



Article

Aphid BCR4 Structure and Activity Uncover a New Defensin Peptide Superfamily

Karine Loth ^{1,2,†}, Nicolas Parisot ^{3,†}, Françoise Paquet ¹, Hugo Terrasson ³, Catherine Sivignon ⁴, Isabelle Rahioui ⁴, Mélanie Ribeiro Lopes ³, Karen Gaget ⁴, Gabrielle Duport ⁴, Agnès F. Delmas ¹, Vincent Aucagne ¹, Abdelaziz Heddi ³, Federica Calevro ⁴ and Pedro da Silva ^{3,*}

¹ Centre de Biophysique Moléculaire, CNRS UPR 4301, 45071 Orléans, France

² UFR Sciences et Techniques, Université d'Orléans, 45071 Orléans, France

³ Univ Lyon, INSA Lyon, INRAE, BF2I, UMR 203, 69621 Villeurbanne, France

⁴ Univ Lyon, INRAE, INSA Lyon, BF2I, UMR 203, 69621 Villeurbanne, France

* Correspondence: pedro.da-silva@insa-lyon.fr

† These authors contributed equally to this work.

Abstract: Aphids (Hemiptera: Aphidoidea) are among the most detrimental insects for agricultural plants, and their management is a great challenge in agronomical research. A new class of proteins, called Bacteriocyte-specific Cysteine-Rich (BCR) peptides, provides an alternative to chemical insecticides for pest control. BCRs were initially identified in the pea aphid *Acyrtosiphon pisum*. They are small disulfide bond-rich proteins expressed exclusively in aphid bacteriocytes, the insect cells that host intracellular symbiotic bacteria. Here, we show that one of the *A. pisum* BCRs, BCR4, displays prominent insecticidal activity against the pea aphid, impairing insect survival and nymphal growth, providing evidence for its potential use as a new biopesticide. Our comparative genomics and phylogenetic analyses indicate that BCRs are restricted to the aphid lineage. The 3D structure of BCR4 reveals that this peptide belongs to an as-yet-unknown structural class of peptides and defines a new superfamily of defensins.

Keywords: aphid BCR; 3D structure; defensin peptide; bioinsecticidal peptide; symbiosis



Citation: Loth, K.; Parisot, N.; Paquet, F.; Terrasson, H.; Sivignon, C.; Rahioui, I.; Ribeiro Lopes, M.; Gaget, K.; Duport, G.; Delmas, A.F.; et al. Aphid BCR4 Structure and Activity Uncover a New Defensin Peptide Superfamily. *Int. J. Mol. Sci.* **2022**, *23*, 12480. <https://doi.org/10.3390/ijms232012480>

Academic Editor: T. G. Emyr Davies

Received: 22 September 2022

Accepted: 12 October 2022

Published: 18 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Insects are among the most important pests of cultured plants and stored products, causing an estimated yearly loss of hundreds of millions of dollars worldwide [1–3]. Aphids, in particular, hold a prominent place among insect pests, as they represent up to 26% of the pests found on the main food crops (maize, wheat, potatoes, sugar beet, barley and tomatoes) grown in temperate climates [4,5]. Aphid control strategies rely almost exclusively on chemical treatments, which cause persistent environmental pollution and lead to the emergence of insect resistance [6]. New ecologically friendly solutions are therefore required to control aphids and other phloem-feeders.

Small Disulfide bond-Rich Proteins (DRPs) extracted from plants or arthropods are promising alternative biopesticide molecules [7–9]. These naturally occurring molecules display a very broad range of biological activities, mainly related to host-defense processes [10], and show a high structural and chemical diversity. Their exceptional stability is particularly appealing for drug discovery purposes [11]. The rather rigid three-dimensional conformations imposed by their polycyclic architecture confer DRPs with a strong resistance towards *in vivo* enzymatic degradation. Taken together, these features contribute to establishing DRPs as an emerging lead compound for the development of novel peptide-based drugs and, more recently, as potential biopesticides in agronomical research [7,8,12,13]. For instance, a knottin DRP extracted from pea seeds, PA1b (Pea Albumin 1, subunit b, 37 amino acids (aa), three disulfide bonds) [9,12,14], is toxic to numerous insects, including aphids, cereal weevils, mosquitos and moths [12].

A new class of DRP, called Bacteriocyte-specific Cysteine-Rich (BCR) peptides, has recently been identified in a major crop pest, the pea aphid, *Acyrtosiphon pisum* [15]. Similar to many other crop pest insects that thrive on unbalanced diets, aphids have evolved long-lasting relationships with endosymbiotic bacteria, and almost all aphids are found in association with the γ 3-proteobacterium *Buchnera aphidicola* [16–18]. This bacterium supplements the host diet with nutrients that are lacking or limited in their habitats, thereby allowing insects to proliferate, causing major economic, social, and health damage. Neither the host nor the endosymbionts can survive independently. *B. aphidicola* is non-culturable, and insects artificially deprived of their endosymbionts (aposymbiotic) cannot survive or reproduce [19]. The maintenance of this association relies on the compartmentalization of endosymbionts in specialized insect cells, called bacteriocytes [20]. BCRs are encoded by seven orphan genes, and are all expressed exclusively in bacteriocytes of both embryonic and adult aphids [15]. This suggests that BCRs may play a role in bacteriocyte homeostasis, presumably in endosymbiont control, as previously described for antimicrobial peptide coleopteracin A in the cereal–weevil endosymbiosis [21]. Consistent with this hypothesis, it has been shown that BCR1, BCR2, BCR3, BCR4, BCR5 and BCR8 exhibit antimicrobial activity or can permeabilize the membrane of *E. coli* cells [22].

Each BCR peptide consists of a secretory signal peptide and a mature peptide (Figure 1) composed of 44–84 amino acids and containing from six cysteine residues in the case of BCR1-2-4-5 and BCR8 to eight in the case of BCR6 [15]. Between BCR1, BCR2, BCR4 and BCR5, the cysteine-rich region is highly divergent, but the six cysteines have almost identical spacing in the predicted proteins. Three of these peptides, BCR1, BCR4 and BCR5, are encoded by genes found within a genomic region of 20 kbp, suggesting that they may have arisen due to a recent tandem gene duplication. No similarity was found between the other BCR family genes [15]. Intriguingly, the pea aphid BCRs show no significant sequence similarity with genes in species outside the aphid lineage.

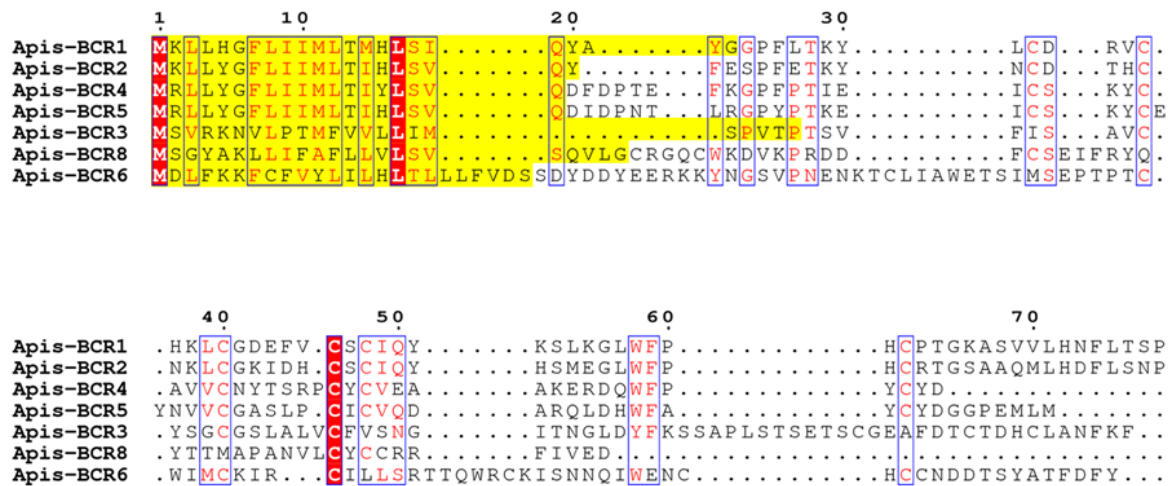


Figure 1. Multiple sequence alignment of *Acyrtosiphon pisum* BCRs. The number and the spacing between these cysteine residues allow to distinguish four BCR sub-families: BCR1-2-4-5, BCR3, BCR8 (containing all six cysteines) and BCR6 (eight cysteines). The signal peptides predicted by SignalP 6.0 server [23] of BCR sequences are highlighted in yellow. The fully conserved residues are framed and highlighted in red.

DRPs have undergone extensive divergent evolution in their sequence structure and function. They are classified based on their secondary structure orientation, cysteine distribution across the sequence, structure similarities in the disulfide bond patterns and precursor gene sequence [10,24]. Based on the sequence analysis, it was assumed that they could have evolved from defensin-type antimicrobial peptides (AMPs), small proteins containing six to eight cysteines, which are universally found in both animals and plants [15,22].

In this study, we focused on BCR4 of *A. pisum*, a benchmark example of a peptide with antimicrobial activities [22]. We discovered that BCR4 has strong insecticidal activity against the pea aphid. Taking advantage of the recent sequencing of several aphid genomes, which enables the study of gene family diversification through comparative and evolutionary analyses [25,26], we conducted a comparative genomic analysis across 22 aphid species for which sequence information was available. This allowed the identification of 76 new BCR sequences, all restricted to aphid species and not related to any known defensins. Finally, we determined the 3D structure of BCR4 and showed that it belongs to a new structural class of disulfide-rich proteins. Overall, the biochemical analyses, the evolutionary history, and the 3D structure of BCR4 give significant insight into the biological and structural properties of BCRs, and provide evidence for the use of this new defensin superfamily as potential new biopesticides.

2. Results

2.1. Total Synthesis of BCR4

To investigate the biological activity of members of the BCR peptide family and to resolve their 3D structure by NMR, a pure sample of synthetic BCR4 was produced through total chemical synthesis. As standard automated Fmoc/tBu solid-phase peptide synthesis (SPPS) was unsuccessful (Supplementary Materials Figures S1 and S2, Table S1), we turned to a native chemical ligation (NCL)-based approach [27], relying on the assembly of two medium-sized peptide segments. Using a recently developed methodology [28–33], we coupled a 20-amino-acid (aa) N-2-hydroxy-5-nitrobenzylcysteine (N-Hnb-Cys) cryptothioester with a 30-aa cysteinyl peptide and obtained the full-length reduced form of BCR4 at high purity (Figure 2). Oxidative folding under thermodynamic control was achieved using a standard protocol [34,35], leading to one major compound featuring three disulfide bridges as evidenced by HRMS analysis (Supplementary Materials Figures S3–S11).

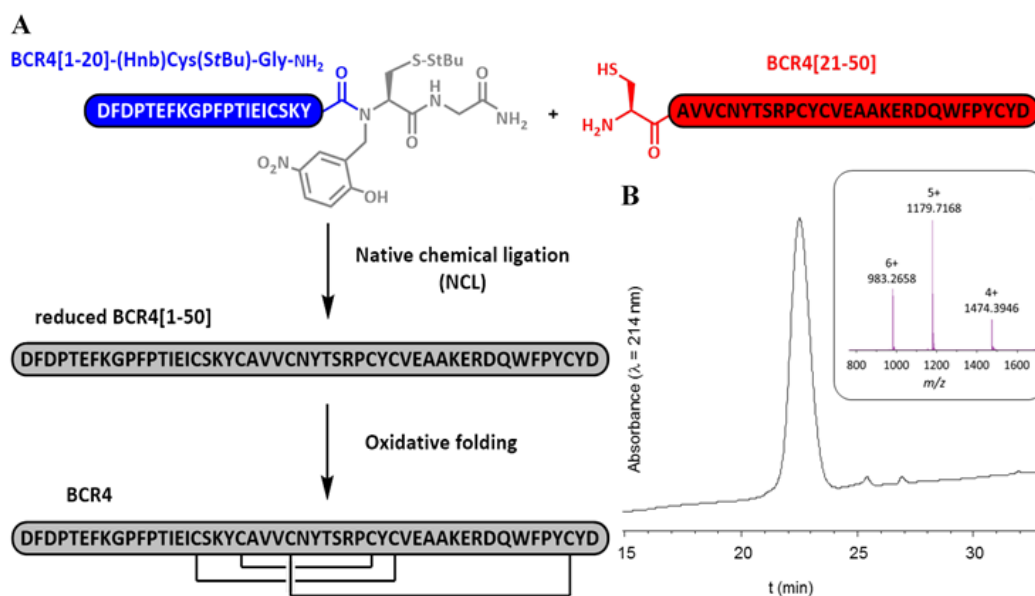


Figure 2. Chemical synthesis of BCR4. (A) Schematic representation of the BCR4 peptide chemical through native chemical ligation (NCL) of two peptide segments BCR4[1-20] in blue and BCR4[21-50] in red, followed by oxidative folding to form the three disulfide bridges. (B) RP-HPLC chromatogram and ESI-HRMS mass spectrum of the purified folded peptide.

2.2. Antimicrobial Activities of BCR4

To assess the antimicrobial activity of BCR4, we tested the effect of various concentrations (ranging from 5 μM to 80 μM) of this peptide on the growth of the Gram-negative bacterium *Escherichia coli* (strain NM522) and the Gram-positive bacterium *Micrococcus*

luteus. Consistent with previous observations [22], a slight inhibition of *E. coli* growth was detected at 5 µM BCR4. This antimicrobial activity increased with BCR4 concentration, and we determined a minimal inhibitory concentration (MIC) of 17.0 ± 2.4 µM, for which no bacterial growth was detected. Comparatively, no antimicrobial activity was detected against the Gram-positive bacterium *Micrococcus luteus* (Table 1).

Table 1. Antimicrobial activities of BCR4 peptide.

	MIC ^a (µM)
<i>Escherichia coli</i> NM522	17.0 ± 2.4
<i>Micrococcus luteus</i>	>80

^a Minimal inhibitory concentration.

2.3. Insect Bioassays

The insecticidal potential of BCR4 was assayed by oral administration of various concentrations of this peptide to the pea aphid and monitoring of survival. For all BCR4 concentrations tested (5–80 µM), aphid mortality was always significantly higher in the BCR4 aphids compared to the control group, demonstrating a specific effect of BCR4 ingestion on aphid survival (Figure 3). This effect was dependent on BCR4 dosage and duration of treatment. The highest concentration tested (80 µM) also had the strongest lethal effect, with a Lethal Time 50% (LT50) of 1.16 days (Table 2) and no surviving aphids after two days of BCR4 treatment (Figure 3).

Table 2. Toxicity of BCR4 peptide on the pea aphid *Acyrtosiphon pisum* (mean values of Lethal Time 50 (LT50) in days, with confidence intervals), analyzed by survival analysis with a log-normal fit.

	Concentration (µM)						
BCR4 peptide	80	40	35	27.5	20	10	5
LT50 (days)	1.16 [0.98–1.37]	1.76 [1.27–2.42]	1.94 [1.47–2.57]	3.24 [2.22–4.72]	3.48 [2.28–5.31]	11.3 [4.16–30.8]	>20

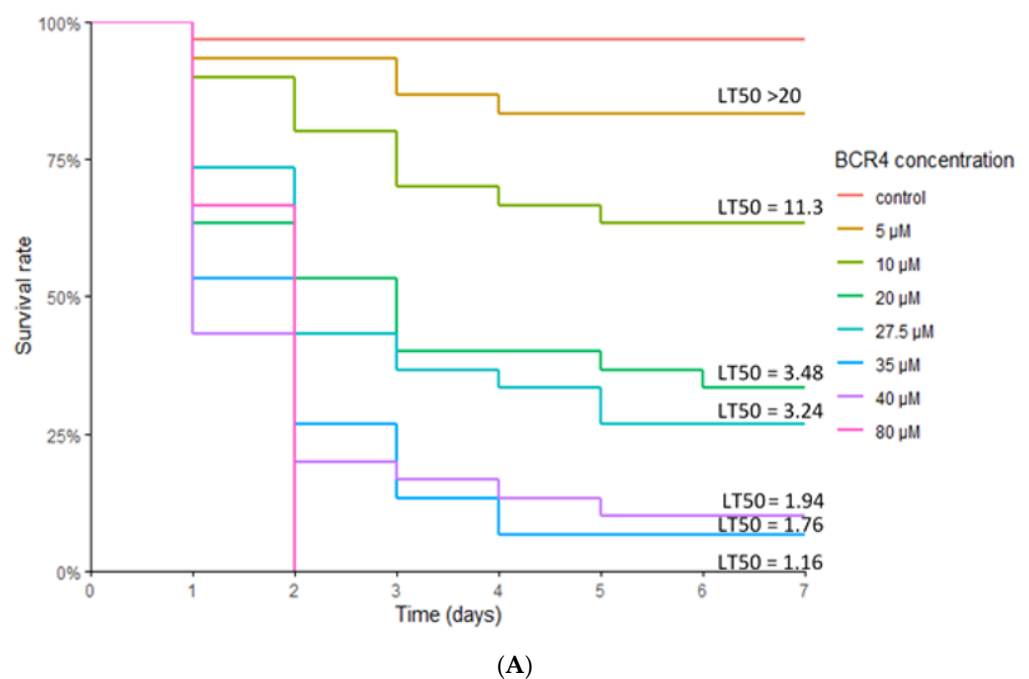


Figure 3. Cont.

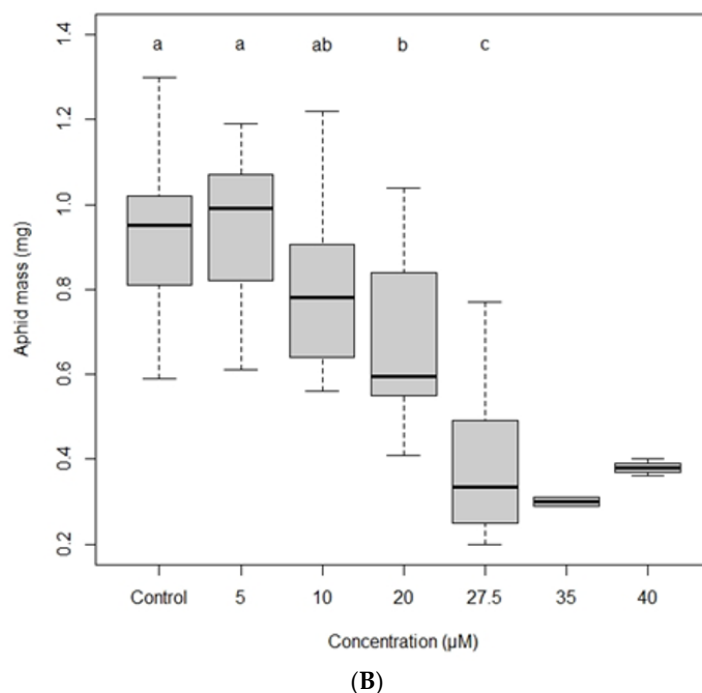


Figure 3. (A) Survival curves of the aphid *Acyrthosiphon pisum* reared on artificial diets containing different concentrations of BCR4 peptide. Mean values of Lethal Time 50 (LT50), in days, are indicated above each curve. (B) Mass (mg) of 7-day-old pea aphid *Acyrthosiphon pisum* subjected to BCR4 treatment. Concentrations labeled with different letters are significantly different ($p < 0.05$). The surviving aphid sample sizes of BCR4 at 35 and 40 μM were too small to be statistically meaningful.

Even for concentrations as low as 20 μM , aphids, on average, did not survive more than 4 days, and only one-third reached adulthood. Apart from this effect on aphid survival, the most striking phenotypical effect observed with BCR4 treatment was a statistically significant growth inhibition of the surviving aphids for doses higher than 5 μM , with, for instance, a 60% weight reduction 7 days after ingestion of BCR4 at 27.5 μM (Figure 3B).

2.4. Phylogeny of the BCR Family

To decipher the molecular evolution of the BCR peptides, we aimed to identify the whole set of proteins homologous to the seven *A. pisum* BCR proteins. Through an exhaustive search of the NCBI nucleotide and protein databases, coupled with specific inspection of the sequenced aphid species available in the AphidBase database [36], we retrieved a total of 76 new BCR sequences across 20 aphid species (Table S2) of the 22 for which sequences are available in publicly available databases. We also found one additional sequence from *A. pisum*, bringing the total number of BCRs in this insect to eight. Interestingly, all these sequences have been found in members of the Aphidoidea super-family, thus confirming that BCR proteins are restricted to the aphid lineage [15]. The complete set of 83 BCR sequences was used for amino acid sequence alignment and phylogenetic tree reconstruction (Figure 4). Based on these, the BCR sequences can be grouped into four subfamilies, including homologs of the BCR1-2-4-5, BCR3, BCR6 and BCR8 sequences of *A. pisum*, respectively. The majority of subfamilies have six cysteine residues except for the BCR6 subfamily, which has eight cysteines. The spacing between cysteines within the BCR sequences is fully conserved within subfamilies, thus supporting within-family homologies and phylogenetic positioning (Table 3).

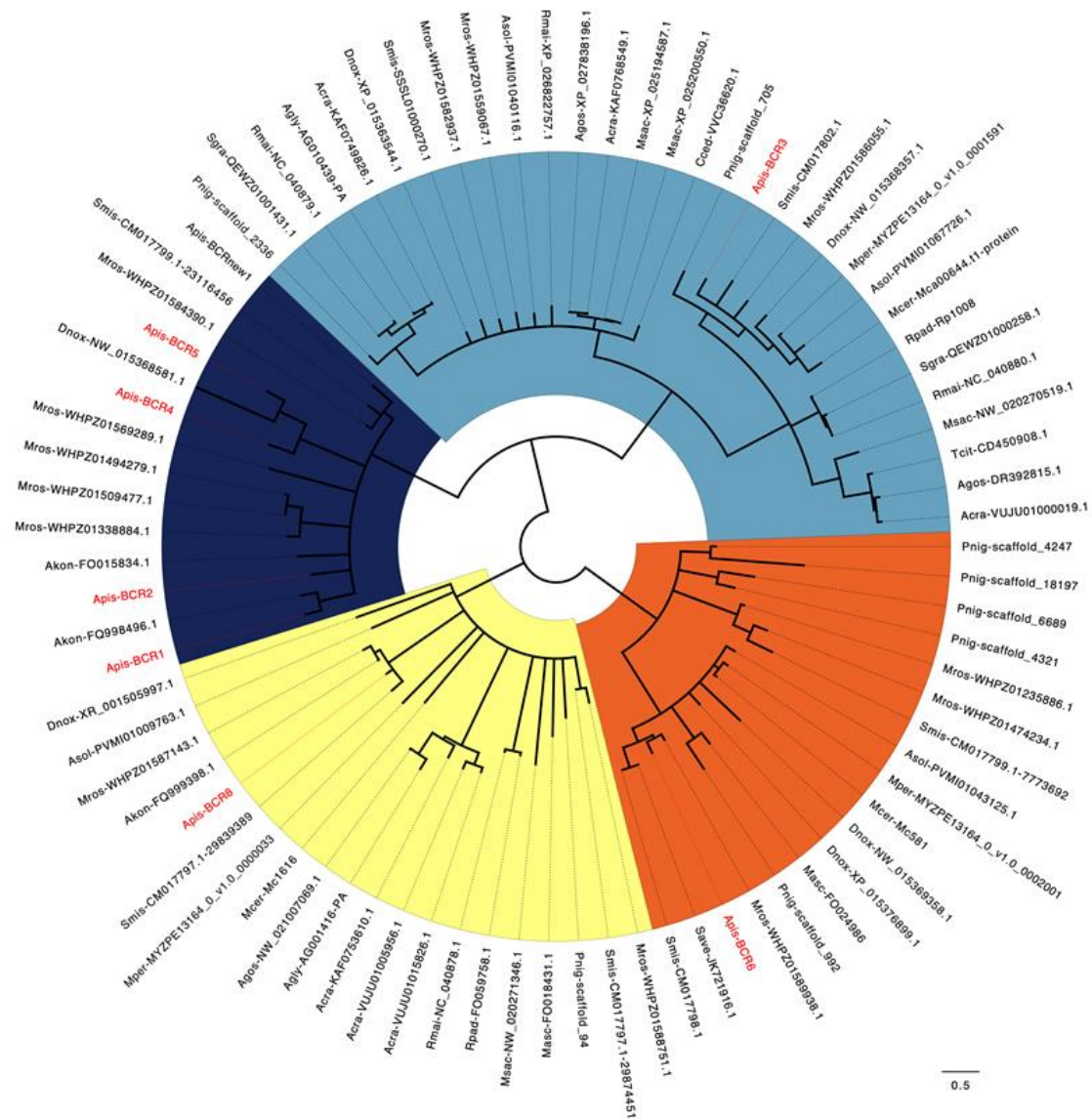


Figure 4. Phylogenetic trees of BCR proteins. Maximum likelihood phylogenetic tree reconstruction was performed using PhyML [37] with an LG 4-rate class model. Branch-support values were calculated using the bootstrap method, with 1000 replicates. Poorly supported branches (<50%) were collapsed using TreeCollapseCL4 [38]. The sequences used for the phylogenetic analysis are listed in Supplementary Materials Table S2. Sequences labeled in red reflect those identified by the Shigenobu and Stern [15] study in the pea aphid (*A. pisum* genome V3.0). Abbreviations: Acra, *Aphis craccivora*; Agly, *Aphis glycines*; Agos, *Aphis gossypii*; Akon, *Acyrtosiphon kondoi*; Apis, *Acyrtosiphon pisum*; Cced, *Cinara cedri*; Dnox, *Diuraphis noxia*; Dvit, *Daktulosphaira vitifoliae*; Masc, *Myzus ascalonicus*; Mcer, *Myzus cerasi*; Mper, *Myzus persicae*; Msac, *Melanaphis sacchari*; Rmai, *Rhopalosiphum maidis*; Rpad, *Rhopalosiphum padi*; Save, *Sitobion avenae*; Sgra, *Schizaphis graminum*; Tcit, *Toxoptera citricida*.

Table 3. Characteristics of the four BCR subfamilies and of the defensin families.

	Number of Cysteines	Consensus Motif	Cysteine Bonding Pattern	Number of Orthologous Sequences in Aphid Species ^a
BCR1-2-4-5	6	CX ₃ CX ₃₋₅ CX ₄₋₆ CXCX ₁₁₋₁₃ C	1:5, 2:4, 3:6	Akon (2), Apis (5), Mros (4), Smis (1), Mros (1), Dnox (1)
BCR3	6	CX ₃ CX ₆ CX ₂₁₋₂₅ CX ₆ CX ₃ C	Unknown	Acra (3), Agly (1), Agos (2), Apis (1), Asol (2), Cced (1), Dnox (2), Mcer (1), Mper (1), Mros (3), Msac (3), Pnig (2), Rma (3), Rpad (1), Sgra (2), Smis (2), Tcit (1)
BCR6	8	CX ₁₅ CX ₃ CX ₃ CX ₁₀ CX ₁₀ CXCC	Unknown	Apis (1), Asol (1), Dnox (2), Masc (1), Mcer (1), Mper (1), Mros (3), Pnig (5), Save (1), Smis (2)
BCR8	6	CX ₃₋₄ CX ₁₀ CX ₁₅₋₁₇ CXCC	Unknown	Akon (1), Acra (3), Agly (1), Agos (1), Apis (1), Asol (1), Dnox (1), Masc (1), Mcer (1), Mper (1), Msac (1), Mros (2), Pnig (1), Rmai (1), Rpad (1), Smis (2)
CSαβ	6 8	CX ₆₋₁₅ CX ₃ CX ₉₋₁₀ CX ₄₋₇ CXC CX ₁₀ CX ₅ CX ₃ CX ₉₋₁₀ CX ₆₋₈ CXCX ₃ C	1:4, 2:5, 3:6 1:8, 2:5, 3:6, 4:7	insects plants
α-defensin	6	CXCX ₃₋₅ CX ₉ CX ₆₋₁₀ CC	1:6, 2:4, 3:5	primates
β-defensin	6	CX ₅₋₇ CX ₃₋₅ CX ₈₋₁₁ CX ₄₋₆ CC	1:5, 2:4, 3:6	vertebrates

^a Abbreviations: Acra, *Aphis craccivora*; Agly, *Aphis glycines*; Agos, *Aphis gossypii*; Apis, *Acyrtosiphon pisum*; Asol, *Aulacorthum solani*; Cced, *Cinara cedri*; Dnox, *Diuraphis noxia*; Mcer, *Myzus cerasi*; Mper, *Myzus persicae*; Msac, *Melanaphis sacchari*; Mros, *Macrosiphum rosae*; Pnig, *Pentalonia nigronervosa*; Rmai, *Rhopalosiphum maidis*; Rpad, *Rhopalosiphum padi*; Save, *Sitobion avenae*; Sgra, *Schizaphis graminum*; Smis, *Sitobion miscanthi*; Tcit, *Toxoptera citricida*. No BCRs were found in *Eriosoma lanigerum* and *Sipha flava*.

Nineteen aphid species have full genome sequences available (*Acyrtosiphon pisum*, *Aphis craccivora*, *Aphis glycines*, *Aphis gossypii*, *Aulacorthum solani*, *Cinara cedri*, *Diuraphis noxia*, *Eriosoma lanigerum*, *Macrosiphum rosae*, *Melanaphis sacchari*, *Myzus cerasi*, *Myzus persicae*, *Pentalonia nigronervosa*, *Rhopalosiphum maidis*, *Rhopalosiphum padi*, *Schizaphis graminum*, *Sipha flava*, *Sitobion avenae*, *Sitobion miscanthi*). This makes it possible to predict the complete set of BCR peptides encoded by those aphid genomes. In this group, we observed a high variability in the number and distribution of BCR peptides among the four subfamilies (Table 3). No BCRs were found in the *E. lanigerum* (Eriosomatinae subfamily) and *S. flava* (Chaitophorinae subfamily) genomes, and only one in *C. cedri* (Lachninae subfamily) (Figure 5). All members of the Aphidinae subfamily appear to have at least two BCR genes, and *M. rosae* has the largest repertoire of BCR sequences, with 13 distinct sequences distributed across the four BCR subfamilies. While members of the BCR3 and BCR8 subfamilies are present in all Aphidinae species included in this study, members of the BCR1-2-4-5 and BCR6 subfamilies appear restricted to the Macrosiphini aphid tribe, suggesting that the genes encoding those BCRs arose from duplication of the former.

2.5. BCR4 Solution Structure

We next determined the 3D structure of BCR4 by NMR spectroscopy. The ¹H NMR and the natural-abundance ¹H-¹⁵N sofast-HMQC spectra [39] of the protein showed a good dispersion of the amide chemical shifts, which is indicative of highly structured peptides. Following standard procedures, the analysis of the set of 2D-TOCSY and NOESY spectra allowed a complete assignment of ¹H chemical shifts. This assignment was facilitated by heteronuclear ¹H-¹⁵N and ¹H-¹³C NMR spectra, particularly in crowded regions of the ¹H TOCSY and NOESY spectra corresponding to side chains (BRMB entry 34197). The 3D structures were calculated by considering a total of 923 distance restraints, 16 hydrogen bonds, 88 dihedral angles, and three ambiguous disulfide bridges (Table 4).

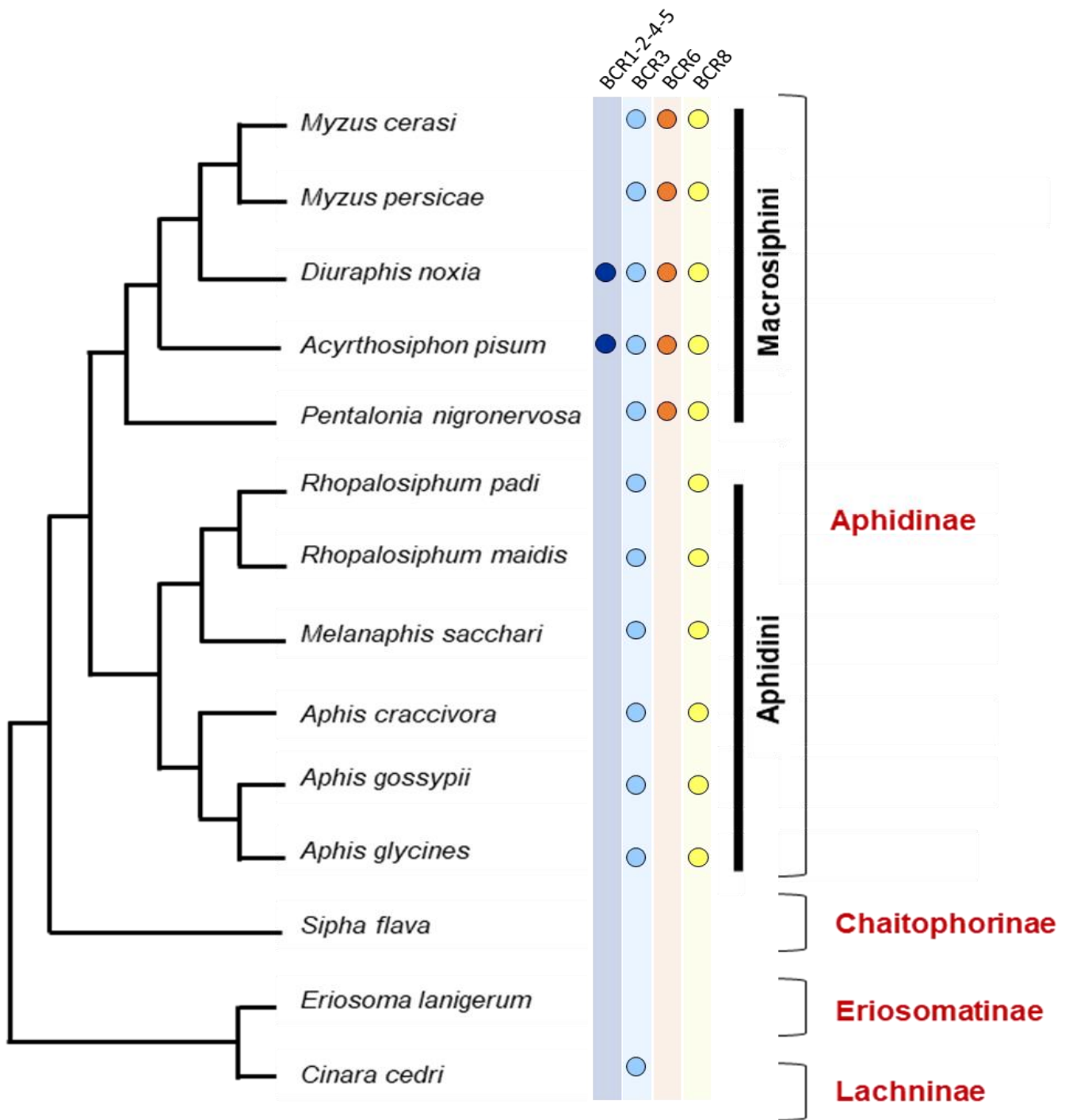


Figure 5. Partitioning of BCR peptides in aphid taxonomic groups. Phylogeny of the 14 members of the aphid group whose genome is annotated and available in databases (among the 19 with sequenced genomes). Their distributions into subfamilies and tribes are indicated in red and black, respectively (adapted from Calevro et al. [4]). Colored dots near each species name indicate whether a member of each BCR subfamily is present in the genome.

Table 4. NMR constraints and structural statistics.

NMR Restraints	
<i>Distance Restraints</i>	
Total NOE	923
Unambiguous	763
Ambiguous	160
Hydrogen bonds	16
<i>Dihedral Angle Restraints</i>	88
<i>Disulfide bridges</i> ^a	3
Structural Statistics (7PQW.pdb)	
<i>Average Violations Per Structure</i>	
NOEs $\geq 0.3\text{\AA}$	0
Hydrogen bonds $\geq 0.3\text{\AA}$	0
Dihedrals $\geq 15^\circ$	3
Dihedrals $\geq 10^\circ$	8
Average pairwise rmsd (\AA)	backbone atoms 0.45 ± 0.14 heavy atoms 1.14 ± 0.15
Ramachandran Analysis	
Most favored region and allowed region	96.9
Generously allowed	0.7
Disallowed	2.4
Energies ($\text{kcal}\cdot\text{mol}^{-1}$)^b	
Electrostatic	-1631.48 ± 29.53
Van der Waals	-334.18 ± 11.14
Total energy	-1196.56 ± 31.80
Residual NOE energy	56.38 ± 5.28

^a Introduced as ambiguous; ^b values are given as mean \pm standard deviation ($n = 20$).

The 200 water-refined structures of BCR4 possess three disulfide bridges with identical pairings: C17–C34, C21–C32 and C25–C48. Among them, 20 structures were selected that were in agreement with all NMR experimental data and the standard covalent geometry. Coordinates were deposited as PDB entry 7PQW. Analysis of the 20 final structures with PROCHECK-NMR [40] showed that almost all of the residues (96.9%) are in the most favored or allowed regions of the Ramachandran diagram (Table 4).

BCR4 is folded into a compact globular unit consisting of an N-terminal tail (D1-T13) and an α -helix (I14-V24), followed by an antiparallel β -sheet composed of two short β -strands, β_1 C34-A37 and β_2 Q43-P46 (Figure 6). The 3D structure is stabilized by three disulfide bridges linking the α -helix to β_1 (C17-C34), to the loop C25-Y33 (C21-C32) and to the C-terminal tail (C25-C48), respectively.

A comparison of the BCR4 structure with 3D structures deposited in the PDB using the DALI server [41] identified only one structure with a somewhat similar fold, as defined by a low Z score of 2.1. This structure corresponds to a chimera peptide, which is not a natural peptide (pdb code 1WT7, named BUTX-MTX [42]). This structural comparison thus identifies the 3D structure of BCR4 as being a new fold, which is consistent with it being unique to the aphid lineage.

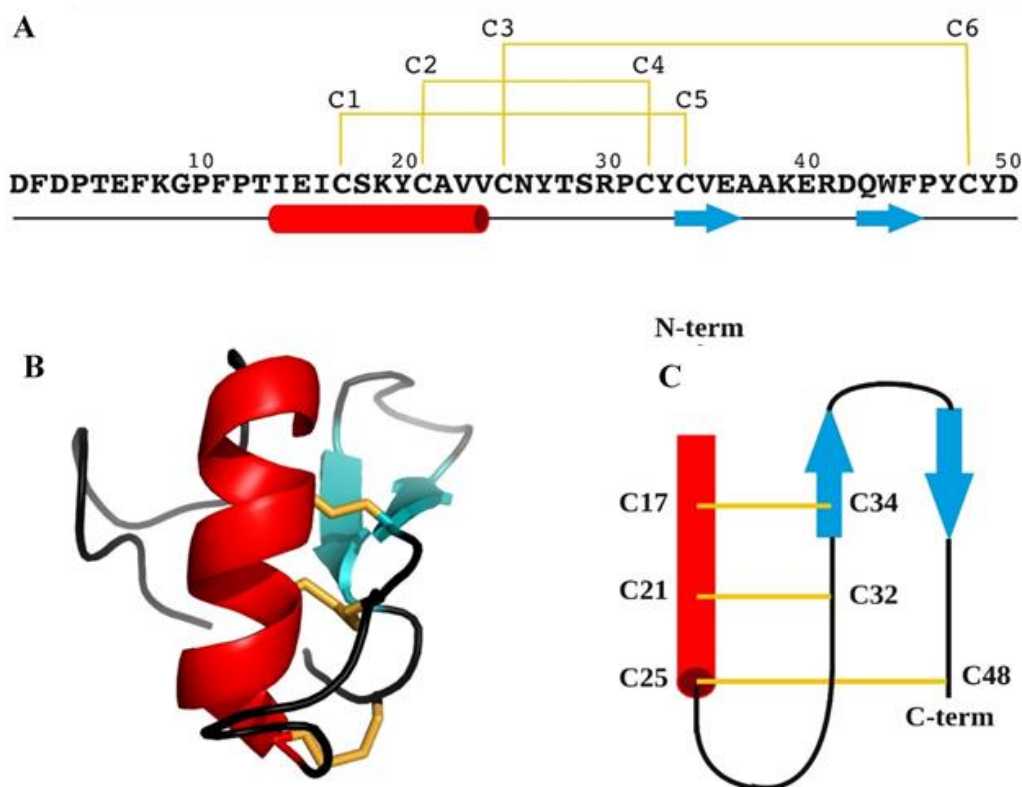


Figure 6. (A) Primary structure of BCR4 peptide. (B) Three-dimensional structure of BCR4. (C) Topology diagram of BCR4. α -helix, β -sheet, and random coil are represented in red, cyan and black, respectively. Disulfide bonds are colored in yellow.

3. Discussion

Thus far, most pest control strategies have relied on the use of systemic chemical pesticides, which are increasingly stigmatized because of their persistence in the environment and their toxicity to non-target organisms [6]. Thus, there is a need for new pest management solutions. Small Disulfide-Rich Proteins (DRPs) extracted from plants or arthropods are promising alternative biopesticide molecules [7,8,12]. Genomic plasticity of DRP-encoding sequences is known to foster the adaptability of organisms and to enable the acquisition of new functions [10]. In this work, we successfully produced in sufficient quantity and purity the folded form of BCR4 (Figure 2), a DRP encoded by the pea aphid genome and part of the BCR family of peptides, which was used to explore the insecticidal and structural properties of this protein. We showed that BCR4 strongly interferes with survival and growth of *A. pisum* in a dose-dependent manner (Figure 3). Its range of activity (5–80 μ M, Figure 3) is similar to that displayed by PA1b, a promising plant biopesticide [9,12] active on the same insect target [43]. Importantly, from a functional perspective and as previously reported [22], BCR4 has a significant bactericidal effect on *E. coli* (Table 1), a free-living relative of *B. aphidicola*, the obligatory aphid endosymbiont [22]. Based on these results, we propose that the ingestion of exogenous BCR may block nymphal development and induce aphid death by interfering with the population density of this endosymbiont, essential for nymphal growth and survival [19].

From an evolutionary perspective, large-scale homology searches through genomic, transcriptomic and proteomic databases were performed to complete the repertoire of BCR peptides, which to date have been limited to the seven sequences identified in the pea aphid genome. Importantly, the 76 additional sequences we found (bringing the total number of BCR sequences identified to 83) are all encoded by aphid genomes (Figure 4). Phylogenetic analysis showed that aphid BCRs are organized into four subfamilies, including BCR1-2-4-5, BCR3, BCR6 and BCR8 sequences (Figure 3), respectively. BCRs from those four subfamilies

all have six or eight cysteine residues, and we observed a very clear intra-group conservation of cysteine topology (e.g., distribution in the sequence and disulfide pairing). However, as for other DRPs, BCRs present high sequence diversity in their inter-cysteine loops, preventing the detection of any firm homology with arthropod defensins at this point [10,24]. While BCR subfamilies had previously been reported [15], we were able to enrich each of them with many new members (Figure 4). We also showed that (i) aphids present varying numbers of BCR-encoding genes, from 1 to 13; and (ii) while BCR3 and BCR8 homologs are widely present in the aphid lineage, genes from the BCR1-2-4-5 and BCR6 subfamilies are restricted to the Macrosiphini tribe, which includes many major agricultural pests. This suggests a complex evolutionary history involving several events of gene duplication and losses with possible functional diversification of the resulting homologs.

The three-dimensional structure of BCR4 was determined by NMR spectroscopy, and subsequent protein fold analysis revealed that BCR4 belongs to an as-yet-unknown structural class of defensin proteins (Figure 6). Defensins are a well-characterized group of DRPs present in all eukaryote genomes [24]. Shaffee's phylogenetic work from Shaffee et al. [10,24] showed that defensins consist of two independent and convergent superfamilies, each derived from independent evolutionary events: the so-called trans-defensin superfamily is uniquely composed of the CS $\alpha\beta$ (cysteine-stabilized α -helix β -sheet) family and the cis-defensin superfamily is composed of the α -, β -, θ -, and big defensin families. In contrast to the CS $\alpha\beta$ family (Figure 7), the CXC motif (X denoting any amino acid residue) of BCR4 is in the first β -strand, leading to a different cysteine bonding pattern, i.e., the one corresponding to the β -defensin fold (the second family of defensin well-known as antimicrobial peptides) [10,24]. The fold of BCR4 is therefore a new type of defensin peptide, suggesting that BCRs constitute a new class of antimicrobial cysteine-rich protein. We here hypothesize that the defensins consist of at least three independent and convergent superfamilies: cis-, trans- and BCR defensins. BCR4 is the first member of this new defensin family with the cysteine bonding pattern of β -defensin and the structural motif of the CS $\alpha\beta$ family (Figure 7).

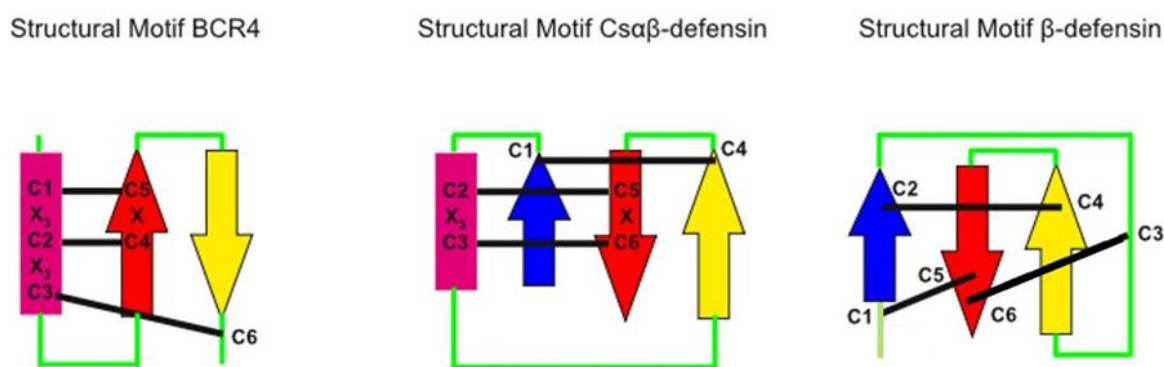


Figure 7. Disulfide connectivities in BCR4 peptide and in CS $\alpha\beta$ and β -defensin groups. α -helices are indicated by rectangle, β -strands are represented by arrows, and disulfides in black lines.

BCRs are orphan genes exclusively expressed in aphid bacteriocytes. This gene family has probably evolved in aphid lineages to ensure several functions related to endosymbiosis, including bacteriocyte homeostasis and endosymbiont control. Studies on the cereal weevil *Sitophilus* have shown that the coleopterucin A (CoIA) AMP selectively targets endosymbionts within bacteriocytes and impairs bacterial cytokinesis, thereby regulating bacterial cell division and preventing bacterial exit from weevil bacteriocytes [21,44]. Similar results were obtained in the *Rhizobium*-legumes symbiosis, where nodule-specific secreted peptides called NCRs have been shown to target bacteria and induce an irreversible elongation of bacteria, rendering them metabolically active but unable to multiply in vitro outside plant nodules [45]. In the actinorhizal plant *Alnus glutinosa*, nodules express defensin-like peptides, including the Ag5 peptide that targets symbiont vesicles and increases the permeability of vesicle membranes [46]. Small peptides are becoming emerging molecules that

target a variety of functions in host–symbiont interaction, including symbiont control and physiology [47]. Such a convergent evolution pattern is especially interesting with respect to its applications.

Overall, the susceptibility of aphids to BCR peptides may lead to the development of effective strategies for controlling such sap sucking pests. The exploration contained in this work may end up in a new protein family targeted to the specific control of aphids, which are some of the most important pests in global agriculture.

4. Materials and Methods

4.1. Peptide Synthesis

BCR4 was synthesized through native chemical ligation (NCL) of two peptide segments, followed by oxidative folding to form the three disulfide bridges, using previously described protocols [28–33]. Briefly, an N-terminal cysteinyl peptide segment (sequence: H-²¹CAVVCNYTSRPCYCVEAAKERDQWFPYCY⁵⁰D-OH) and a C-terminal crypto-thioester segment (sequence: H-¹DFDPTEFKGPFPTIEICSK²⁰Y-(Hnb)C(*St*Bu)G-NH₂) were synthesized on a Prelude peptide synthesizer (Protein Technologies, Tucson, AZ, USA) using standard Fmoc/*t*Bu chemistry at a 25 μmol scale and starting from a Tentagel resin equipped with a Rink's amide linker. The *N*-2-hydroxy-5-nitrobenzyl (Hnb) group was automatically introduced through resin reductive amination. Both peptide segments were purified by C₁₈ reverse phase RP-HPLC on a Nucleosil C₁₈ column (300 Å, 250 × 10 mm) (Macherey-Nagel, Düren, Germany), maintained at 25 °C. Solvent A was 0.1% trifluoroacetic acid (TFA) in water and B was 0.1% TFA in acetonitrile. A linear gradient from 20% B to 55% B was applied for 30 min at a flow rate of 3 mL/min and yielded the pure reduced form of BCR4 through collection of the peak identified as the target compound. The purified peptides were then chemoselectively coupled under standard NCL conditions (2 mM peptide segments, 50 mM tris-carboxyethylphosphine, 200 mM 4-mercaptophenylacetic acid, 6 M guanidinium chloride, 200 mM phosphate buffer, pH 6.5, 37 °C, 24 h) [28,29] and purified as before to retrieve the pure reduced form of BCR4 (sequence: H-¹DFDPTEFKGPFPTIEICSKYCAVVCNYTSRPCYCVEAAKERDQWFPYCY⁵⁰D-OH). The purified compound was further analyzed by ESI-HRMS on a Bruker maXisTM 22 ultra-high-resolution Q-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany), in positive mode and a [M+H]⁺ *m/z* ratio of 5897.5883 was found (theoretical monoisotopic *m/z* calculated for C₂₆₆H₃₇₉N₆₂O₇₉S₆: 5897.5869). Subsequently, the pure reduced form of BCR4 was oxidatively folded in vitro (10 μM peptide concentration, 0.1 mM oxidized glutathione, 1 mM glutathione, 1 mM EDTA, 100 mM TRIS, pH 8.5, 20 °C, 48 h) [34,48] before purification to homogeneity by C₁₈ RP-HPLC. Analysis of the purified folded product via ESI-HRMS analysis gave results consistent with the formation of three disulfide bridges ([M+H]⁺ *m/z* obtained: 5891.5437; theoretical monoisotopic *m/z* calculated for C₂₆₆H₃₇₃N₆₂O₇₉S₆: 5891.5400).

4.2. Antimicrobial Assays

E. coli and *M. luteus* were grown in Lysogeny Broth (LB) and Terrific Broth (TB) medium (Sigma-Aldrich, Saint-Louis, MO, USA), respectively. Antimicrobial assays were performed in sterile 96-well plates with a final volume of 100 μL per well, composed of 50 μL of culture and 50 μL of serially diluted peptides (5–80 μM). *E. coli* and *M. luteus* were added at an OD₆₀₀ of 0.006 and 0.05, respectively. Plates were incubated at 30 °C for 24 h and growth was measured at 600 nm using a Power wave XS-Biotek plate reader (Biotek Instrument, Colmar, France). The lowest concentration of BCR4 peptide showing complete inhibition was taken as minimal inhibitory concentration (MIC). All analyses were performed in triplicate, with the results expressed as mean ± standard deviation of mean (SEM).

4.3. Insects and Insect Assays

The aphid clone used was *A. pisum* LL01, a long-established alfalfa-collected clone containing only the primary endosymbiont *B. aphidicola*. Aphids were maintained on young broad bean plants (*Vicia faba* L. cv. Aguadulce) at 21 °C, with a photoperiod of 16 h light–8 h dark to obtain strictly parthenogenetic aphid matriline, that were reared and synchronized as previously described [20].

For toxicity analyses, three groups of ten 1st instar nymphs (aged between 0 and 24 h) were collected and placed in *ad hoc* feeding chambers containing an AP3 artificial diet [49] supplemented with different doses (5 µM to 80 µM) of solubilized BCR4 peptide. Toxicity was evaluated by scoring survival daily over the whole nymphal life of the pea aphid (7 days) [50]. Growth was also measured by weighing adult aphids on a Mettler AE163 analytical microbalance (Mettler Toledo, Columbus, OH, USA) at the closest 10 µg, following the protocol described previously [51,52].

Aphid mortality data for all BCR4 concentrations were analyzed separately in a parametric survival analysis with a log-normal fit. Aphid weights were analyzed by ANOVA followed by Tukey–Kramer HSD test for comparing multiple means. All analyses were performed with JMP software version 11 (SAS Institute Cary USA, MacOS version).

4.4. Sequence Analysis and Phylogeny

Homologous BCR protein sequences were retrieved using a combination of TBLASTN and BLASTP [53,54] against the aphid genomes available from (i) AphidBase [36], (ii) the NCBI genome database and (iii) the whole NCBI non-redundant protein, nucleotide and EST databases (with *A. pisum* BCR as blast seeds; see Supplementary Materials Table S2 for a complete list of BCR protein sequences used for the phylogenetic analysis). BCR protein sequences were subjected to multiple sequence alignments using the MUSCLE program [55]. Subsequently, a phylogenetic tree was constructed, using the PhyML method [37] implemented in the Seaview software (v5.0.4) [37,56] (LG model with 4 rate classes), and the reliability of each branch was evaluated using the bootstrap method, with 1000 replicates. Poorly supported branches (<50%) were collapsed using TreeCollapseCL4 [38]. Graphical representation and editing of the phylogenetic tree were performed with FigTree (v1.4.3) [57].

4.5. NMR Experiments

Prior to NMR analysis, the synthesized BCR4 peptides were dissolved in H₂O:D₂O (9:1 ratio) at a concentration of 0.6 mM and pH was adjusted to 4.8. Then, 2D ¹H NOESY, 2D ¹H TOCSY, a sofast-HMQC [39] (¹⁵N natural abundance) and a ¹³C-HSQC (¹³C natural abundance) were performed at 298 K on an Avance III HD BRUKER 700 MHz spectrometer equipped with a cryoprobe. ¹H chemical shifts were referenced to the water signal (4.77 ppm at 298 K). NMR data were processed using the Topspin software version 3.2TM (Bruker, Billerica, MA, USA) and analyzed with CCPNMR version 2.2.2 [58].

4.6. Structure Calculations

Structures were calculated using the Crystallography and NMR System [59,60] through the automatic assignment software ARIA2 version 2.3 [61] with NOE-derived distances, hydrogen bonds (in accordance with the observation of typical long- or medium-distance NOE cross peak network for β-sheets and α-helices respectively—HN/HN, HN/Hα, Hα/Hα), backbone dihedral angle restraints (determined with the DANGLE program [62]) and three ambiguous disulfide bridges. The ARIA2 protocol, with default parameters used, simulated annealing with torsion angle and Cartesian space dynamics. The iterative process was repeated until the assignment of the NOE cross peaks was complete. The last run for BCR4 was performed with 1000 initial structures and 200 structures were refined in water. Twenty structures were selected on the basis of total energies and restraint violation statistics, to represent the structure of BCR4 in solution. The quality of the final

structures was evaluated using PROCHECK-NMR [40] and PROMOTIF [63]. The figures were prepared with PYMOL [64].

4.7. Structure Relationship of BCR4 Structure

Comparison of the BCR4 3D structure to the PDB database was performed using the DALI server [41].

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232012480/s1>, References [65,66] are cited in the supplementary materials.

Author Contributions: Conceptualization, K.L., N.P., F.P., A.F.D., V.A., A.H., F.C. and P.d.S.; methodology, K.L., N.P., F.P., A.F.D., V.A., A.H., F.C. and P.d.S.; validation, K.L., N.P., F.P., A.F.D., V.A., F.C. and P.d.S.; formal analysis, K.L., N.P., F.P., A.F.D., V.A., F.C. and P.d.S.; investigation, K.L., N.P., F.P., H.T., C.S., I.R., M.R.L., A.F.D., V.A., A.H., F.C. and P.d.S.; resources, K.G. and G.D.; data curation, K.L., N.P., F.P., M.R.L., A.F.D., V.A., F.C. and P.d.S.; writing—original draft preparation, K.L., N.P., F.P., A.F.D., V.A. and P.d.S.; writing—review and editing, K.L., N.P., F.P., M.R.L., A.F.D., V.A., A.H., F.C. and P.d.S.; visualization, K.L., N.P., F.P., A.F.D., V.A., A.H., F.C. and P.d.S.; supervision, V.A., A.F.D., F.C. and P.d.S.; project administration, V.A. and P.d.S.; funding acquisition, N.P., F.P., V.A., A.F.D., F.C. and P.d.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by INRAE (Institut National de Recherche pour l’Agriculture, l’Alimentation et l’Environnement), INSA Lyon (Institut National des Sciences Appliquées Lyon) and the French ANR-119-CE11-0004-01 (Biofamily) program grant.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are contained within the manuscript and its supporting information.

Acknowledgments: The authors thank the NMR sub-platform of the MO2VING facility (Orléans, France). We thank Philippe Marceau and Jean-Baptiste Madinier for assistance in solid phase peptide synthesis and Patrick Callaerts for careful reading of the English language.

Conflicts of Interest: The authors declare no conflict of interests.

References

- Bradshaw, C.J.A.; Leroy, B.; Bellard, C.; Roiz, D.; Albert, C.; Fournier, A.; Barbet-Massin, M.; Salles, J.-M.; Simard, F.; Courchamp, F. Massive yet grossly underestimated global costs of invasive insects. *Nat. Commun.* **2016**, *7*, 12986. [CrossRef] [PubMed]
- Diagne, C.; Leroy, B.; Gozlan, R.E.; Vaissiere, A.C.; Assailly, C.; Nuninger, L.; Roiz, D.; Jourdain, F.; Jaric, I.; Courchamp, F. InvaCost, a public database of the economic costs of biological invasions worldwide. *Sci. Data* **2020**, *7*, 277. [CrossRef] [PubMed]
- Oerke, E.C.; Dehne, H.W. Safeguarding production—Losses in major crops and the role of crop protection. *Crop Prot.* **2004**, *23*, 275–285. [CrossRef]
- Calevro, F.; Tagu, D.; Callaerts, P. *Acyrtosiphon pisum*. *Trends Genet. TIG* **2019**, *35*, 781–782. [CrossRef]
- Dedryver, C.A.; Le Ralec, A.; Fabre, F. The conflicting relationships between aphids and men: A review of aphid damage and control strategies. *Comptes. Rendus. Biol.* **2010**, *333*, 539–553. [CrossRef] [PubMed]
- Goulson, D. Call to restrict neonicotinoids. *Science* **2018**, *360*, 973. [CrossRef] [PubMed]
- Huang, Y.H.; Du, Q.; Craik, D.J. Cyclotides: Disulfide-Rich peptide toxins in plants. *Toxicon* **2019**, *172*, 33–44. [CrossRef] [PubMed]
- King, G.F. Tying pest insects in knots: The deployment of spider-venom-derived knottins as bioinsecticides. *Pest. Manag. Sci.* **2019**, *75*, 2437–2445. [CrossRef] [PubMed]
- Gressent, F.; Da Silva, P.; Eyraud, V.; Karaki, L.; Royer, C. Pea Albumin 1 subunit b (PA1b), a promising bioinsecticide of plant origin. *Toxins* **2011**, *3*, 1502–1517. [CrossRef]
- Shafee, T.M.; Lay, F.T.; Phan, T.K.; Anderson, M.A.; Hulett, M.D. Convergent evolution of defensin sequence, structure and function. *Cell. Mol. Life Sci.* **2017**, *74*, 663–682. [CrossRef]
- Gonzalez-Castro, R.; Gomez-Lim, M.A.; Plisson, F. Cysteine-Rich Peptides: Hyperstable Scaffolds for Protein Engineering. *Chembiochem. A Eur. J. Chem. Biol.* **2021**, *22*, 961–973. [CrossRef]
- Rahioui, I.; Eyraud, V.; Karaki, L.; Sasse, F.; Carre-Pierrat, M.; Qin, A.; Zheng, M.H.; Toepfer, S.; Sivignon, C.; Royer, C.; et al. Host range of the potential biopesticide Pea Albumin 1b (PA1b) is limited to insects. *Toxicon* **2014**, *89*, 67–76. [CrossRef] [PubMed]

13. Bell, J.; Sukiran, N.A.; Walsh, S.; Fitches, E.C. The insecticidal activity of recombinant nemertide toxin alpha-1 from *Lineus longissimus* towards pests and beneficial species. *Toxicon* **2021**, *197*, 79–86. [[CrossRef](#)] [[PubMed](#)]
14. Delobel, B.; Grenier, A.M.; Gueguen, J.; Ferrasson, E.; Mbaiguinam, M. Utilisation d'un polypeptide dérivé d'une albumine PA1b de légumineuse comme insecticide (Brevet-98/05877). *French Patent* **1998**, *98*, 11.
15. Shigenobu, S.; Stern, D.L. Aphids evolved novel secreted proteins for symbiosis with bacterial endosymbiont. *Proc. Biol. Sci.* **2013**, *280*, 20121952. [[CrossRef](#)] [[PubMed](#)]
16. Baumann, P. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* **2005**, *59*, 155–189. [[CrossRef](#)] [[PubMed](#)]
17. Douglas, A.E. Multiorganismal insects: Diversity and function of resident microorganisms. *Annu. Rev. Entomol.* **2015**, *60*, 17–34. [[CrossRef](#)] [[PubMed](#)]
18. Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp APS. *Nature* **2000**, *407*, 81–86. [[CrossRef](#)] [[PubMed](#)]
19. Brinza, L.; Viñuelas, J.; Cottret, L.; Calevro, F.; Rahbé, Y.; Febvay, G.; Duport, G.; Colella, S.; Rabatel, A.; Gautier, C.; et al. Systemic analysis of the symbiotic function of Buchnera aphidicola, the primary endosymbiont of the pea aphid *Acyrtosiphon pisum*. *Comptes. Rendus. Biol.* **2009**, *332*, 1034–1049. [[CrossRef](#)] [[PubMed](#)]
20. Simonet, P.; Gaget, K.; Balmand, S.; Ribeiro Lopes, M.; Parisot, N.; Buhler, K.; Duport, G.; Vulsteke, V.; Febvay, G.; Heddi, A.; et al. Bacteriocyte cell death in the pea aphid/Buchnera symbiotic system. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E1819–E1828. [[CrossRef](#)] [[PubMed](#)]
21. Login, F.H.; Balmand, S.; Vallier, A.; Vincent-Monegat, C.; Vignerot, A.; Weiss-Gayet, M.; Rochat, D.; Heddi, A. Antimicrobial peptides keep insect endosymbionts under control. *Science* **2011**, *334*, 362–365. [[CrossRef](#)] [[PubMed](#)]
22. Uchi, N.; Fukudome, M.; Nozaki, N.; Suzuki, M.; Osuki, K.I.; Shigenobu, S.; Uchiumi, T. Antimicrobial Activities of Cysteine-rich Peptides Specific to Bacteriocytes of the Pea Aphid *Acyrtosiphon pisum*. *Microbes Environ.* **2019**, *34*, 155–160. [[CrossRef](#)]
23. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gislason, M.H.; Pihl, S.I.; Tsirigos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **2022**, *40*, 1023–1025. [[CrossRef](#)]
24. Shafee, T.M.; Lay, F.T.; Hulett, M.D.; Anderson, M.A. The Defensins Consist of Two Independent, Convergent Protein Superfamilies. *Mol. Biol. Evol.* **2016**, *33*, 2345–2356. [[CrossRef](#)]
25. Huygens, C.; Ribeiro Lopes, M.; Gaget, K.; Duport, G.; Peignier, S.; De Groef, S.; Parisot, N.; Calevro, F.; Callaerts, P. Evolutionary diversification of insulin-related peptides (IRPs) in aphids and spatiotemporal distribution in *Acyrtosiphon pisum*. *Insect Biochem. Mol. Biol.* **2022**, *141*, 103670. [[CrossRef](#)]
26. Ribeiro Lopes, M.; Parisot, N.; Gaget, K.; Huygens, C.; Peignier, S.; Duport, G.; Orleans, J.; Charles, H.; Baatsen, P.; Jousset, E.; et al. Evolutionary novelty in the apoptotic pathway of aphids. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 32545–32556. [[CrossRef](#)] [[PubMed](#)]
27. Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779. [[CrossRef](#)] [[PubMed](#)]
28. Lelievre, D.; Terrier, V.P.; Delmas, A.F.; Aucagne, V. Native Chemical Ligation Strategy to Overcome Side Reactions during Fmoc-Based Synthesis of C-Terminal Cysteine-Containing Peptides. *Org. Lett.* **2016**, *18*, 920–923. [[CrossRef](#)] [[PubMed](#)]
29. Terrier, V.P.; Adihou, H.; Arnould, M.; Delmas, A.F.; Aucagne, V. A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters. *Chem. Sci.* **2016**, *7*, 339–345. [[CrossRef](#)]
30. Terrier, V.P.; Delmas, A.F.; Aucagne, V. Efficient synthesis of cysteine-rich cyclic peptides through intramolecular native chemical ligation of N-Hnb-Cys peptide crypto-thioesters. *Org. Biomol. Chem.* **2017**, *15*, 316–319. [[CrossRef](#)]
31. Abboud, S.A.; Amoura, M.; Madinier, J.B.; Renoux, B.; Papot, S.; Piller, V.; Aucagne, V. Enzyme-Cleavable Linkers for Protein Chemical Synthesis through Solid-Phase Ligations. *Angew. Chem.* **2021**, *60*, 18612–18618. [[CrossRef](#)]
32. Abboud, S.A.; Aucagne, V. An optimized protocol for the synthesis of N-2-hydroxybenzyl-cysteine peptide crypto-thioesters. *Org. Biomol. Chem.* **2020**, *18*, 8199–8208. [[CrossRef](#)] [[PubMed](#)]
33. Abboud, S.A.; Cisse, E.H.; Doudeau, M.; Benedetti, H.; Aucagne, V. A straightforward methodology to overcome solubility challenges for N-terminal cysteinyl peptide segments used in native chemical ligation. *Chem. Sci.* **2021**, *12*, 3194–3201. [[CrossRef](#)] [[PubMed](#)]
34. Derache, C.; Meudal, H.; Aucagne, V.; Mark, K.J.; Cadene, M.; Delmas, A.F.; Lalmanach, A.C.; Landon, C. Initial insights into structure-activity relationships of avian defensins. *J. Biol. Chem.* **2012**, *287*, 7746–7755. [[CrossRef](#)] [[PubMed](#)]
35. Martinez, G.; Hograindleur, J.P.; Voisin, S.; Abi Nahed, R.; Abd El Aziz, T.M.; Escoffier, J.; Bessonnat, J.; Fovet, C.M.; De Waard, M.; Hennebicq, S.; et al. Spermaurin, an La1-like peptide from the venom of the scorpion *Scorpio Maurus palmatus*, improves sperm motility and fertilization in different mammalian species. *Mol. Hum. Reprod.* **2016**, *23*, 116–131. [[CrossRef](#)]
36. Legeai, F.; Shigenobu, S.; Gauthier, J.P.; Colbourne, J.; Rispe, C.; Collin, O.; Richards, S.; Wilson, A.C.; Murphy, T.; Tagu, D. AphidBase: A centralized bioinformatic resource for annotation of the pea aphid genome. *Insect Mol. Biol.* **2010**, *19*, 5–12. [[CrossRef](#)]
37. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 390–401. [[CrossRef](#)]
38. Hodcroft, E. TreeCollapseCL 4. Available online: <http://emmahodcroft.com/treecollapsecl.html> (accessed on 22 July 2022).

39. Schanda, P.; Kupce, E.; Brutscher, B. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR* **2005**, *33*, 199–211. [[CrossRef](#)]
40. Laskowski, R.A.; Rullmann, J.A.C.; MacArthur, M.W.; Kaptein, R.; Thornton, J.M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **1996**, *8*, 477–486. [[CrossRef](#)]
41. Holm, L. Dali server: Structural unification of protein families. *Nucleic Acids Res.* **2022**, *50*, W210–W215. [[CrossRef](#)]
42. M'Barek, S.; Chagot, B.; Andreotti, N.; Visan, V.; Mansuelle, P.; Grissmer, S.; Marrakchi, M.; El Ayeb, M.; Sampieri, F.; Darbon, H.; et al. Increasing the molecular contacts between maurotoxin and Kv1.2 channel augments ligand affinity. *Proteins* **2005**, *60*, 401–411. [[CrossRef](#)] [[PubMed](#)]
43. Gressent, F.; Duport, G.; Rahioui, I.; Pauchet, Y.; Bolland, P.; Specty, O.; Rahbe, Y. Biological activity and binding site characteristics of the PA1b entomotoxin on insects from different orders. *J. Insect Sci.* **2007**, *7*, 12. [[CrossRef](#)]
44. Login, F.H.; Heddi, A. Insect immune system maintains long-term resident bacteria through a local response. *J. Insect Physiol.* **2013**, *59*, 232–239. [[CrossRef](#)] [[PubMed](#)]
45. Van de Velde, W.; Zehirov, G.; Szatmari, A.; Debreczeny, M.; Ishihara, H.; Kevei, Z.; Farkas, A.; Mikulass, K.; Nagy, A.; Tiricz, H.; et al. Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **2010**, *327*, 1122–1126. [[CrossRef](#)]
46. Carro, L.; Pujic, P.; Alloisio, N.; Fournier, P.; Boubakri, H.; Hay, A.E.; Poly, F.; Francois, P.; Hocher, V.; Mergaert, P.; et al. Alnus peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing Frankia. *ISME J.* **2015**, *9*, 1723–1733. [[CrossRef](#)]
47. Mergaert, P.; Kikuchi, Y.; Shigenobu, S.; Nowack, E.C.M. Metabolic Integration of Bacterial Endosymbionts through Antimicrobial Peptides. *Trends Microbiol.* **2017**, *25*, 703–712. [[CrossRef](#)] [[PubMed](#)]
48. Silva, P.; Strzepa, A.; Jouvensal, L.; Rahioui, I.; Gressent, F.; Delmas, A.F. A folded and functional synthetic PA1b, an interlocked entomotoxic miniprotein. *Biopolymers* **2009**, *92*, 436–444. [[CrossRef](#)] [[PubMed](#)]
49. Febvay, G.; Delobel, B.; Rahbe, Y. Influence of the Amino-Acid Balance on the Improvement of an Artificial Diet for a Biotype of *Acyrtosiphon-Pisum* (Homoptera, Aphididae). *Can. J. Zool.* **1988**, *66*, 2449–2453. [[CrossRef](#)]
50. Simonet, P.; Gaget, K.; Parisot, N.; Duport, G.; Rey, M.; Febvay, G.; Charles, H.; Callaerts, P.; Colella, S.; Calevro, F. Disruption of phenylalanine hydroxylase reduces adult lifespan and fecundity, and impairs embryonic development in parthenogenetic pea aphids. *Sci. Rep.* **2016**, *6*, 34321. [[CrossRef](#)]
51. Loth, K.; Costechareyre, D.; Effantin, G.; Rahbe, Y.; Condemine, G.; Landon, C.; da Silva, P. New Cyt-like delta-endotoxins from *Dickeya dadantii*: Structure and aphicidal activity. *Sci. Rep.* **2015**, *5*, 8791. [[CrossRef](#)]
52. Rahbe, Y.; Febvay, G. Protein Toxicity to Aphids—An Invitro Test on *Acyrtosiphon-Pisum*. *Entomol. Exp. Appl.* **1993**, *67*, 149–160. [[CrossRef](#)]
53. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.H.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)] [[PubMed](#)]
54. Boratyn, G.M.; Schaffer, A.A.; Agarwala, R.; Altschul, S.F.; Lipman, D.J.; Madden, T.L. Domain enhanced lookup time accelerated BLAST. *Biol. Direct.* **2012**, *7*, 727–741. [[CrossRef](#)] [[PubMed](#)]
55. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
56. Gouy, M.; Guindon, S.; Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **2010**, *27*, 19854763. [[CrossRef](#)]
57. Drummond, A.J.; Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **2007**, *7*, 214. [[CrossRef](#)]
58. Vranken, W.F.; Boucher, W.; Stevens, T.J.; Fogh, R.H.; Pajon, A.; Llinas, P.; Ulrich, E.L.; Markley, J.L.; Ionides, J.; Laue, E.D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005**, *59*, 687–696. [[CrossRef](#)]
59. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2007**, *2*, 2728–2733. [[CrossRef](#)]
60. Brunger, A.T.; Adams, P.D.; Clore, G.M.; DeLano, W.L.; Gros, P.; Grosse-Kunstleve, R.W.; Jiang, J.S.; Kuszewski, J.; Nilges, M.; Pannu, N.S.; et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D* **1998**, *54*, 905–921. [[CrossRef](#)]
61. Rieping, W.; Habeck, M.; Bardiaux, B.; Bernard, A.; Malliavin, T.E.; Nilges, M. ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **2007**, *23*, 381–382. [[CrossRef](#)]
62. Cheung, M.S.; Maguire, M.L.; Stevens, T.J.; Broadhurst, R.W. DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J. Magn. Reson.* **2010**, *202*, 223–233. [[CrossRef](#)] [[PubMed](#)]
63. Hutchinson, E.G.; Thornton, J.M. PROMOTIF—A program to identify and analyze structural motifs in proteins. *Protein Sci.* **1996**, *5*, 212–220. [[CrossRef](#)] [[PubMed](#)]
64. De Lano, W.L. *Pymol*; DeLano Scientific: South San Francisco, CA, USA, 2002.
65. Piszkievicz, D.; Landon, M.; Smith, E.L. Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations. *Biochem. Biophys. Res. Commun.* **1970**, *40*, 1173–1178. [[CrossRef](#)]
66. Segalas, I.; Thai, R.; Menez, R.; Vita, C. A particularly labile Asp-Pro bond in the green mamba muscarinic toxin Mtx2—Effect of protein conformation on the rate of cleavage. *FEBS Lett.* **1995**, *371*, 171–175. [[CrossRef](#)]