Cfcec+1	M	N	F	A	R	V	L	F
G6DQH6 CecA_DANPL	M	D	F	S	K	I	F	F
P01507.1 CecA_HYACE	M	D	F	S	K	I	F	F
W5U4W8 Cec_HERIL	M	N	F	A	K	L	F	V
P14954 CECA_DROME	M	N	F	A	K	L	F	V
Q5W8G6 Cec_ACALU	M	N	F	A	K	L	F	V
B0WWG8 Cec_CULQU	M	N	F	N	K	L	F	V

575

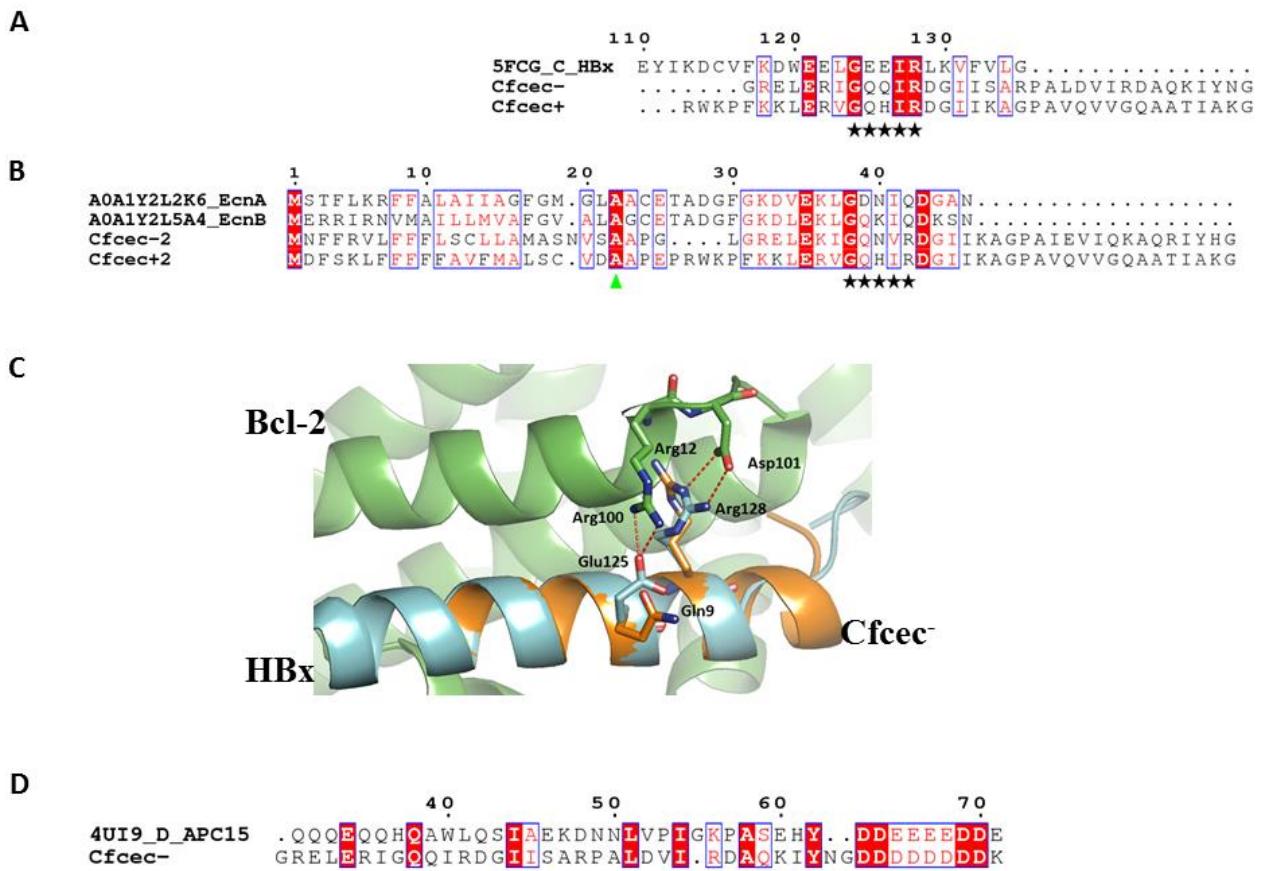
576

577 B



578

579 Figure 2



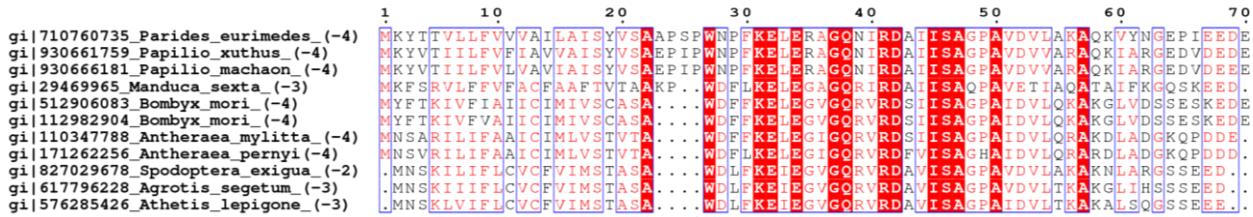
580

581 Figure 3

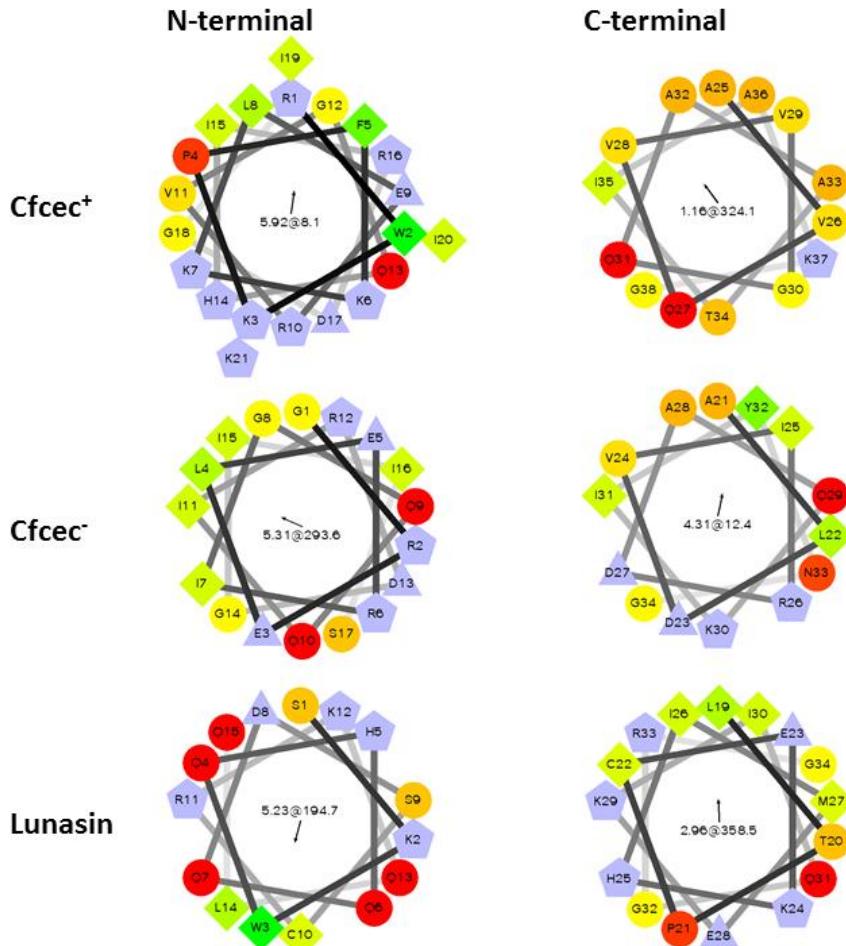
582

583 **Supplementary Material**

584
585



586
587 Figure S1. Lepidopteran cecropins with C-terminal poly-L-aspartate/glutamate.
588 Accession number, species name and net charge (in brackets) are at the beginning of
589 entry. Green and blue arrows identify the C-terminal amino acid of the signal peptide
590 (SP) and the propetide (PP), respectively.
591



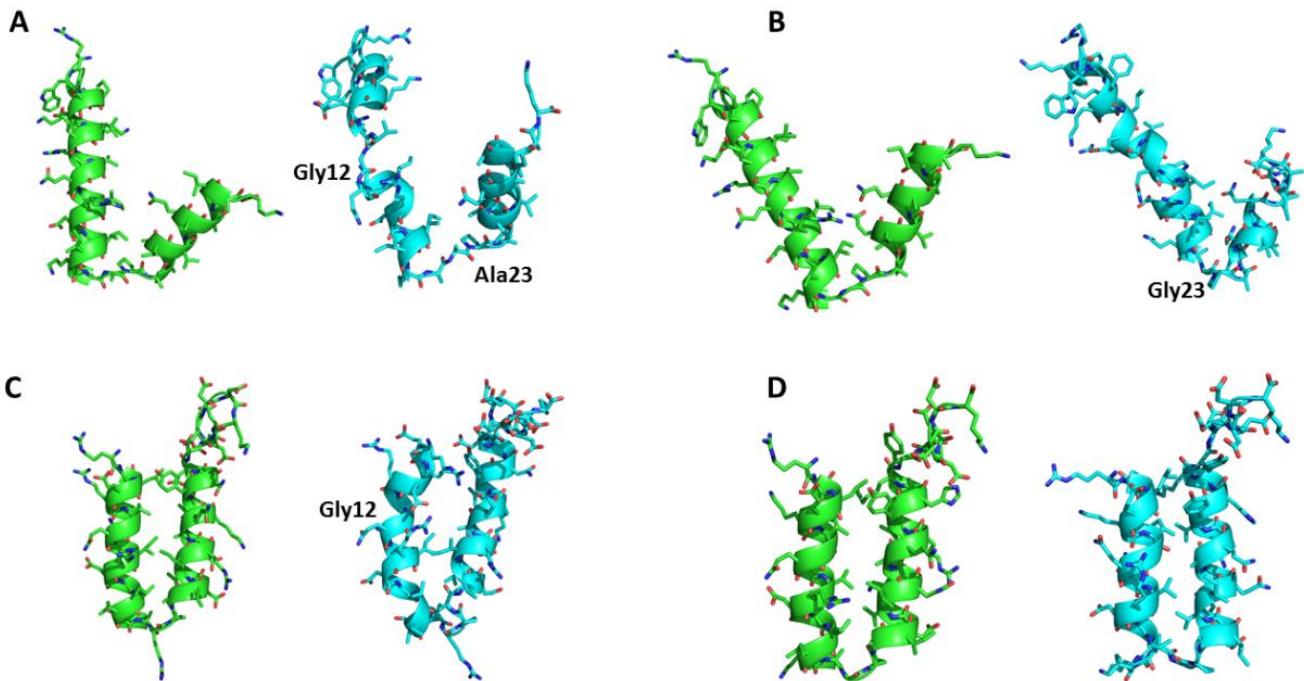
592

593 Figure S2. Helical-wheel diagram to illustrate the amphipathic properties of alpha helices
594 in Cfcec⁺, Cfcec⁻ and lunasin. The plot reveals whether hydrophobic amino acids are
595 concentrated on one side of the helix, usually with polar or hydrophilic amino acids on
596 the other side. The hydrophobic residues as diamonds, hydrophilic residues as red circles,
597 potentially negatively charged as triangles, and potentially positively charged as
598 pentagons. The value in the centre of each helix represents the mean amphipathic
599 moment $\langle \mu H \rangle$. The length and the direction of the $\langle \mu H \rangle$ vector depend on the
600 hydrophobicity and the position of the side chain along the helix axis. A large $\langle \mu H \rangle$
601 value means that the helix is amphipathic perpendicular to its axis.
602 <http://rzlab.ucr.edu/scripts/wheel/wheel.cgi>.

603

604

605



606

607 Figure S3. Snapshots of molecular dynamics (MD) simulation of the Cfcec⁺ and Cfcec⁻
608 peptides. These figures show that the flexible regions of peptides are situated at the Gly
609 and Ala residues. (A) Cfcec⁺1, (B) Cfcec⁺2, (C) Cfcec⁻1 and (C) Cfcec⁻2. Simulation time
610 of 0 ns (green structure) and 1 ns (cyan structure).